Equilibrium occurrence and unobserved competition*

Alexandra Belova†   Philippe Gagnepain‡   Stéphane Gauthier§

April 8, 2021

Abstract

We build an index for the likelihood that an equilibrium obtains based on the postulate that it is more likely to occur when it is the only rationalizable outcome. The market should be in equilibrium if the index is lower than 1, conditional on the ability of the researcher to observe the full set of products/services proposed by competing firms. If not, the relevant threshold for the index should be adjusted below 1. Assessing the plausibility of the occurrence of an equilibrium can thus be used to delineate empirically the scope of the relevant market, which is a well-known concern for competition authorities. We illustrate our framework with data from the U.S. air transportation industry.

JEL classification numbers: C51, D21, L13, L40, L93.

Keywords: Rationalizability, Nash equilibrium, relevant market, airline industry.

1 Introduction

Economic analysis often refers to the Nash equilibrium concept at the moment of describing strategic interactions between agents. Then every agent is assumed to be able to forecast correctly the behavior of the other agents. The recent literature in industrial organization shows that this assumption is more likely to hold in a stable environment, e.g., if firms operate in markets that are geographically close to the market of their competitors (Aguirregabiria and Magesan (2020)) or if they can accumulate experience and gradually learn how the others behave. However, this literature also suggests that beliefs may loose accuracy in a more disturbed environment subject to frequent entries and exits of participants. In this situation the equilibrium reference becomes less relevant (Doraszelski et al. (2018)).[1]

Although agents may not succeed to forecast the behavior of their competitors, they are usually considered as rational in the sense of maximizing their objective given their expectations on what the others do. An important insight from the concept of rationalizability is that rationality, even pushed at a high degree of sophistication, may be not enough to reach a Nash equilibrium. A player is rational at level $k = 1$ if she plays a best-response given her beliefs; at level $k = 2$, she is rational and believes that the other players are rational; at level $k = 3$, she also believes that the others believe that the others are rational. Reproducing this process of higher-order beliefs about rationality eventually yields the set of rationalizable outcomes, which always includes but does not necessarily reduce to Nash equilibria (Bernheim (1984), Pearce (1984), Moulin (1979)).

In this article, we build an index for the likelihood that an equilibrium obtains, based on the postulate that an equilibrium is more likely to occur when it is the only rationalizable outcome (Guesnerie (1992)). The index uses the characterization of rationalizable outcomes as being those surviving an iterative process of elimination of dominated strategies. The process eliminates every outcome close to an equilibrium, but not the equilibrium itself, if it is locally contracting around this equilibrium. Contraction obtains when the spectral radius of the Jacobian matrix governing locally the process is less than 1. We derive a proxy for the spectral radius and use it as an index for the likelihood that the observed market will reach an equilibrium: the theory predicts that the market should be in equilibrium if the index is lower than 1.

To test for this prediction, we use data on the U.S. air transportation industry.[2] We

---

[1]Recent discussions of the plausibility of the Nash equilibrium in Industrial Organization frameworks also include Aguirregabiria and Xie (2020), Aguirregabiria and Jeon (2020), Kashaev (2016) or Kosenkova (2017). Similar concerns also appear in macroeconomic models following Woodford (2003): see for instance Farhi and Werning (2017), Angeletos and Lian (2018), Giacomini et al. (2020) or Landier et al. (2020). Most of this literature appeals to the level-k thinking developed by Crawford et al. (2013) and rationalizability-based arguments. Experimental evidence and direct elicitation of beliefs are provided in Alaoui and Penta (2016), Aradillas-Lopez and Tamer (2008), Coibion and Gorodnichenko (2015), Coibion et al. (2018) or Kneeland (2015).

[2]Our model generalizes Desgranges and Gauthier (2016) to the case of heterogeneous production facilities, which is known to be a crucial ingredient in the airline industry.

merge several Department of Transportation databases for the period 2003:2016 to estimate cost and demand functions for a large number of air routes. This information allows us to recover the value of the Nash equilibrium index on every route as well as the hypothetical volume of transported passengers by each airline in the equilibrium.

A difficulty of our exercise relies on the fact that, in order to compute accurately the Nash equilibrium on each route, we would need to be able to specify the strategic game between airlines correctly and choose the right functional form of costs and demand. The airline industry is quite challenging on this front as many ingredients such as multimarket contacts, collusion, or dynamic choices are potentially relevant and might yield misspecification issues on specific routes if they are omitted or misinterpreted. Our view here is not to intend to provide the best possible representation of the industry but to focus instead on the interpretability and tractability of our framework. In this sense, the reader should see our contribution as an exploration of rationalizability conditional on the assumptions of our model which is clearly a simplified representation of the data reality.

With our setup, we are able to compute the difference between the actual observed production levels and those that would prevail in the equilibrium. A first contribution of the present article is to show that our index is a reliable indicator of this difference: a low value of the index is associated with a situation where the actual strategies of the airlines stand close to the Nash equilibrium of our model. The index thus could be viewed as a competition authority tool to assess the status of counterfactual analysis referring to Nash equilibria. The index is below 1 in approximately 90% of the markets in our dataset, and we find that a 10% increase of the index is associated with a 7% increase in the difference between observed and Nash quantities. Descriptive statistics show that the Nash equilibrium is less likely to occur on short distance and intense traffic routes linking populated cities. Low cost companies also complicate the convergence toward the equilibrium whereas greater concentration tends to yield a lower value of the index.

On several routes where the index is below 1, the empirical difference between the observed and the Nash quantities is still significant, though lower than in routes where the index exceeds 1. This does not accord well with a strict interpretation of the theoretical model as the difference should then be 0. We show how to reconcile theory and data by accounting for the size of the relevant market which consists of the full set of services or goods supplied by the competing firms under scrutiny (Davis and Garcés (2009)). We argue that the relevant threshold for the Nash equilibrium index should be adjusted further down below 1, the higher the share of missing products not identified in the set of potential substitutes on a given route. As the threshold goes down, the gap between the observed and the theoretical Nash quantities is reduced, which is another signal that our model performs reasonably well. As a by-product, a second contribution of our paper is therefore to assess empirically the scope of the relevant market in a given industry, which is a well known important issue for competition authorities. As an illustration, we propose a natural experiment, the Wright amendment, which restricted flights from the Dallas Love airport in order to promote the development of the Dallas/Fort Worth airport (Ciliberto et Tamer (2009)). The estimated relevant threshold in routes from Dallas/Fort Worth is sharply

reduced after the repeal of the amendment, suggesting that the relevant market indeed expands over the period.

Finally, we discuss two features of our model. First we wonder whether rationalizability-based inferences are the only factor explaining the departures of actual from equilibrium strategies. Adaptive learning could also contribute to explain convergence toward the equilibrium, as airlines may use past observations to form their anticipations on the decisions that their competitors will take. We find that adaptive learning is relevant on top of rationalizability. Second we are interested in the welfare impact of Nash departures. Empirical studies on firms' entry and exit usually assume that the market is in equilibrium both before and after entry/exit and so refer to the corresponding prices and quantities in each period. Here, we view entry/exit as potential perturbations that may lead to a multiplicity of rationalizable outcomes. In the case of the New York-Tampa route, we find that firms schedule too many seats following entry: consumer surplus is higher than what one would get in equilibrium.

The rest of our article is organized as follows. We build the theoretical model applied to the airline industry in Section 2 and Section 3 discusses the data and the empirical model. Section 4 reports estimation results on the index that governs the plausibility of the equilibrium. It shows that the index is correlated with observed departures from Nash behavior. Section 5 extends our analysis to the identification of the relevant market. In section 6 we discuss adaptive learning and illustrate the welfare impact of nonequilibrium behavior. Finally section 7 concludes.

2   Theoretical benchmark

2.1   General framework

In the airline industry a market is defined as the set of air services offered by different carriers in a route linking a pair of origin and destination airports or cities. The airlines compete in the route for carrying freight and passengers. They all face the same demand function for transportation services but they typically differ according to their technological and organizational characteristics summarized by their cost structure. Some airlines may use a few large capacity aircraft to spread the cost of booking airport slots, boarding passengers and operating flights on few departures, whereas others rely on lower capacity aircraft but schedule more departures. Such choices result from medium-run intermittent contractual negotiations between airlines and airports and long-run airlines capacity investment policies. Over a shorter horizon, the fact that Delta Airlines in the route linking Chicago and Atlanta allocates a A320 Airbus to a booked slot at O'Hara airport from 8h00 to 8h30am every Monday is essentially given.

Over this shorter horizon, airlines instead rely on yield management to control prices and quantities of transported passengers given the available aircraft allocated to the route (Borenstein and Rose, 1994). Following Ciliberto and Tamer (2009), we allow for a firm f specific cost function $c_{af}(q)$ for transporting q passengers using a type a aircraft that varies

with aircraft type as well as additional firm characteristics, e.g., negotiated input (fuel and employees) prices. We assume that the cost function is twice differentiable, increasing and convex with the number of transported passengers. A polar but plausible case involves a low marginal cost for transporting one additional passenger except when the total number of passengers approaches the aircraft capacity.

Let $\mathcal{A}_f$ be the given set of aircraft used by airlines f in the route and n_{af} be the given number of flights operated by the airlines using type a aircraft. The total cost of firm f for transporting q_f passengers is

$$C_f(q_f) = \min_{(q_{af})} \left\{ \sum_{a \in \mathcal{A}_f} n_{af} c_{af}(q_{af}) \mid \sum_{a \in \mathcal{A}_f} n_{af} q_{af} \geq q_f \right\}, \tag{1}$$

which is also an increasing and convex $\mathcal{C}^2$-function with the number q_f of passengers. As is well known from Kreps and Scheinkman (1983), a Cournot Nash equilibrium obtains as the outcome of a two-stage game where firms first choose their capacity and then compete over prices. This seems to fit well the case of the airline industry where firms have first to decide on airport slots and quantity of aircraft-seats available.[3] Assuming a Cournot-Nash behavior, firm f takes as given the number Q_{-f} of passengers transported by the other airlines and produces

$$q_f \in \arg \max_q P(Q_{-f} + q)q - C_f(q),$$

where $P(Q)$ is the inverse demand function, and $Q = Q_{-f} + q_f$ is the total number of passengers transported in the route. Assuming that the marginal revenue $P(Q_{-f} + q)q$ is decreasing in q, the best choice for firm f is $q_f = R_f(Q_{-f})$, where the best-response function $R_f(\cdot)$ is decreasing.

A Cournot-Nash equilibrium is a F component vector $\mathbf{q}^* = (q_f^*)$ such that $q_f^* = R_f(Q_{-f}^*)$ for all f, with Q_{-f}^* being the aggregate production of others in the equilibrium. Thus, in an equilibrium, every airline f is assumed to predict correctly the number Q_{-f}^* of passengers transported by its competitors.

Our article provides an empirical assessment of this assumption appealing to rationalizability. Rationalizable outcomes can be characterized by referring to an iterative process of elimination of dominated strategies (see, e.g., Osborne and Rubinstein, 1994). The process starts from the assumption of common knowledge among airlines that at some initial step $\tau = 0$ production satisfies

$$q_f \in [q_f^{\text{inf}}(0), q_f^{\text{sup}}(0)] \tag{2}$$

for every f, with $q_f^{\text{inf}}(0) \leq q_f^* \leq q_f^{\text{sup}}(0)$. Assuming the equilibrium amounts to require $q_f^{\text{inf}}(0) = q_f^* = q_f^{\text{sup}}(0)$ for all f. In the sequel we assume that firms restrict attention to a neighborhood of the equilibrium, with $(q_f^{\text{inf}}(0), q_f^{\text{sup}}(0))$ is close to, but different from (q_f^*, q_f^*).

---

[3]See also Brander and Zhang (1990), Brueckner (2002) or Basso (2008).

By individual rationality airlines only select volumes of transported passengers that are best-response to decisions consistent with (2). Airlines f thus chooses some production in a new interval $[q_f^{\inf}(1), q_f^{\sup}(1)]$ at step $\tau = 1$ of the process, with[4]

$$q_f^{\inf}(1) = R_f \left(\sum_{k \neq f} q_k^{\sup}(0) \right), \quad q_f^{\sup}(1) = R_f \left(\sum_{k \neq f} q_k^{\inf}(0) \right).$$

This reasoning applies to every airlines. Thus, in the immediate vicinity of an equilibrium, we have

$$q_f^{\inf}(1) - q_f^* = R_f'(Q_{-f}^*) \sum_{k \neq f} [q_k^{\sup}(0) - q_k^*] \tag{3}$$

and

$$q_f^{\sup}(1) - q_f^* = R_f'(Q_{-f}^*) \sum_{k \neq f} \left[q_k^{\inf}(0) - q_k^* \right] \tag{4}$$

for all f. By common knowledge of rationality and the slopes $R_f'(Q_{-f}^*)$ of the firms' best-response functions, one can iterate the argument. The iterative process of elimination of dominated strategies defined by (3) and (4) is governed by the $F \times F$ matrix $\mathbf{B}$ whose every entry in the f-th row equals $R_f'(Q_{-f}^*)$ except the diagonal entry (in the f-th column) which is 0. It has the Cournot-Nash equilibrium $\mathbf{q}^*$ is a fixed point.

Level-k thinking popularized by Crawford and Iriberri (2007) would iterate the process k times, for some finite number k. If iterated ad infinitum, the process eventually pins down the equilibrium if and only if the spectral radius of $\mathbf{B}$ is less than 1. If the Nash equilibrium is locally the only rationalizable outcome, one can argue that firms should eventually convince themselves that their competitors will behave according to Nash. In this case we say that the equilibrium is locally 'stable'. Otherwise, if the spectral radius is greater than 1, there are multiple rationalizable outcomes and the iterative process can no longer justify that firms eventually pin down their Nash productions. We say that the equilibrium then is 'unstable'. The following proposition gives a condition for local stability.

Proposition 1. *The Nash equilibrium is locally stable if and only if*

$$S(\mathbf{q}^*) = \sum_f \frac{R_f'(Q_{-f}^*)}{R_f'(Q_{-f}^*) - 1} < 1 \tag{5}$$

where

$$R_f'(Q_{-f}) = -\frac{P''(Q) q_f + P'(Q)}{P''(Q) q_f + 2P'(Q) - C_f''(q_f)}, \tag{6}$$

and

$$C_f''(q_f) \sum_{a \in \mathcal{A}_f} \frac{n_{af}}{c_{af}''(q_{af})} = 1.$$

---

[4]If $q_f^{\inf}(1) \leq q_f^{\inf}(0)$ then we set $q_f^{\inf}(1) = q_f^{\inf}(0)$. Similarly, if $q_f^{\sup}(1) \geq q_f^{\sup}(0)$ then we set $q_f^{\sup}(1) = q_f^{\sup}(0)$. The process therefore remains at (2) if both $q_f^{\inf}(1) \leq q_f^{\inf}(0)$ and $q_f^{\sup}(1) \geq q_f^{\sup}(0)$. No strategy is eliminated. The same procedure applies to every step $\tau \geq 1$ of elimination.

Proof. See Appendix A ■

Proposition 1 forms the basis of our empirical illustration by providing us with a simple criterion for the plausibility of the occurrence of the Nash equilibrium. It predicts that the spread between the theoretical Nash equilibrium productions $\mathbf{q}^*$ and the actual observed productions should be magnified if the 'stability index' $S(\mathbf{q}^*)$ defined in (5) is greater than a threshold of 1.

Condition (5) shows that local stability of the Nash equilibrium obtains if firms are not too sensitive to the production of others, i.e., $R'_f(Q^*_{-f})$ is close to 0, which accords with the early insights developed by Guesnerie (1992) for the competitive case. The intuition is that firms find it difficult to understand the behavior of others when others are sensitive to their beliefs.

2.2 Linear-quadratic specification

Proposition 1 considers a general formulation of demand and cost which complicates the implementation of comparative statics exercises. The direct effect of a perturbation on the stability index, which assumes an equilibrium production fixed at its initial level, can be called into question when accounting for the change in the equilibrium production itself. The indirect effect through the equilibrium production is neutralized in the particular case where demand is linear and costs are quadratic. Then, the slope of the best-reaction function, and so the value of the stability index, no longer depends on the number $\mathbf{q}^*$ of transported passengers in the equilibrium. Our empirical illustration relies on this particular framework.

With a linear demand function,

$$P(Q) = \delta_0 - \delta Q, \quad \delta_0 > 0, \delta > 0, \tag{7}$$

and a quadratic cost,

$$c_{af}(q) = \frac{q^2}{2\sigma_{af}},$$

where $\sigma_{af} > 0$ is a technological parameter that is specific to aircraft $\times$ airlines, the cost function solution to the program (1) is

$$C_f(q) = \frac{q^2}{2\sigma_f}, \quad \sigma_f = \sum_{a \in \mathcal{A}_f} n_{af}\sigma_{af}. \tag{8}$$

The parameter σ_f plays a central role in our model. It can be interpreted by noticing that both the marginal cost $C'_f(q)$ associated with (8) and its derivative $C''_f(q)$ are decreasing with σ_f for any given production q. As $C_f(0) = 0$ a higher value of σ_f implies a production efficiency gain (a lower production cost), which is made possible thanks to an increase in the

individual σ_{af} or because the capacity n_{af} goes up. This efficiency gain however comes with greater flexibility captured by dampened marginal costs. This makes firms more sensitive to expected changes in the production of others: the slope of the best-response function of firm f

$$R'_f(Q^*_{-f}) = -\frac{\delta\sigma_f}{2\delta\sigma_f + 1} \tag{9}$$

is decreasing with σ_f. This bundle of efficiency gains and greater flexibility drives a trade-off between surplus maximization in the Nash equilibrium and stability of this equilibrium illustrated by Proposition 2.

Proposition 2. *The transfer of an additional aircraft to some airlines in the linear-quadratic setup increases the aggregate equilibrium production Q^* but it locally destabilizes the equilibrium, i.e., it leads to an increase in the index $S(\mathbf{q}^*)$.*

Proof. See Appendix B. ∎

An additional aircraft allocated to the route corresponds to a higher transportation seat capacity, and so corresponds to an increase in the σ_f parameter. By Proposition 2 one should consequently observe in the data that routes with high traffic display a higher spread between the theoretical Nash equilibrium and the actual production.

The next result controls for route size by considering a transfer of aircraft between two airlines in the same route. This allows us to highlight the impact of the distribution of transportation capacities across airlines.

Proposition 3. *An aircraft reallocation from airlines f to airlines f' in the linear-quadratic setup increases the aggregate equilibrium production Q^* if and only if $\sigma_f > \sigma_{f'}$. This reallocation locally destabilizes the equilibrium, i.e., it leads to an increase in the index $S(\mathbf{q}^*)$, if and only if $\sigma_f > \sigma_{f'}$.*

Proof. See Appendix C. ∎

The trade-off between efficiency and stability illustrated in Proposition 2 is still valid. However, Proposition 2 would not allow us to discuss the impact of the transfer considered in Proposition 3 as the contributing airlines f entails an efficiency loss and a stability gain whereas airlines f', which enjoys the additional aircraft, is associated with an efficiency gain and a stability loss. Proposition 3 actually obtains by comparing the magnitudes of these two changes. It leads to the new testable prediction that some asymmetry in the airlines capacity in a given route, with large seat capacity firms competing against smaller ones, should be associated with a theoretical Nash equilibrium production closer to the actual one.

3 Empirical illustration to the airline industry

Our theoretical analysis predicts that the Nash equilibrium production should stand far from the observed production when the stability index $S(\mathbf{q}^*)$ is high (Proposition 1), the total production capacity is high (Proposition 2) and distributed evenly across firms (Proposition 3). We test these predictions on the US airline industry using data from the Bureau of Transportation Statistics over the 2003-2016 period. We estimate the demand for airlines tickets and aircraft costs fitting the linear-quadratic setup developed in Section 2.2. This allows us to compute the stability index $S(\mathbf{q}^*)$ and the volumes of transported passengers in the Nash equilibrium $\mathbf{q}^*$ by each airline, which can then be compared to the actual observed number of transported passengers. The results of these tests appear to be consistent with our theoretical predictions, though the linear-quadratic baseline model is plausibly a crude approximation of data reality. The constraints on available slots, aircrafts and seats possibly make the quadratic cost formulation a relevant approximation, while linear demand can be interpreted as a first-order approximation provided that prices and quantities do not vary too much in the route.

3.1 Costs

A market consists of all the flights between two endpoint cities, identified by their City Market ID number assigned by the U.S. Department of Transportation (DOT). We estimate cost functions using information contained in schedule P-5.1 of the Air Carrier Financial Reports and the Air Carrier Statistics T-100 Domestic Segment released by the Bureau of Transportation Statistics of the U.S. DOT.

Table 1: CARRIER COST DESCRIPTIVE STATISTICS

	Mean	Standard deviation	Min	Max
Aircraft costs (thousands of USD)	148,887	187,083	329	1,895,361
Passengers per carrier (in thousands)	1,931.1	2,657.2	16.242	26,000
Number of operated segments	196.41	112.96	17	604
Salary[1] (in thousands of USD)	22.59	6.51	8.19	46.34
Average fuel price[1] per 1000 gallons (in USD)	2,199.3	649.6	834.4	6,867.9

Number of observations: 3,575

1. Prices and costs are adjusted using the transportation sector price index of the Bureau of Labor Statistics, `http://www.bls.gov/cpi/`

Schedule P-5.1 includes data on input prices, maintenance expenses, equipment depreciation, rental costs, and total operating expenses disaggregated by airlines × aircraft type. Costs are available for each aircraft type, but the data does not include cost broken out by airline route. This data limitation obliges us to refer to an output defined at the aircraft level as well. As airlines produce passenger transportation services on non-stop flights using a single aircraft, the aircraft level coincides with the segment level, i.e., direct non-stop flights. A flight from city A to city C that entails a stop at city B consists of the two segments AB and BC, and we estimate separate costs for each of the two segments. At

the segment level we can ultimately recover the costs at airlines $\times$ segment $\times$ aircraft type level that appear in the theoretical model.

The financial information in Schedule P-5.1 is merged with Air Carrier Statistics T-100 Domestic Segment (U.S. Carriers) which contains domestic non-stop segment monthly data reported by U.S. air carriers, including origin and destination points, number of passengers carried, flight frequency, aircraft type and route length. The T-100 consists of more than three million observations over the sample window 2003:2016. We select the segments with distance above 100 miles, with more than 10 passengers per flight, with at least eight departures and 600 passengers during every quarter[5]. This yields a database with $3,575$ observations at the carrier $\times$ segment level that contains 22 carriers operating on $1,298$ segments and carrying 80 percent of the passengers transported in the U.S. domestic market. Some descriptive statistics for the average carrier are provided in Table 1.

The estimation of an aircraft cost function is based on the quadratic specification

$$c_{afst} = \frac{q_{afst}^2}{2\sigma_{afst}}, \tag{10}$$

which applies to a given flight operated by airline f with a type a aircraft during period t on segment s. The number of passengers q_{afst} transported on segment s by airlines f using aircraft type a during quarter t is readily available in the T-100 dataset.

The details of the cost estimation procedure are in Appendix D.

Table 2: ROUTE SUB-SAMPLE SELECTION FROM DB1B DATASET

	Number of routes	Mean	Standard deviation	Min	Max
Original base	11141				
Passengers per route (in thousands)		18.146	65.839	0.010	2986.0
Share of direct tickets per route		0.46	0.39	0	1
Direct routes (share of direct flights > 0.6)	3337				
Passengers per route (in thousands)		38.118	95.359	0.050	2986.0
Share of direct tickets per route		0.87	0.12	0.60	1
Routes in the final sample	379				
Passengers per route (in thousands)		265.9	263.8	9.9	2986.0
Share of direct tickets per route		0.90	0.09	0.60	1

3.2 Demand

Demand is estimated from the Airline Origin and Destination Survey (DB1B) database over the 2003:2016 period. The DB1B sample contains more than 4 million observations at the ticket level for every quarter. In order to match the segment perspective used on the supply side, we restrict our attention to markets with a high enough proportion of direct

---

[5]The T-100 dataset has been frequently used in the economic literature interested in airline competition. For further discussion on data selection, see for instance Ciliberto and Tamer (2009).

flights. We have chosen a threshold that eliminates routes with less than 60 percent of direct non-stop flight tickets. As shown in Table 2 there remains $3,337$ routes in the DB1B database. We disregard routes with one airlines in a monopoly situation, to which the theoretical part does not apply, and we only keep routes matching a segment existing in our final T-100 dataset. Then we use a cleaning process that mirrors the one applied to the T-100 dataset: we remove routes with distance below 100 miles and get rid of the routes with few passengers. We also discard tickets with extreme reported prices in the bottom and top 5% quantiles of the price per mile distribution, and we remove routes observed during less than 12 years in our 14-year sample window. We are eventually left with 379 routes. Compared to the original sample, our dataset tends to be biased toward routes with greater passenger traffic.

To estimate demand we compute the average quarterly number of passengers booking a flight on the selected segments and the corresponding average fare for each of these 379 routes. Fares are adjusted using the transportation sector price index of the Bureau of Labor Statistics. The database is enlarged with temperature and population of origin and destination cities.[6] Descriptive statistics are presented in Table 3.

We start from the linear demand specification (7),

$$Q_{st} = \gamma_{st}^0 + \gamma_{st} P_{st}, \tag{11}$$

where P_{st} and Q_{st} are respectively the average price level and the aggregate quantity of passengers transported in segment s during period t. Both are computed from the DB1B dataset.

The details of the demand estimation procedure are in Appendix E.

Table 3: Route descriptive statistics

	Mean	Standard deviation	Min	Max
Population in larger city (million)	6.419	4.552	0.928	18.663
Population in smaller city (million)	1.785	1.518	0.011	12.368
Average ticket price on the route* (USD)	182.8	74.6	23.9	586.8
Medium ticket price on the route* (USD)	161.7	54.2	8.2	467.1
Average price per mile (in USD)	0.286	0.108	0.079	0.654
Distance between two cities (thousand km)	0.783	0.557	0.013	2.918
Passengers per route (thousand)	265.9	263.8	9.9	2986.0
Number of airlines	3.084		2	7

Number of observations: 20,808

* Corrected by the consumer price index for transportation sector

---

[6]This information is obtained from `ggweather.com` and `citypopulation.de`.

4   Empirical results

The cost estimation procedure is based on $1,298$ segments of the T-100 dataset whereas we estimate demand from 379 routes of the DB1B dataset. In order to match supply and demand information, all the subsequent results are based on the 301 common segments that appear in the two samples.

4.1   Stability index

The cost efficiency parameter $\hat{\sigma}_{afst}$ obtains by replacing the parameters that appear in (15) with their estimated values reported in Appendix D.[7] We use (8) to aggregate $\hat{\sigma}_{afst}$ over aircraft groups to get the sufficient statistics $\hat{\sigma}_{fst}$ for the contribution of market supply to the stability of the Nash equilibrium. On the demand side, the results in Appendix E give the estimated slopes of the demand function δ_{st} that summarizes the contribution of market demand to stability. We therefore have all the information that is needed to compute the slope of the best-response function (9) in the linear-quadratic specification of the theoretical model, and so the stability index $S(\mathbf{q}^*)$ in (5).[8]

Table 4: SUFFICIENT STATISTICS FOR STABILITY

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Cost efficiency σ_{fst}	0.002	0.28	0.515	0.762	0.965	5.935
Inverse demand slope δ_{st}	0.915	1.28	1.495	1.609	1.816	5.113
Best-response slope $R'(\mathbf{q}^*)$	-0.482	-0.375	-0.309	-0.299	-0.231	-0.002
Stability index $S(\mathbf{q}^*)$	0.167	0.458	0.568	0.629	0.774	1.794

Table 4 shows that $S(\mathbf{q}^*)$ exceeds 1 for only 9 percent of the segment $\times$ quarter observations. Our theoretical model thus predicts successful coordination on the equilibrium in most markets, conditional on the fact that the set of the observed services $\times$ airlines coincides with the complete set of the services $\times$ airlines that constitute the relevant market.

The variability in the stability index mostly comes from segment heterogeneity: although our sample window includes the 2007 crisis and the Great Recession we find that time only explains 5 percent of the variance of the stability index. Table 5 provides us with a more detailed picture of the stability index by reporting the results of the regression $\log(S_{st}) = X'\beta + \text{period}_t + \varepsilon_{st}$ where the right-hand side variables in X consist of exogenous route characteristics as well as potentially endogenous competition indicators.

Higher values of the stability index apply to segments linking two distant and densely populated cities. Population matters even after controlling for the number of passengers.

---

[7]The constant term β_0 is set to the estimate $\hat{b}_0 = \beta_0 + d_0$ and the statistics $\hat{\sigma}_{afst}$ can only be recovered at the coarse level of aggregation of the aircraft group g rather than the fine level of the aircraft type a.

[8]See Figure 2 in Appendix G for the distribution of $\hat{\sigma}_{fst}$, $\hat{\delta}_{st}$, and the best-response slope.

Table 5: A picture of the stability index

	Dependent variable: $\log(S_{st})$
Constant	−2.958
	(0.069)
MARKET STRUCTURE	
$\log(\texttt{Distance}_s)$	−0.031
	(0.004)
$\log(\texttt{Lowest population}^1_{st})$	0.022
	(0.003)
$\log(\texttt{Highest population}^2_{st})$	0.019
	(0.008)
$\log(\texttt{Nb of airlines}_{st})$	0.060
	(0.044)
$\log(\texttt{Herfindahl index}_{st})$	−0.685
	(0.017)
$\log(\texttt{Share of direct flights}_{st})$	0.097
	(0.032)
$\log(\texttt{Nb of passengers}_{st})$	0.251
	(0.017)
AIRLINES CLUSTER	
Cluster 3 (AA)	−0.044
	(0.006)
Cluster 6 (UA-AA)	−0.036
	(0.008)
Cluster 4 (UA)	−0.034
	(0.003)
Cluster 2 (No high market share airlines)	−0.033
	(0.006)
Cluster 8 (WN-AA)	−0.031
	(0.003)
Cluster 7 (WN-UA-AA)	−0.012
	(0.010)
Cluster 9 (WN-UA)	−0.002
	(0.009)
Cluster 1 (WN)	reference
Cluster 5 (DL)	0.007
	(0.005)
Cluster 10 (WN-DL)	0.015
	(0.006)
Number of observations	7,422
R^2	0.897
Adjusted R^2	0.896
Residual Std. Error	0.123 (df = 7350)
F Statistic	899.472 (df = 71; 7350)

Note: robust standard errors clustered by route are in parenthesis

In accordance with Proposition 2, routes with high traffic display a higher stability index. Such segments are likely to involve high competition intensity. We have looked at two measures of competition. A higher number of competitors does not influence the value of the stability index, but a lower Herfindahl index, which reflects greater similarity of passenger transportation market shares among the competing airlines, is associated with higher values of the stability index. This fits Proposition 3, using market shares as proxies for airlines seat capacity.

The bottom of Table 5 emphasizes the types of interactions among airlines that give rise to higher indexes by clustering routes based on the operating airlines. Unlike the clustering relying on exogenous route characteristics that was used to estimate the cost function, the decision to enter is probably correlated with components of the stability index. We apply the Fanny clustering by Kaufman and Rousseeuw (1990) to the four airlines with the greatest market share in our whole sample: American Airlines (AA), Delta Airlines (DL), United Airlines (UA) and Southwest Airlines (WN). The clusters and an associated measure of within cluster similarity are reported in Table 6 (the highest within cluster similarity, which is normalized to 1, obtains when all the observations in the cluster are identical).

Columns 2 to 5 of Table 6 give the percentage of observations classified in the clusters where the four main airlines is active. In cluster 1 the low-cost airlines WN is present in every segment and quarter, occupying a quasi-monopoly position by competing against airlines with low market shares in the whole sample. Cluster 2 only consists of segments × quarter operated by low market share airlines. In cluster 10, UA appears in 17% of observations, and it always compete against WN and DL. The lower within cluster similarity found for cluster 10 shows that the competition structure changes over segments and time: some observations correspond to a duopoly with WN and DL possibly competing against smallest airlines, whereas other observations in cluster 10 involve WN and DL competing against UA and/or AA as well as smaller airlines.

Table 5 delineates three groups of airlines clusters. A first group with low stability indexes is formed by routes where UA and AA hold a dominant position and routes where they compete against each other. This group also includes routes where no major airlines is active. There is a second set of routes with intermediate values of the index, where WN compete against AA and UA. The presence of the low-cost company WN thus tends to be associated with higher values of the stability index. In the last group, with the highest values of the index, we find routes where WN and the major airlines DL interact.

4.2   Assessing the Nash equilibrium

We are now in a position to assess whether the spread between the actual and the theoretical Nash numbers of transported passengers is positively correlated with the indexes S_{st} over segments s and quarters t. We start from the squared spread between actual and Nash productions as

$$\|\mathbf{q}_{st} - \mathbf{q}_{st}^*\|^2 = \sum_{f \in \mathcal{F}_{st}} \left(q_{fst} - q_{fst}^*\right)^2,$$

Table 6: AIRLINES NETWORK FROM THE MAIN AIRLINES

cluster	WN	UA	AA	DL	Within homogeneity	Nb of segments	Nb of observations
1	100	0	0	0	1	83	1,448
2	0	0	0	0	1	71	631
5	0	0	0	100	1	62	1,071
8	100	0	100	0	1	57	738
10	100	17	8	100	0.580	57	719
9	100	100	0	0	1	56	957
4	0	100	0	15	0.740	50	486
7	100	100	100	5	0.920	43	614
3	0	0	100	32	0.570	30	490
6	0	100	100	36	0.590	21	268

where $\mathcal{F}_{st}$ represents the set of airlines active in segment s during quarter t. In order to avoid a mechanical bias with a higher spread in segments with intense traffic, the above spread is normalized by the squared number of passengers

$$\|\mathbf{q}_{st}^*\|^2 = \sum_{f \in \mathcal{F}_{st}} q_{fst}^{*2}.$$

Our normalized measure for the spread between Nash and actual production (at the segment $\times$ period level) is finally

$$\Delta_{st} = \frac{\|\mathbf{q}_{st} - \mathbf{q}_{st}^*\|}{\|\mathbf{q}_{st}^*\|}.$$
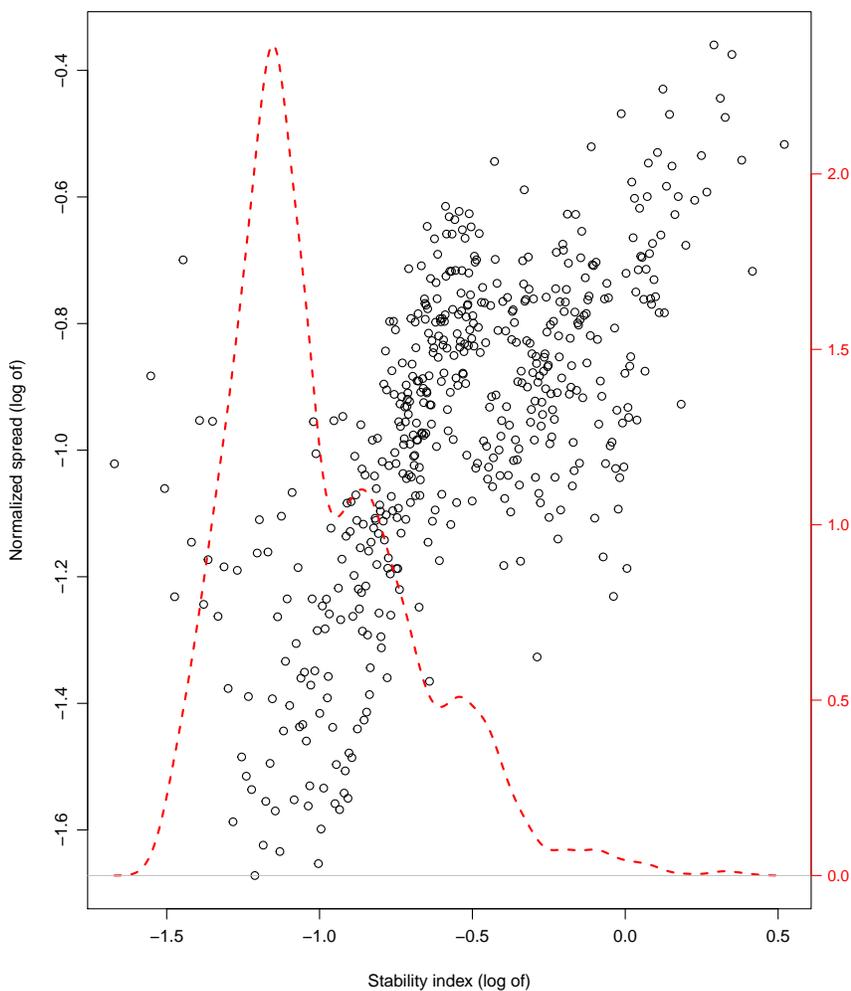

Figure 1: Density and average spread per range 2e-02 stability indexes

The equilibrium production q_f^* of airlines f is again assumed to the linear-quadratic specification. It obtains using the parameters σ_{fst} and δ_{st} estimated in Section 4. Some

16

details about the resulting spread density distribution are given in Appendix G. Figure 1 shows the high positive correlation between the stability index S_{st} and the spread Δ_{st}. We observe a smooth rise in the spread from 0 (where actual and Nash number of transported passengers coincide) when the index increases. The spread reaches its highest level around 0.8: the departure from Nash then represents 80% of the Nash production. The density of $S_{st} \times \Delta_{st}$ observations depicted in dotted red shows that this pattern applies to most observations, with a stability index between $\exp(-1.5) \simeq 0.2$ and $\exp(-0.25) \simeq 0.8$.

A strict account of the theoretical model consists in two regimes: the actual observed number of transported passengers coincides with the Nash equilibrium output if the stability index is below 1, while it differs from the Nash output if the stability index is above 1. To assess this prediction we consider the model

$$\Delta_{st} = b_1 \mathbf{1}[S_{st} \geq 1] + \texttt{Segment}_s + \texttt{Quarter}_t + \varepsilon_{st}$$

where $\mathbf{1}[S_{st} \geq 1]$ is a dummy that takes value 1 if the stability index S_{st} in route s at time t is greater than 1, and 0 otherwise. The spread between actual and Nash passengers reduces to the fixed effects in the low stability index regime ($S_{st} < 1$), and changes by b_1 in the high stability index regime ($S_{st} \geq 1$). The parameter b_1 should be positive for the model to be consistent with the theoretical prediction. The results reported in Table 7 show that the deviation from Nash in the Nash volume of passengers increases by 0.11 when switching to the high stability index regime. With an average fixed effect of 0.06, the total impact would thus equal 0.17.

Table 7: NASH DEPARTURE AND STABILITY INDEX: TWO-REGIME CASE

	Dependent variable: Nash spread Δ_{st}
	(1)
High stability index regime (b_1)	0.114
	(0.008)
Quarter fixed effect	Yes
Segment fixed effect	Yes
Observations	7,422
R^2	0.908
Adjusted R^2	0.903
Residual Std. Error	0.141 (df = 7065)
F Statistic	194.379 (df = 357; 7065)

Note: Robust standard errors reported at the segment level are in parenthesis

An alternative continuous measure obtains by focusing on the econometric relationship

$$\log \Delta_{st} = b_0 + b_1 \log S_{st} + \varepsilon_{st}.$$

The results of this regression are reported in Table 8. All the three variants reported in this Table display a positive high significant relation between the theoretical stability index and departures from Nash outcomes. We conclude that the stability index can be considered as a reliable predictor of the occurrence of the Nash equilibrium.

Table 8: Nash departure and stability index

| | Dependent variable: Nash spread $\log \Delta_{st}$ | | |
	(1)	(2)	(3)
Stability index $\log S_{st}$ (b_1)	0.539	0.574	0.713
	(0.131)	(0.145)	(0.089)
Constant (b_0)	−0.829	−0.731	−0.560
	(0.077)	(0.117)	(0.085)
Quarter fixed effect	No	Yes	Yes
Segment fixed effect	No	No	Yes
Number of observations	7,422	7,422	7,422
R^2	0.096	0.117	0.596
Adjusted R^2	0.096	0.110	0.575
Residual Standard Error	0.632 (df = 7420)	0.627 (df = 7365)	0.433 (df = 7065)
F-statistic	788.090 (df = 1; 7420)	17.459 (df = 56; 7365)	29.230 (df = 356; 7065)

Note: Robust standard errors reported at the segment level are in parenthesis

The variants differ according to fixed effects that are controlled for. The inclusion of time fixed effects does not yield substantial changes in the results. Instead, unobserved heterogeneity across segments explains an important large part of Nash departures. Our preferred specification in the third column includes both segments and quarter fixed effects: a 1 percent increase in the stability index is associated with a 0.713 percent rise in the spread between actual and Nash numbers of transported passengers.

The recent literature provides mixed evidence about whether firm size asymmetry helps stability by facilitating equilibrium coordination. In Kumar et al. (2015), competitive pressure improves firms' attention to global macroeconomic indicators and so makes forecasts about current market outcomes more accurate. In Byrne and de Roos (2019), the dominant firm BP in the retail gasoline industry is able to orient price strategies of the smaller retailers toward some focal point. The results reported in Tables 6 and 8 show that coordination failures on Nash outcomes are more likely in competitive markets with high traffic where competitors have similar market shares. The presence of Delta Airlines (DL), which completed its merger with Northwest (NW) from 2010, favors non-Nash behavior whereas Southwest (WN) complicates convergence toward a Nash equilibrium. This echoes

the empirical literature that documents the ability of Southwest to trigger fierce reactions by its competitors (see Morrison (2001) and Goolsbee and Syverson (2008)).

5   Nash departures and the relevant market

There are many circumstances in which it is difficult to identify the relevant market associated with a particular origin and destination city pair, i.e., the market that includes all the products/services that can be seen as relevant substitutes. One may think of multimarket contacts through common endpoints or transportation services as a composite good that consist of differentiated items, e.g., economy versus business class services, or non-stop direct versus indirect flights. Alternative products that entail other modes of transportation such as car or railway can also make identification more difficult. Neglecting part of the potential goods or services in the relevant market leads to underestimate the overall transportation capacity that the airlines have to predict. The econometrician working on a subsample of products could wrongly conclude in favor of market stability while instability actually prevails in the unobserved relevant market.

Proposition 1 suggests that departures from Nash equilibria should occur in routes with a stability index above 1. We adapt now this result to account for potential unobserved competition and characterize a new stability index threshold S_s^* in route s. This new threshold S_s^* falls below 1 if the full set of products which constitute the relevant market is not accounted for by the econometrician. Appendix F discusses the theoretical insights of our approach. In this section, we describe how to compute S_s^* and identify those routes where the relevant market is not observed.

5.1   Stability index threshold

We exploit two main ingredients, which are the spread between observed and Nash quantities and the estimated stability index, in order to identify the S_s^* threshold. We expect that a route s with a high spread Δ_{st} is more likely to originate from a high spread regime, i.e., a regime where $S_{st} \geq S_s^*$. Instead a low spread is more likely to be associated with a low spread regime where $S_{st} < S_s^*$.

Our estimation strategy works as follows. In a first step, we compute the empirical probability π_{st} that route s falls in the low spread regime during quarter t using the machine learning EM algorithm developed by Dempster, Laird and Rubin (1977). The details of the estimation procedure are provided in Appendix H. In a second step, we obtain individual estimates of the threshold $\hat{S}_{st}^*$ by matching π_{st} to the probability $\Pr(S_{st} \leq \hat{S}_{st}^*)$ referring to the empirical distribution of the estimated stability index, where the stability index S_{st} for all s and t is estimated from (5).

To switch from $\hat{S}_{st}^*$ to one single statistic $\hat{S}_s^*$ for each segment s, we set $\hat{S}_s^* = \max_t \hat{S}_{st}^*$. This should be seen as a conservative choice that classifies segment s as in equilibrium if it is in equilibrium during at least one period t. In most routes this choice is without loss

of great generality because the time dispersion of $\hat{S}_{st}^*$ within routes is low (see Figure 5 in the Appendix).

Table 9: Estimated stability index threshold distribution

			$\hat{S}_s^*$		
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.1862	0.8360	0.8698	0.8097	0.8714	0.8715

Number of routes: 301

Table 9 provides summary statistics about the estimated $\hat{S}_s^*$ at the route level. All values fall below 1, ranging from 0.18 to 0.87, which suggests that all routes in our dataset suffer to a certain extend from unobserved competition. The scope of unobserved competition is however limited as most routes involve a threshold above 0.8. With these new estimates we find that one-third of the routes can be regarded as falling in the high spread regime, whereas our initial threshold equal to 1 led to significantly more optimistic conclusions (less than 10% of the markets were supposed to be out of an equilibrium).

Table 10 illustrates how the estimated statistics $\hat{S}_s^*$ interacts with several explanatory variables. A high value of the threshold (a narrower scope for unobserved competition) is associated with routes linking less populated cities, which probably indicates that the share of indirect services supplied by competing airlines is lower in this case. More concentrated markets are associated with lower $\hat{S}_s^*$, which is not completely surprising given that these markets are generally operated by big traditional airlines supplying large menus of indirect services in addition to direct flights.

The effect of the observed market size when one accounts for unobserved competition can be put in perspective with our previous results in Table 5. We report in Table 11 information about the 10 routes with respectively the largest and smallest average number of passengers per quarter. The 10 largest routes display a larger average spread Δ_s and these departures from Nash behavior appear to reflect the combination of both higher stability index S_s and lower route threshold $\hat{S}_s^*$. No largest route in our sample eventually reaches equilibrium ($S_s > \hat{S}_s^*$) and instability here comes with high share of unobserved products ($\hat{S}_s^*$ is further away from 1); the two routes Atlanta–Miami and Boston–Washington are detected to be especially subject to high unobserved competition.

5.2 The Wright amendment repeal

The Wright amendment (WA) investigated in Ciliberto and Tamer (2009) serves as an interesting natural experiment which can be exploited in order to provide additional feedback about the validity of our methodology. The aim of the WA was to restrict airline services out of the Dallas Love airport (DAL) in order to stimulate the activity of the Dallas/Fort Worth airport (DFW).

Table 10: Stability index thresholds

	Dependent variable: $\hat{S}_s^*$	
	(1)	(2)
Route distance	7.515e-05	6.186e-05
	(1.618e-05)	(1.842e-01)
Least population endpoint	-2.789e-08	-1.080e-08
	(5.947e-09)	(7.549e-09)
Highest population endpoint	-4.314e-09	-1.984e-09
	(2.063e-09)	(2.288e-09)
Least temperature endpoint	-1.845e-03	-5.621e-04
	(1.373e-03)	(1.358e-03)
Least temperature endpoint	-1.410e-04	-1.384e-04
	(1.188e-03)	(1.152e-03)
Passenger per quarter		-2.743e-04
		(5.819e-05)
Share of direct flights		2.089e-02
		(1.313e-01)
Nb of airlines		-1.224e-02
		(2.038e-02)
Herfindahl index		-3.222e-01
		(1.312e-01)
Constant	0.933	1.055
	(6.872e-02)	(1.842e-01)
Nb of (route) observations	301	301
R^2	0.135	0.219
Adjusted R^2	0.120	0.195
Residual Std. Error	0.126 (df = 295)	0.121 (df = 291)
F Statistic	9.197 (df = 5; 295)	9.052 (df = 9; 291)

Note: Standard errors are in parenthesis

In 1980, the WA gets effective and states that airline services in DAL using large aircraft could be provided only to airports within Texas and its four neighboring U.S. states, namely Arkansas, Louisiana, New Mexico and Oklahoma (Allen (1989)). Flights to other states are allowed only on small aircraft. Airlines could not offer connecting flights, through service on another airline, or through ticketing beyond the five-state region. In October 2006 a partial repeal is decided and the full repeal gets effective in 2014.

Table 11: Relevant market for 20 airlines routes

Origin (US DOT city market)	Destination (US DOT city market)	Passengers (per quarter – thousand)	Distance (km)	Spread (within route average)	Stability index (within route average)	Threshold S_s^* (upper bound)
10 routes with the greatest average traffic per quarter						
Boston, MA	Washington, DC	1117	358	0.70	1.52	0.45
Orlando, FL	New York City, NY	1029	937	0.58	1.29	0.54
San Francisco, CA	Los Angeles, CA	948	332	0.46	1.40	0.60
Atlanta, GA	New York City, NY	924	748	0.81	1.02	0.54
Atlanta, GA	Miami, FL	905	584	0.90	0.95	0.31
Washington, DC	Chicago, IL	854	596	0.65	1.03	0.79
Atlanta, GA	Washington, DC	839	551	0.65	0.62	0.58
Las Vegas, NV	San Francisco, CA	761	399	0.59	1.06	0.84
Denver, CO	Los Angeles, CA	740	846	0.56	1.01	0.73
Phoenix, AZ	Los Angeles, CA	739	350	0.50	0.76	0.65
10 routes with the least average traffic per quarter						
Burlington, VT	Philadelphia, PA	44	335	0.27	0.29	0.87
Chicago, IL	Fort Wayne, IN	41	157	0.49	0.35	0.79
Denver, CO	Sioux Falls, SD	40	483	0.20	0.27	0.87
Jackson, WY	Salt Lake City, UT	39	205	0.79	0.25	0.85
Philadelphia, PA	Syracuse, NY	36	228	0.54	0.26	0.35
Philadelphia, PA	Richmond, VA	36	198	0.64	0.26	0.30
Chicago, IL	Sioux Falls, SD	34	462	0.35	0.25	0.82
Philadelphia, PA	Rochester, NY	33	257	0.35	0.24	0.84
Greensboro/High Point, NC	Philadelphia, PA	32	365	0.44	0.26	0.64
Denver, CO	Jackson, WY	32	406	0.16	0.21	0.87
Denver, CO	Rapid City, SD	32	300	0.23	0.19	0.85

The abrogation of airline service restrictions from DAL in a Southwest stronghold area implies greater competitive pressure on DFW, where American Airlines operates direct non-stop long-haul flights. The abrogation of service restrictions affects the expansion of the size of the relevant market of services including the Dallas/Fort Worth area as an origin or destination point, depending on whether the point of destination of origin belongs to the so-called five-state region or not. Consider the case of the Dallas-Washington market for instance: under the WA, all flights had to go through DFW because no services were allowed from/to DAL; all of the airline services of the relevant market Dallas-Washington would therefore be products operated from/to DFW. After the abrogation of the WA, the same relevant market would typically include all airline services from both DAL and DFW. If the econometrician has only data on airline services from/to DFW, she does not suffer from any missing information as long as the WA is effective (in which case the stability index threshold should be close to 1); after the abrogation of the WA however, a significant share of information would be missing, and this should be reflected in a fall in the stability index threshold. The results reported in Table 12 are largely consistent with these predictions.

Table 12: THE RELEVANT MARKET OF DALLAS/FORT WORTH

Origin (US DOT city market)	Destination (US DOT city market)	Sub-period	Passengers (per quarter – thousand)	Distance (km)	Spread (within route average)	Stability index (within route average)	Threshold S_*^* (upper bound)
AREA UNRESTRICTED BY WRIGHT AMENDMENT[1]							
Dallas/Fort Worth, TX	Lubbock, TX	2003:1-2006:2	161	282	0.29	0.45	0.86
Dallas/Fort Worth, TX	Lubbock, TX	2006:3-2014:2	164	282	0.35	0.47	0.87
Dallas/Fort Worth, TX	Lubbock, TX	2014:3-2016:4	138	282	0.12	0.50	0.86
Dallas/Fort Worth, TX	San Antonio, TX	2003:1-2006:2	480	247	0.45	0.58	0.61
Dallas/Fort Worth, TX	San Antonio, TX	2006:3-2014:2	502	247	0.51	0.59	0.79
Dallas/Fort Worth, TX	San Antonio, TX	2014:3-2016:4	472	247	0.36	0.60	0.80
Dallas/Fort Worth, TX	New Orleans, LA	2003:1-2006:2	196	437	0.30	0.45	0.83
Dallas/Fort Worth, TX	New Orleans, LA	2006:3-2014:2	243	437	0.34	0.54	0.87
Dallas/Fort Worth, TX	New Orleans, LA	2014:3-2016:4	299	436	0.47	0.71	0.79
Dallas/Fort Worth, TX	Tulsa, OK	2003:1-2006:2	209	237	0.19	0.66	0.87
Dallas/Fort Worth, TX	Tulsa, OK	2006:3-2014:2	222	237	0.45	0.58	0.86
Dallas/Fort Worth, TX	Tulsa, OK	2014:3-2016:4	199	237	0.37	0.53	0.82
Albuquerque, NM	Dallas/Fort Worth, TX	2003:1-2006:2	221	569	0.35	0.63	0.73
Albuquerque, NM	Dallas/Fort Worth, TX	2006:3-2014:2	276	569	0.38	0.54	0.87
Albuquerque, NM	Dallas/Fort Worth, TX	2014:3-2016:4	231	569	0.18	0.52	0.87
Dallas/Fort Worth, TX	Kansas City, MO	2003:1-2006:2	221	460	0.48	0.49	0.64
Dallas/Fort Worth, TX	Kansas City, MO	2006:3-2014:2	285	460	0.40	0.56	0.84
Dallas/Fort Worth, TX	Kansas City, MO	2014:3-2016:4	285	460	0.27	0.61	0.87
Dallas/Fort Worth, TX	Austin, TX	2003:1-2006:2	454	189	0.50	0.58	0.51
Dallas/Fort Worth, TX	Austin, TX	2006:3-2014:2	480	189	0.57	0.59	0.60
Dallas/Fort Worth, TX	Austin, TX	2014:3-2016:4	487	189	0.48	0.60	0.59
AREA RESTRICTED BY WRIGHT AMENDMENT							
Dallas/Fort Worth, TX	Philadelphia, PA	2003:1-2006:2	155	1302	0.30	0.45	0.85
Dallas/Fort Worth, TX	Philadelphia, PA	2006:3-2014:2	207	1302	0.41	0.51	0.86
Dallas/Fort Worth, TX	Philadelphia, PA	2014:3-2016:4	274	1300	0.66	0.68	0.65
Dallas/Fort Worth, TX	Denver, CO	2003:1-2006:2	373	641	0.42	0.76	0.79
Dallas/Fort Worth, TX	Denver, CO	2006:3-2014:2	406	641	0.51	0.86	0.78
Dallas/Fort Worth, TX	Denver, CO	2014:3-2016:4	512	641	0.61	1.08	0.27
Dallas/Fort Worth, TX	Atlanta, GA	2003:1-2006:2	513	732	0.35	0.83	0.84
Dallas/Fort Worth, TX	Atlanta, GA	2006:3-2014:2	516	730	0.36	0.88	0.87
Dallas/Fort Worth, TX	Atlanta, GA	2014:3-2016:4	584	725	0.55	0.95	0.42
Dallas/Fort Worth, TX	Phoenix, AZ	2003:1-2006:2	335	868	0.43	0.60	0.83
Dallas/Fort Worth, TX	Phoenix, AZ	2006:3-2014:2	319	866	0.41	0.59	0.87
Dallas/Fort Worth, TX	Phoenix, AZ	2014:3-2016:4	422	868	0.71	0.78	0.67
Dallas/Fort Worth, TX	Seattle, WA	2003:1-2006:2	236	1660	0.50	0.46	0.77
Dallas/Fort Worth, TX	Seattle, WA	2006:3-2014:2	250	1660	0.53	0.49	0.69
Dallas/Fort Worth, TX	Seattle, WA	2014:3-2016:4	294	1660	0.54	0.54	0.49
Dallas/Fort Worth, TX	Salt Lake City, UT	2003:1-2006:2	178	988	0.27	0.48	0.82
Dallas/Fort Worth, TX	Salt Lake City, UT	2006:3-2014:2	162	988	0.30	0.51	0.87
Dallas/Fort Worth, TX	Salt Lake City, UT	2014:3-2016:4	204	988	0.30	0.62	0.86
Dallas/Fort Worth, TX	Chicago, IL	2003:1-2006:2	426	802	0.49	0.54	0.41
Dallas/Fort Worth, TX	Chicago, IL	2006:3-2014:2	443	800	0.67	0.76	0.41
Dallas/Fort Worth, TX	Chicago, IL	2014:3-2016:4	588	795	0.80	1.03	0.22
Dallas/Fort Worth, TX	Charlotte, NC	2003:1-2006:2	147	936	0.28	0.47	0.86
Dallas/Fort Worth, TX	Charlotte, NC	2006:3-2014:2	248	935	0.30	0.49	0.87
Dallas/Fort Worth, TX	Charlotte, NC	2014:3-2016:4	320	931	0.57	0.50	0.69
Dallas/Fort Worth, TX	Orlando, FL	2003:1-2006:2	305	984	0.63	0.49	0.81
Dallas/Fort Worth, TX	Orlando, FL	2006:3-2014:2	277	983	0.60	0.45	0.74
Dallas/Fort Worth, TX	Orlando, FL	2014:3-2016:4	344	977	0.71	0.68	0.24
Dallas/Fort Worth, TX	Las Vegas, NV	2003:1-2006:2	382	1055	0.65	0.66	0.42
Dallas/Fort Worth, TX	Las Vegas, NV	2006:3-2014:2	352	1055	0.62	0.58	0.52
Dallas/Fort Worth, TX	Las Vegas, NV	2014:3-2016:4	445	1055	0.72	0.80	0.26
Dallas/Fort Worth, TX	San Francisco, CA	2003:1-2006:2	436	1456	0.71	0.56	0.30
Dallas/Fort Worth, TX	San Francisco, CA	2006:3-2014:2	427	1458	0.75	0.53	0.45
Dallas/Fort Worth, TX	San Francisco, CA	2014:3-2016:4	498	1455	0.91	0.80	0.17

The amendment initially restricted service from Dallas Love outside Texas, Arkansas, Louisiana, New Mexico and Oklahoma; and Missouri from 2005.

We propose to test empirically this prediction with our data, using only information on services from/to DFW. We consider three sub-periods, namely 2003:1-2006:2 (before the announcement of the repeal of the WA), 2006:3-2014:2 (from the announcement to the repeal of the WA) and 2014:2-2016:4 (after the repeal of the WA). Our subsample contains 18 routes that include DFW at some endpoints. Table 12 shows that the stability index threshold is stable across the three periods in every market that makes a connection between the Dallas/Fort Worth area and a city market in the WA zone (i.e., a city market located in Texas, Arkansas, Louisiana, New Mexico or Oklahoma). We do not detect any systematic change in the stability index or the spread between actual and Nash volumes of transported passengers in these routes.

The situation is however quite different for markets connecting the Dallas/Fort Worth area to cities outside the Wright zone. Indeed we find that the threshold is sharply reduced after the repeal of the amendment (i.e., from 2006:3-2014:2 to 2014:2-2016:4), whereas the announcement effect seems to be not significant (there is no significant difference between 2003:1-2006:2 and 2006:3-2014:2). We also find a larger spread after the repeal as we observe higher stability indexes. This probably suggests that the repeal of the WA introduced some instability in each market from/to the Dallas/Fort Worth area.

6   Discussion

The stability index governs a process where airlines iteratively eliminate dominated choices of volumes of passenger-carrying services that can be viewed as taking place within every period t. Empirical evidence from surveys about expectations documents an extrapolative nature of expectations based on past observations and forecasts (see Ma, Sraer and Thesmar (2018) for recent examples). In our context, it seems plausible that airlines also process past information, e.g., past capacity choices of their competitors, to form/revise their beliefs about period t choices of their competitors.

In this section we suggest that airlines both rely on rationalizability-based inferences and process past information using adaptive learning. We also illustrate possible welfare consequences of instability by considering situations where the entry and exit of firms creates instability and impedes firms from reaching a Nash equilibrium. Airlines appear overly optimistic by scheduling too much seats, which yields a short run consumer surplus above its Nash equilibrium value.

6.1   Adaptive learning

Here we test whether the role of the stability index only transits through adaptive learning. We restrict our attention to the special case where airlines hold identical beliefs over the individual productions of their competitors. Then, a reduced form for the actual realization of the spread Δ_{st} that fits the local stability properties of our theoretical setup is $\Delta_{st} = S_{st}\Delta_{st}^e$ where Δ_{st}^e stands for the common forecast about the (normalized) spread between the actual and Nash volume of passengers in segment s during quarter t. We introduce

adaptive learning by assuming that the belief Δ_{st}^e is determined by a log-linear function of past forecast errors,

$$\Delta_{st}^e = \prod_{i>0} \Delta_{s(t-i)}^{\gamma_i},$$

where γ_i is a parameter that weights the realized error i quarters before the formation of the forecast Δ_{st}^e. With this rule, the actual dynamics of the spread writes

$$\log \Delta_{st} = \log S_{st} + \sum_{i>0} \gamma_i \log \Delta_{s(t-i)}.$$

This specification highlights that $\log \Delta_{st}$ should be in the same proportion as $\log S_{st}$ if the role played by the stability index S_{st} only transits through adaptive learning. Hence we consider the following econometric specification:

$$\log \Delta_{st} = b_1 \log S_{st} + \sum_{i>0} \gamma_i \log \Delta_{s(t-i)} + \texttt{Segment}_s + \texttt{Quarter}_t + \varepsilon_{st}.$$

If the estimated parameter $\hat{b}_1$ differs from 1, we can reject the null hypothesis that only adaptive learning matters in the determination of the current departure from Nash.

The model is estimated using the extended linear GMM estimator based upon lagged differences of Δ_{st} in addition to lagged levels of Δ_{st} as instruments (Blundell and Bond, 1998). The results are provided in Table 10 together with several specification tests. The Sargan test is not completely convincing as the statistic shown presents significant evidence against the null hypothesis that the overidentifying restrictions are valid. The latter implies that we need in principle to reconsider our model or our instruments, unless we attribute the rejection to heteroskedasticity in the data-generating process. The presence of heteroskedasticity is a realistic assumption, as suggested by Arellano and Bond (1991). Another important assumption is that, when the errors ε are independently and identically distributed, the first difference errors are first-order serially correlated, which is confirmed here: first, the test statistic presents strong evidence against the null hypothesis of zero autocorrelation in the first-differenced errors at order 1; second, it presents no significant evidence of serial correlation in the first-differenced errors at order 2, which strengthens the validity of our results.

The results reported in Table 13 show that past forecasting errors do indeed matter. The estimate $\hat{b}_1$ is usually significantly different from 1. One can thus reject the null hypothesis that the role played by the stability index is only adaptive learning based. The specification used in Column 1 shows that errors made by airlines two and three quarters earlier have a significant impact on the current Nash spread, but not the errors made in the last quarter; the short-run stickiness of seat capacity adjustments makes the last quarter seat capacity the best proxy for the current one, but this information may not be publicly available when airlines form their forecasts. In every specification airlines put less weight on far distance past errors, $|\hat{\gamma}_2| > |\hat{\gamma}_3|$.

The full sample of observations is used in the first two columns of Table 13. The impact $\hat{b}_1$ of the stability index on the current departure from Nash remains very similar

to the one found when neglecting the effect of past departures from Nash (in Column 3 of Table 8). The results in the last two columns apply to sub-samples that only consist of segments where there is no major airlines among the four companies selected in the clustering made in Table 6 (Column 3), or segments where only these airlines are active (Column 4). The specifications in these last two columns suggest that large airlines display a greater confidence in past strategic choices of their competitors than small airlines to form their current forecasts as the difference between $\hat{b}_1$ and 1 seems to loose significance in column 4.

Table 13: ADAPTIVE LEARNING

| | Dependent variable: $\log \Delta_{st}$ | | | |
	(1)	(2)	(3)	(4)
$\log \Delta_{s(t-1)}$	0.258			
	(0.577)			
$\log \Delta_{s(t-2)}$	−0.382	−0.320	−0.348	−0.301
	(0.142)	(0.057)	(0.074)	(0.080)
$\log \Delta_{s(t-3)}$	−0.175	−0.227	−0.247	−0.221
	(0.137)	(0.033)	(0.050)	(0.038)
$\log S_{st}$	0.705	0.774	0.705	0.799
	(0.200)	(0.089)	(0.123)	(0.125)
Nb of segments	301	301	178	123
Quarters used (over 56)	1-51	1-51	1-47	2-51
Nb of observations	7422	7422	3510	3912
Sargan test (p-value)	0.01	0.05	0.151	0.331
Autocorrelation test (1) (p-value)	0.287	1.77e-07	1.71e-04	5.92e-05
Autocorrelation test (2) (p-value)	0.519	0.676	0.241	0.320
Wald test for coefficients (p-value)	$\leq$ 2.22e-16	$\leq$ 2.22e-16	5.01e-12	$\leq$ 2.22e-16
Wald test for time dummies (p-value)	4.55e-11	5.54e-09	1.44e-05	$\leq$ 2.22e-16

Note: We use the robust covariance matrix proposed by Windmeijer (2005). Standard errors are in parenthesis

6.2 Welfare impact of entry

A standard problem faced by competition authorities is to evaluate the impact of a new firm entry (or exit) on the consumer surplus. The empirical framework usually assumes that a Nash equilibrium is obtained before and after entry (Scenario 1). Hence, the change in consumers' surplus is

$$\Delta^*_{Surplus} = -\frac{1}{2} \left(p_1^* Q_1^* - p_0^* Q_0^* \right),$$

where Q_0^* and Q_1^* are the equilibrium productions before and after entry, respectively.

Here, we relax this hypothesis and assume instead that higher competition following entry possibly entails instability so that firms may not be able to play Nash quantities. We are thus interested in a scenario where airlines play a Nash equilibrium before entry while

26

the equilibrium becomes unstable after entry and entails a production level Q_1 which is different from the Nash quantity and is directly observed in our sample (Scenario 2). The change in consumers' surplus after entry can then be computed as

$$\Delta_{Surplus} = -\frac{1}{2}\left(p_1 Q_1 - p_0^* Q_0^*\right).$$

Table 14: WELFARE IN THE NEW-YORK/TAMPA SEGMENT

Period	2011:4	2012:1
Competing careers	CO, WN, DL, B6	UA, WN, DL, B6, AA
Stability index S_{st}	0.974	1.055
Nash Quantity (thousands)	405.5	417.1
Observed Quantity (thousands)	405.5	426,5
Nash price (in USD)	336.2	321.2
Observed price (in USD)	336.2	305.9
Nash Surplus (in USD)	133,498.6	141,640.4
$\Delta_{Surplus}^*$ (in USD)		+8,141.8
$\Delta_{Surplus}^*$ per passenger (in USD)		+19.5
Real Surplus (in USD)	133,498.6	148,088.9
$\Delta_{Surplus}$ (in USD)		+14,590.3
$\Delta_{Surplus}$ per passenger (in USD)		+34.2

To evaluate the difference between $\Delta_{Surplus}^*$ and $\Delta_{Surplus}$, we pick up a specific market, namely New-York/Tampa, where a change in the number of the competitors makes the value of the stability index jump above 1. In this market, Continental (CO), Southwest (WN), Delta (DL), and Jetblue (B6) compete against each other in 2011:4, whereas the next quarter witnesses the exit of Continental, and the simultaneous entry of American Airlines and United Airlines. We compute the equilibrium productions Q_0^* and Q_1^* as suggested in Section 4.2. The corresponding prices p_0^* and p_1^* obtain using the estimated demand slope and intercept on segment s at time t from $\hat{\delta}_{st}^0$ and $\hat{\delta}_{st}$ in Table 17.

As shown in Table 14, the Nash quantity Q_1^* is greater than Q_0^* while the Nash price decreases from 336.2 USD to 321.2 USD. Thus, the increase in total surplus $\Delta_{Surplus}^*$ per passenger in Scenario 1 is equal to 19.5 USD. In Scenario 2, where carriers make prediction mistakes in 2012:1, the increase in total surplus $\Delta_{Surplus}$ per passenger is equal to 34.2 USD. Hence assuming Nash behavior after entry in this particular market leads to an under-evaluation of the surplus gain.

The change in total production from one period to another can be decomposed at the airline level, as illustrated in Table 15. We find that Delta (DL) behaves very much in line with Nash whereas JetBlue (B6) and American Airlines (AA) produce much more than what the Nash equilibrium would predict.

Table 15: Passengers in the New-York/Tampa segment

Airlines	Nash Quantity 2011:4	Nash Quantity 2012:1	Observed Quantity 2012:1
CO	105,9	-	-
NW	58.6	51.6	36.9
DL	101.8	94.9	89.7
B6	118.8	104.6	141.7
AA	-	28.5	19.7
UA	-	109.4	138.2

7　Conclusion

Greater competition is often viewed as driving welfare gains from lower equilibrium prices; our article shows that it may also compromise the occurrence of an equilibrium. Thus, in markets where the usual indicators of high competitive pressure are present, i.e., those where several airlines with similar market shares or competitive low cost companies are present, the traditional equilibrium welfare analysis has to be worked out carefully. Eventually the equilibrium would be a reliable reference in only $70 - 90\%$ of the routes.

Our analysis is subject to a number of potential limitations that could be analyzed in future work.

1. Data from the U.S. Department of Transportation provide us with information about airlines costs at the non-stop flight segment level. We therefore estimate demand at the same level, i.e., we restrict ourselves to routes where the share of direct flights is high enough. Our identification procedure for the scope of the relevant market however suggests that indirect flights matter, especially in routes with large flows of passengers, and so should be explicitly taken into account. Any initiative that could ease the combination of the two types of information is obviously welcome.

2. Our analysis abstracts from dynamic aspects that are certainly important in shaping the regular interactions between the airline companies that compete in a route. The insights from our robustness check suggest that firms certainly retrieve valuable information from bad prediction in the past. Brandenburger, Danieli and Friedenberg (2019) makes progress toward identification of the level of rationality in this context. An empirical application on the airline industry may be more challenging to implement as medium and long-run strategies also encompass both slot portfolios, which requires introducing airports into the analysis, as well as capacity choices.

3. A relevant choice for the measure of the discrepancy between the actual and equilibrium strategies regarding volumes of transported passengers needs a suitable equilibrium reference. Our article refers to a version of a non-cooperative Cournot game with linear passengers demand and quadratic airline costs, but a positive spread could also come from an inadequate reference. There are two main sources of model misspecification in our context. The first one relates to the restrictive linear-quadratic modeling of market fundamentals that precludes multiplicity of Nash equilibria. In the presence of multiple Nash equilibria, local rationalizability can still be employed as a selection device to eliminate unstable equilibria from the set of empirically relevant outcomes. Still multiplicity of locally stable equilibria is possible, in which case there is no longer any obvious equilibrium reference. The second source of misspecification relates to the short-run strategies assigned to airlines. Volumes of passengers is only part of yield management strategies: individual price setting certainly also enters the airlines choice set; and, as suggested for instance by Ciliberto and Williams (2014), explicit alliances or collusive behaviors weaken the status of the non-cooperative equilibrium reference.

References

[1] Aguirregabiria, V. and J. Jeon (2020). Firms' beliefs and learning: Models, identification, and empirical evidence. Review of Industrial Organization, 56(2), 203-235.

[2] Aguirregabiria, V. and E. Xie (2020). Identifcation of Non-Equilibrium Beliefs in Games of Incomplete Information Using Experimental Data. Journal of Econometric Methods, 1.

[3] Aguirregabiria, V. and A. Magesan (2020). Identification and estimation of dynamic games when players' beliefs are not in equilibrium. The Review of Economic Studies, 87(2), 582-625.

[4] Alaoui, L. and A. Penta (2016). Endogenous depth of reasoning. The Review of Economic Studies, 83(4), 1297-1333.

[5] Allen, E. A. (1989). The Wright Amendment: The Constitutionality and Propriety of the Restrictions on Dallas Love Field. J. Air L. & Com., 55, 1011.

[6] Angeletos, G. M. and C. Lian (2018). Forward guidance without common knowledge. American Economic Review, 108(9), 2477-2512.

[7] Aradillas-Lopez, A. and E. Tamer (2008). The identification power of equilibrium in simple games. Journal of Business & Economic Statistics, 26(3), 261-283.

[8] Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297.

[9] Basso, L. J. (2008). Airport deregulation: Effects on pricing and capacity. International Journal of Industrial Organization, 26(4), 1015-1031.

[10] Bernheim, B. D. (1984). Rationalizable strategic behavior. Econometrica: Journal of the Econometric Society, 1007-1028.

[11] Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics, 87(1), 115-143.

[12] Borenstein, S. and N.L. Rose (1994). Competition and price dispersion in the US airline industry. Journal of Political Economy, 102(4), 653-683.

[13] Brandenburger, A., Danieli, A. and A. Friedenberg (2019). Identification of Reasoning about Rationality.

[14] Brander, J.A. and A. Zhang (1990). Market conduct in the airline industry: an empirical investigation. The RAND Journal of Economics, 567-583.

[15] Brueckner, J.K. (2002). Airport congestion when carriers have market power. American Economic Review, 92(5), 1357-1375.

[16] Byrne, D.P. and N. De Roos (2019). Learning to coordinate: A study in retail gasoline. American Economic Review, 109(2), 591-619.

[17] Ciliberto, F. and E. Tamer. (2009). Market structure and multiple equilibria in airline markets. Econometrica, 77(6), 1791-1828.

[18] Ciliberto, F., and J.W. Williams (2014). Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry. The RAND Journal of Economics, 45(4), 764-791.

[19] Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), 2644-78.

[20] Coibion, O., Gorodnichenko, Y., Kumar, S. and J. Ryngaert (2018). Do you know that I know that you know? higher-order beliefs in survey data. National Bureau of Economic Research 24987.

[21] Crawford, V. P. and N. Iriberri (2007). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?. Econometrica, 75(6), 1721-1770.

[22] Crawford, V. P., Costa-Gomes, M. A. and N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. Journal of Economic Literature, 51(1), 5-62.

[23] Davis, P. and E. Garces (2009). Quantitative techniques for competition and antitrust analysis. Princeton University Press.

[24] A.P. Dempster, Laird, N.M. and D. Rubin (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society 39(1).

[25] Desgranges, G. and S. Gauthier (2016). Rationalizability and efficiency in an asymmetric Cournot oligopoly. International Journal of Industrial Organization, 44, 163-176.

[26] Doraszelski, U., Lewis, G. and A. Pakes (2018). Just starting out: Learning and equilibrium in a new market. American Economic Review, 108(3), 565-615.

[27] Farhi, E. and I. Werning (2017). Fiscal unions. American Economic Review, 107(12), 3788-3834.

[28] Giacomini, R., Skreta, V. and J. Turen (2020). Heterogeneity, Inattention, and Bayesian Updates. American Economic Journal: Macroeconomics, 12(1), 282-309.

[29] Goolsbee, A., & Syverson, C. (2008). How do incumbents respond to the threat of entry? Evidence from the major airlines. The Quarterly journal of economics, 123(4), 1611-1633.

[30] Guesnerie, R. (1992). An exploration of the eductive justifications of the rational-expectations hypothesis. The American Economic Review, 1254-1278.

[31] Kashaev, N. (2016). Testing for Nash behavior in entry games with complete information, Pennsylvania State University working paper.

[32] Kaufman, L., and Rousseeuw, P. J. (1990). Partitioning around medoids (program pam). Finding groups in data: an introduction to cluster analysis, 344, 68-125.

[33] Kneeland, T. (2015). Identifying Higher-Order Rationality. Econometrica, 83(5), 2065-2079.

[34] Kosenkova, L. (2017). Nonparametric inference in asymmetric first-price auctions with k-rationalizable beliefs, Pennsylvania State University working paper.

[35] Kreps, D.M., and J.A. Scheinkman (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. The Bell Journal of Economics, 326-337.

[36] Kumar, S., Afrouzi, H., Coibion, O., and Y. Gorodnichenko (2015). Inflation targeting does not anchor inflation expectations: Evidence from firms in New Zealand (No. w21814). National Bureau of Economic Research.

[37] Landier, A., Ma, Y. and D. Thesmar (2019). Biases in expectations: Experimental evidence. Available at SSRN 3046955.

[38] Ma, Y., D. Sraer and D. Thesmar (2018). The aggregate cost of systematic forecast errors, working paper, https://sites.google.com/site/dthesmar/

[39] Morrison, S.A. (2001). Actual, adjacent, and potential competition estimating the full effect of Southwest Airlines. Journal of Transport Economics and Policy (JTEP), 35(2), 239-256.

[40] Moulin, H. (1984). Dominance solvability and Cournot stability. Mathematical social sciences, 7(1), 83-102.

[41] Nachbar, J.H. (1990). Evolutionary selection dynamics in games: Convergence and limit properties, International journal of game theory 19 (1), 59-89.

[42] Oum, T.H., Waters, W.G. and J.S. Yong, (1992). Concepts of price elasticities of transport demand and recent empirical estimates: an interpretative survey. Journal of Transport Economics and policy, 139-154.

[43] Osborne, M.J. and A. Rubinstein (1994). A course in game theory. MIT press.

[44] Pearce, D.G. (1984). Rationalizable strategic behavior and the problem of perfection. Econometrica: Journal of the Econometric Society, 1029-1050.

[45] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. Journal of econometrics, 126(1), 25-51.

[46] Woodford, M. (2003). Optimal interest-rate smoothing. The Review of Economic Studies, 70(4), 861-886.

Appendix

A   Proof Proposition 1

By Lemma 3 in Desgranges and Gauthier (2016) the matrix $\mathbf{B}$ has a spectral radius less than 1 if and only if (5) is satisfied. The slope in (6) obtains by differentiating the first-order condition for firm f profit maximization. The expression of $C_f''(q_f^*)$ obtains by applying the generalized envelope theorem to the cost minimization program (1). This yields $C_f'(q_f) = \lambda$, where λ is the non-negative Lagrange multiplier associated with the production constraint in (1). Therefore,

$$C_f''(q_f) = \frac{d\lambda}{dq_f}.$$

The first-order conditions for the cost minimization problem (1) are

$$\lambda = c_{af}'(q_{af}) \text{ for all } a \in \mathcal{A}_f \quad \text{and} \quad \sum_{a \in \mathcal{A}_f} n_{af} q_{af} = q_f$$

for every f. Differentiating these first-order conditions yields

$$d\lambda = c_{af}''(q_{af}) dq_{af} \text{ for all } a \in \mathcal{A}_f \quad \text{and} \quad \sum_{a \in \mathcal{A}_f} n_{af} dq_{af} = dq_f$$

for every f. Reintroducing the expression of dq_{af} into the last equality we finally obtain

$$d\lambda \sum_{a \in \mathcal{A}_f} \frac{n_{af}}{c_{af}''(q_{af})} = dq_f.$$

This completes the proof.

B   Proof Proposition 2

In the linear-quadratic specification, the stability index

$$S(\mathbf{q}^*) = \sum_f \frac{\delta \sigma_f}{1 + 3\delta \sigma_f}. \tag{12}$$

does not depend on production. The equilibrium production of firm f satisfies

$$q_f^* = \frac{\sigma_f}{1 + \sigma_f \delta}(\delta_0 - \delta Q^*), \tag{13}$$

Summing gives the aggregate equilibrium production

$$Q^* = \frac{\displaystyle\sum_f \frac{\delta_0\,\sigma_f}{1+\delta\sigma_f}}{1+\displaystyle\sum_f \frac{\delta\,\sigma_f}{1+\delta\sigma_f}}. \tag{14}$$

Both the index $S(\mathbf{q}^*)$ and the production Q^* are increasing with σ_f (which itself increases with σ_{af}). An additional marginal aircraft transfer corresponds to $d\sigma_f > 0$. The result follows.

C   Proof Proposition 3

It is similar to the proof of proposition 2 but now exploits second-order derivatives of $S(\mathbf{q}^*)$ and Q^*. In the linear-quadratic specification, $S(\mathbf{q}^*)$ and Q^* are increasing concave functions of σ_f (which is itself increasing linear in σ_{af}). A marginal aircraft transfer $dn_{af} = -dn_{af'} = -1$ yields $d\sigma_f = -\sigma_{af} < 0$ and $d\sigma_{f'} = -d\sigma_f > 0$. The result follows.

D   Costs estimation

In order to estimate the parameter σ_{afst} that enters the stability index at the segment level using the aggregate cost in (8), we express this parameter as

$$\frac{1}{2\sigma_{afst}} = \beta_0\,\xi_{ft}\,\mu_s\,\nu_a \tag{15}$$

where β_0 is a constant term, ξ_{ft} varies across firms and time periods, whereas μ_s and ν_a are segment and aircraft fixed effects, respectively. The variable ξ_{ft} accounts for unobserved characteristics of airlines productive efficiency, e.g., managerial effort or marketing strategies, each of which plausibly varies over time. In (15), segment and aircraft fixed effects are restricted to be time invariant, but our final cost specification includes time fixed effects common to segments and aircraft. Using (10) and (15) the aggregate cost C_{aft} given in (8) becomes

$$C_{aft} = \beta_0\,\xi_{ft}\,\nu_a \sum_s \mu_s\,n_{afst}\,q_{afst}^2. \tag{16}$$

We argue that unobserved managerial efforts and/or marketing strategies in ξ_{ft} are correlated with the input prices that airlines bargain with input providers. The contribution ξ_{ft} is modeled as a linear function of wages and fuel prices faced by each airline f during period t; in other words,

$$\log\xi_{ft} = b\,\log\mathtt{Wage}_{ft} + (1-b)\,\log\mathtt{PFuel}_{ft} + \xi_f + \mathtt{Quarter}_t + \mathtt{Year}_t, \tag{17}$$

where ξ_f is a carrier fixed effect, and $\texttt{Quarter}_t$ and $\texttt{Year}_t$ are quarter and year time dummies. The property of linear homogeneity of degree one in input prices guarantees that the corresponding coefficients sum to 1.

The $3,575$ observation data used to estimate cost only entails $1,298$ different segments (see Table 1), which prevents us from estimating reliable individual segment fixed effects μ_s for every segment. To circumvent this difficulty we assume that

$$\mu_s = d_0 + d_1 \texttt{Distance}_s + d_2 \texttt{Temperature}_s, \tag{18}$$

where $\texttt{Distance}_s$ is the segment length measured as the geographical distance between two cities, and $\texttt{Temperature}_s$ is the average temperature at the departure and arrival cities over the whole sample window.

The cost function to be estimated obtains by reintroducing (17) and (18) into (16). The final expression of this function is

$$\log C_{aft} = \log b_0 + b \log \texttt{Wage}_{ft} + (1-b) \log \texttt{PFuel}_{ft}$$

$$+ \log \left[\sum_s \left(1 + d_1^* \texttt{Distance}_s + d_2^* \texttt{Temperature}_s \right) n_{afst} \, q_{afst}^2 \right]$$

$$+ \xi_f + \nu_g + \texttt{Quarter}_t + \texttt{Year}_t + \varepsilon_{aft}, \tag{19}$$

where ε_{aft} is an error term. To get this expression, we have made three main simplifications. First the constant $\log b_0$ replaces the sum of the two constants $\log \beta_0$ and $\log d_0$ that cannot be estimated separately. Second, we normalize the coefficients $d_1^* = d_1/d_0$ and $d_2^* = d_2/d_0$. Third, in the data airlines have preferences for specific aircraft types that gives rise to a high correlation between the airlines and aircraft fixed effects ξ_f and ν_a. This prevents us to keep track of both airlines and aircraft unobserved heterogeneity at the detailed level of the aircraft type. We therefore work with a more aggregated aircraft group by clustering the 29 different aircraft types into 12 groups referring to model characteristics and carrier[9]. This provides us with an aircraft group g fixed effect ν_g that replaces the original aircraft fixed effect ν_a.

Table 16 reports the maximum likelihood estimates of the parameters of the cost function (19) for three variants. The results are very similar in each case, except from the impact of distance between the two cities located at the endpoints of the segment. Costs increase with input prices; they are also higher on cold weather segments and during the colder first and fourth quarters. There is a positive significant correlation between route distance and temperature in our data that makes the impact of distance on cost in variant (2) magnified when temperature is omitted. The likelihood of the variant (1), which only differs from the one in (3) by excluding aircraft group fixed effects on the segment, shows

---

[9]Aircraft in the same cluster belong to the same generation of models and have similar size. For example, Boeing 737-300, Boeing 737-400 and Boeing 737-500 are allocated to the same cluster, whereas next generation larger Boeing 737-800 and Boeing 737-900 are in another cluster. There remain small clusters with rare aircraft types like Aerospatiale/Aeritalia ATR-72 or Saab-Fairchild 340/B.

Table 16: ESTIMATED COST FUNCTION

	Dependent variable: $\log C_{aft}$		
	(1)	(2)	(3)
Constant ($\log b_0$)	2.513	1.127	2.182
	(0.34)	(0.31)	(0.45)
$\log$ Wage (b)	0.558	0.597	0.602
	(0.08)	(0.07)	(0.07)
Distance (d_1^*)	0.025	0.709	0.205
	(0.02)	(0.37)	(0.11)
Temperature (d_2^*)	-0.006		-0.006
	(0.00)		(0.00)
Quarter 2	-0.170	-0.157	-0.164
	(0.01)	(0.01)	(0.01)
Quarter 3	-0.199	-0.180	-0.193
	(0.01)	(0.01)	(0.01)
Quarter 4	-0.096	-0.086	-0.089
	(0.01)	(0.01)	(0.01)
Aircraft group f.e. (ν_g)	No	Yes	Yes
Airlines f.e. (ξ_f)	Yes	Yes	Yes
Year f.e. (Year_t)	Yes	Yes	Yes
Number of observations	3,575	3,575	3,575
Log-Likelihood	-1010	-453	-407

Note: standard errors are in parenthesis

the importance of taking into account the aircraft type in airline costs. In the absence of control for the aircraft type, we find that a greater distance does not involve higher costs for the operating airlines. The expected positive impact is recovered once the control is introduced, which reflects the fact that airlines allocate specific aircraft types conditionally on the length of each segment.

E   Demand estimation

We assume that the two parameters entering the linear demand function depend on route distance and population size of endpoint cities, $\texttt{Pop}^1_{st}$ and $\texttt{Pop}^2_{st}$ (with $\texttt{Pop}^1_{st} \leq \texttt{Pop}^2_{st}$). We also assume that the slope γ_{st} may directly depend on time. The expressions of the intercept and the slope rewrite as

$$\gamma^0_{st} = \alpha_0 + \alpha_1 \texttt{Pop}^1_{st} + \alpha_2 \texttt{Pop}^2_{st}, \tag{20}$$

and

$$\gamma_{st} = \alpha_3 + \alpha_4 \texttt{Distance}_s + \alpha_5 \texttt{Pop}^1_{st} + \alpha_6 \texttt{Pop}^2_{st} + \texttt{Quarter}_t + \texttt{Year}_t. \tag{21}$$

The demand equation that we estimate is (11) with intercept and slope given by (20) and (21), respectively.

The (one quarter) lagged price can be used to deal with the joint determination of the price P_{st} and the quantity Q_{st}. Table 17 presents the demand function estimated from the 379 selected routes. The price variable is the route average ticket price in columns 1 and 2 and the median price in columns 3 and 4. The specification in columns 2 and 4 fits the specification described in (7), (20) and (21). For comparison purposes, we report in columns 1 and 3 the results in the case where the slope of the demand function is assumed to be independent from route characteristics.

We observe a higher demand for routes linking densely populated cities. In such markets demand displays lower price sensitivity. The greater distance between origin and destination points makes substitution with alternative transport facilities more difficult as α_4 is positive. Moreover, α_5 and α_6 are both positive, which suggests that consumers departing from or arriving to cities with a larger population (i.e., hubs) are less sensitive to price fluctuations.

The contribution of demand to market stability captured by the index $S(\mathbf{q}^*)$ for route s at time t only relies on the slope $\delta = \delta_{st} = -1/\gamma_{st}$ in Table 17. The results are very similar when we use average and median lagged prices as instruments. In the sequel we shall use the specification in column 2 referring to the average price. To provide the reader with an order of magnitude for this slope from this specification, we obtain $\gamma_{st} = -0.986 + 0.168 \times 0.783 + 0.004 \times 1.785 + 0.051 \times 6.419 \simeq -0.519$ plus quarter and year fixed effects in the average segment in Table 3. Including the time fixed effects one gets $\gamma_{st} = -0.621$: a 10 USD increase in the fare yields $6,210$ passengers less in the route (see Oum et al., 1992, for an overview of transport demand elasticity estimates).

Table 17: ESTIMATED DEMAND FUNCTION

| | Dependent variable: number of passengers | | | |
	(1)	(2)	(3)	(4)
Intercept γ_{st}^0 – eq. (20)				
Constant α_0	133.646	207.210	126.531	196.328
	(16.24)	(16.58)	(16.26)	(16.62)
Pop[1] (α_1)	18.116	19.037	18.829	18.518
	(2.45)	(2.45)	(2.46)	(2.46)
Pop[2] (α_2)	47.003	22.234	46.288	17.078
	(3.02)	(3.70)	(3.03)	(3.82)
Slope γ_{st} – eq. (21)				
Constant (α_3)	-0.505	-0.986	-0.583	-1.060
	(0.02)	(0.04)	(0.02)	(0.04)
Distance (α_4)		0.168		0.148
		(0.02)		(0.053)
Pop[1] (α_5)		0.004		0.013
		(0.00)		(0.00)
Pop[2] (α_6)		0.051		0.078
		(0.01)		(0.01)
Year f.e.	Yes	Yes	Yes	Yes
Quarter f.e.	Yes	Yes	Yes	Yes
R^2	0.176	0.204	0.170	0.198
Obs.	20808	20808	20808	20808
F test (route f.e.)	1003.8	982.9	1008.7	978.1

Note: fixed-effects (within) IV regression. Standard errors are in parenthesis

(1), (2) - mean price is employed, (3), (4) - median price is employed

F   Relevant market

In order to highlight the possible biases associated with missing competition, we examine a simple variant of our theoretical setup where the relevant market consists of two substitutable items. The demand for item m $(m = 1, 2)$ is $P^m(Q^m, Q^{-m})$, where Q^m and Q^{-m} stand for the aggregate demands for the two items. On the supply side we assume that the set of firms producing the two items are disjoint. This is a strong assumption if the missing information about the relevant market mostly consists of indirect flights as traditional carriers often offer both direct and indirect flights. The assumption could be better suited on short distance routes which involve few alternatives to direct flights, with airline services competing against rail transportation or other inland transport modes for instance. In addition, we simplify the exposition by working under two symmetry assumptions:

Assumption A1. Symmetry of demand: $P^{-m}(Q^{-m}, Q^m) = P^m(Q^{-m}, Q^m)$ for all (Q^{-m}, Q^m) and all m.

Assumption A2. Symmetry of supply: every firm f producing item 1 has a mirror firm $F + f$ producing item 2, i.e., $\sigma_f = \sigma_{F+f}$ for every $f = 1, \ldots, F$, with F being the total number of airlines producing any given item.

Firm f in market m produces q_f that maximizes its profit given the aggregate production Q^m_{-f} of the other firms in market m and the aggregate production Q^{-m} in the other market. Its best-response can be written $q_f = R^m_f(Q^m_{-f}, Q^{-m})$, which is decreasing with its two arguments if the two items are substitutes. In view of Assumptions A1 and A2, we focus attention on a symmetric Nash equilibrium where $q^*_f = q^*_{F+f}$ for every $f \leq F$. In equilibrium the aggregate output is Q^* in each sub-market. The iterated process of elimination of dominated strategies in the neighborhood of such an equilibrium is now driven by the partial derivatives

$$R'_{f1}(\mathbf{q}^*) = \frac{\partial R^m_f}{\partial Q^m_{-f}}\left(Q^*_{-f}, Q^*\right), \quad \text{and } R'_{f2}(\mathbf{q}^*) = \frac{\partial R^m_f}{\partial Q^{-m}}\left(Q^*_{-f}, Q^*\right)$$

of the best-response function of firms f and $F + f$ with respect to its first and second argument, respectively. In this setup the stability condition given in Proposition 1 becomes:

Proposition 4. *A symmetric Nash equilibrium is locally stable if and only if*

$$\sum_{f \leq F} \frac{R'_{f1}(\mathbf{q}^*) + R'_{f2}(\mathbf{q}^*)}{R'_{f1}(\mathbf{q}^*) - 1} < \frac{1}{2}. \tag{22}$$

Proof. The new $\mathbf{B}$ matrix that governs the iterated process of elimination of dominated strategies is a $2F \times 2F$ matrix whose f-th row, $f \leq F$, is 0 at f-th column, is $R'_{f2}(\mathbf{q}^*)$ in every column $j > F$ and $R'_{f1}(\mathbf{q}^*)$ otherwise. Similarly, its $(F + f)$-th row is 0 at $F + f$-th

column, is $R'_{f2}(\mathbf{q}^*)$ in every column $j \leq F$ and $R'_{f1}(\mathbf{q}^*)$ otherwise. Matrix $\mathbf{B}$ is contracting if and only if the spectral radius of the positive matrix $-\mathbf{B}$ is lower than 1. Let e some eigenvalue e of $-\mathbf{B}$ and $\mathbf{v}$ be the associated $2F$-eigenvector $(v_1, \ldots, v_{2F})$. From $e\mathbf{v} = -\mathbf{B}\mathbf{v}$, we have: for all $f \leq F$,

$$ev_f - R'_{f1}v_f = -\sum_{k \leq F} R'_{f1}v_k - \sum_{F < k \leq 2F} R'_{f2}v_k$$

$$\Leftrightarrow v_f = -\frac{R'_{f1}}{e - R'_{f1}} \sum_{k \leq F} v_k - \frac{R'_{f2}}{e - R'_{f1}} \sum_{F < k \leq 2F} v_k$$

and

$$v_{F+f} = -\frac{R'_{f2}}{e - R'_{f1}} \sum_{k \leq F} v_k - \frac{R'_{f1}}{e - R'_{f1}} \sum_{F < k \leq 2F} v_k.$$

All the derivatives of the best-response functions are evaluated at $\mathbf{q}^*$. Summing over firms yields

$$\sum_{f \leq F} v_f = -\sum_{f \leq F} \frac{R'_{f1}}{e - R'_{f1}} \sum_{k \leq F} v_k - \sum_{f \leq F} \frac{R'_{f2}}{e - R'_{f1}} \sum_{F < k \leq 2F} v_k$$

and

$$\sum_{F < f \leq 2F} v_{F+f} = -\sum_{F < f \leq 2F} \frac{R'_{f2}}{e - R'_{f1}} \sum_{k \leq F} v_k - \sum_{F < f \leq 2F} \frac{R'_{f1}}{e - R'_{f1}} \sum_{F < k \leq 2F} v_k.$$

The symmetry properties of $\mathbf{B}$ imply that the eigenvectors are such that $v_f = v_{F+f}$ for all $f \leq F$. Hence the two previous equations reduce to

$$\sum_{f \leq 2F} v_f = -\sum_{f \leq 2F} \frac{R'_{f1} + R'_{f2}}{e - R'_{f1}} \sum_{k \leq 2F} v_k.$$

Eigenvalues e of $-\mathbf{B}$ thus are solutions to

$$G(e) \equiv -\sum_{f \leq 2F} \frac{R'_{f1} + R'_{f2}}{e - R'_{f1}} - 1 = 0$$

The function G is continuous decreasing for all $e \geq 0$, with $G(0) > 0 > -1 = G(+\infty)$. There is consequently a unique $e \geq 0$ solution to $G(e) = 0$. This is the spectral radius. Because G is decreasing, this eigenvalue is lower than 1 if and only if $G(1) < 0$, or equivalently,

$$-\sum_{f \leq 2F} \frac{R'_{f1} + R'_{f2}}{1 - R'_{f1}} - 1 < 0.$$

The result follows from $R'_{fm} = R'_{(F+f)m}$ for every $f \leq F$ and $m = 1, 2$. ■

Proposition 4 departs from Proposition 1 in two respects. First, the new stability index that appears in the left-hand side of (22), computed using information from sub-market m

only, now accounts for the cross derivative $R'_{f2}(\mathbf{q}^*)$ because firms active in sub-market m need to predict the behavior of firms in the other sub-market. With two substitutable items, $R'_{f2}(\mathbf{q}^*) \leq 0$ and so the new index tends to be higher than the one derived in Proposition 1. The need of predicting the behavior of competitors in the other sub-market makes stability more difficult to achieve.

Second, in the right-hand side, there is a reduction from 1 to 1/2 in the route threshold S^*_s above which stability is lost. In a route s where $S^*_s < S_s < 1$, the equilibrium is unstable even though the stability index S_s computed in Proposition 1 is below 1. A threshold S^*_s below 1 can be used as a signal that the relevant market of route s is larger than the observed one.

In practice, in the situation where the econometrician has no a priori knowledge about how far the relevant market goes beyond the observed one, the derivative $R'_{f2}(\mathbf{q}^*)$ cannot be computed. In the empirical illustration we neglect this derivative by setting $R'_{f2}(\mathbf{q}^*)$ equal to 0, and continue to refer to the stability index given in Proposition 1. Rewriting (22) as

$$\sum_{f \leq F} \frac{R'_{f1}(\mathbf{q}^*)}{R'_{f1}(\mathbf{q}^*) - 1} < \frac{1}{2} - \sum_{f \leq F} \frac{R'_{f2}(\mathbf{q}^*)}{R'_{f1}(\mathbf{q}^*) - 1} \tag{23}$$

shows that we underestimate the true right-hand side of (22) when we set $R'_{f2}(\mathbf{q}^*)$ to 0.

G   Stability index threshold estimate

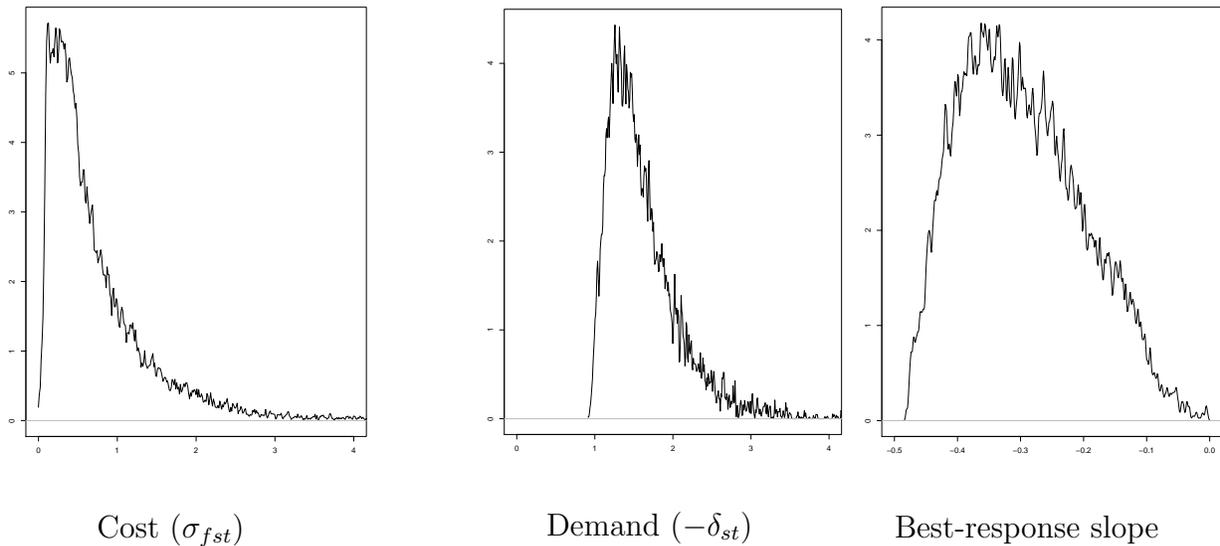| Cost (σ_{fst}) | Demand $(-\delta_{st})$ | Best-response slope |

Figure 2: Probability density functions

The left-panel of Figure 2 depicts the density of $\hat{\sigma}_{fst}$ at the airlines $\times$ segment $\times$ quarter level. Its distribution is positively skewed with a long right tail (Table 4 shows that the

median estimated value of this parameter is 0.515 whereas its average of 0.762). The statistics $\hat{\sigma}_{fst}$ is inversely related to the marginal cost. Hence Figure 2 shows that the marginal cost for transporting one additional passenger quickly increases with the total number of transported passengers. The small subset succeeding to contain marginal costs consists of the largest airlines: we find that the average number of transported passengers is 3.6 times higher for observations in the last quartile of the $\hat{\sigma}_{fst}$ distribution ($\sigma_{fst} \geq 0.965$).

The shape of the density of $\hat{\delta}_{st}$ shown in the middle panel in Figure 2 is quite different. Price sensitivity of demand indeed appears symmetrically distributed and displays a high concentration around an average sensitivity of $1/0.621 = 1.609$.

The homogeneity of demand behavior (the low variability of $\hat{\delta}_{st}$) across segments makes the distribution of the slope of the best-response $R'_{fst}(\mathbf{q}^*)$ in (9) at the airlines $\times$ segment $\times$ quarter level to be mostly driven by cost heterogeneity (the high variability of $\hat{\sigma}_{fst}$). It is reported in the right panel in Figure 2. Airlines are found to display significant inertia in how they react to the production of others. The average slope equals -0.299, i.e., an average airlines would only transport 30 passengers less when its competitors are expected to transport 100 additional passengers. Short-run inertia is plausible in the airline industry because the number of passengers can be adjusted within the limit of the route transportation seat capacity, which is mostly fixed over a quarter. The theory predicts that this inertia favors stability of the Nash equilibrium, but we know from (5) that stability relies on the sum of all the slopes of the best-response functions of the airlines active on the segment. Despite individual inertia, stability would be lost with more than 4 identical competing airlines characterized by a -0.299 average best-response slope.

Figure 3 depicts in solid black the density of the spread between actual and Nash production at the more disaggregated segment $\times$ airlines $\times$ period level. This distribution combines situations where airlines produce below the equilibrium number of passengers and situations where they instead produce above it. The first (resp. last) type of situations reflects an underestimated (resp., overestimated) expected price. Figure 3 highlights some symmetry between under and overestimated prices when forecast errors are small (the spread stands below 0.25). For larger values of the spread, the density corresponding to an overestimated production of others / an underestimated price is depicted in dashed blue. Figure 3 clearly shows limited errors in this case, compared to the case where airlines underestimate the production of others / overestimate prices (depicted in dashed red). If one considers a predicted low price as a symptom of firms' pessimism, Figure 3 suggests limited forecasting errors from pessimistic airlines. This would accord with the recent findings based on surveys about financial analysts, CEO or Chief Financial Officers expectations or proxies for managerial expectations that errors relate to overconfidence and optimism (see, e.g., Ma, Sraer and Thesmar (2018)).
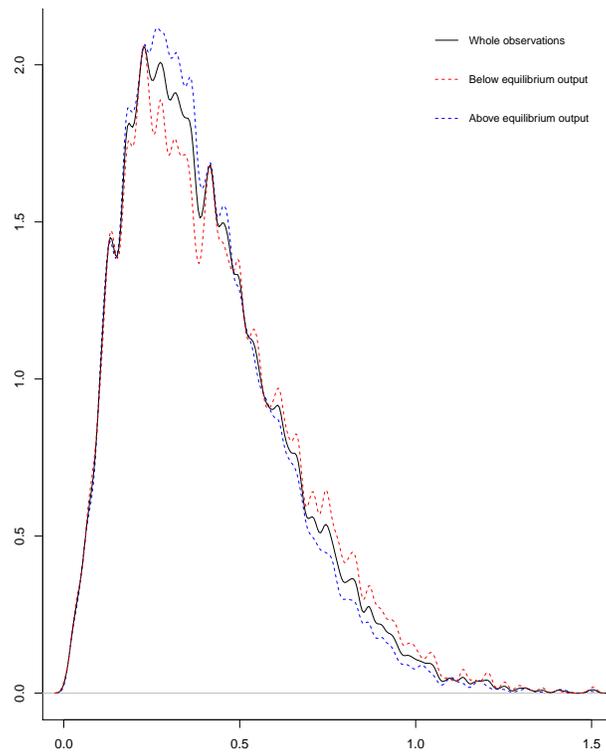
Figure 3: Spread density distribution

H Spread regimes from the EM algorithm

The EM algorithm is designed to cluster points (the various spreads in our setup) that do not come with any specific label (a low or high spread regime). We assume that the distribution of the spread Δ_{st} reproduced in plain black in Figure 4 arises from a mixture of two Gaussian distributions: the first distribution is associated with a low spread regime (where the Nash equilibrium is a plausible outcome of competition); the second distribution is associated with a high spread regime. The EM algorithm aims at generating the probabilities that a spread point originates from any of the two regimes. This probabilities are then used to derive the individual market threshold we are interested in.

In order to initialize the EM algorithm, we compute a preliminary allocation of all the spread points to two different sets with the help of a standard k-means clustering technique. The average spread in the resulting low spread group G_1 is 0.26 (standard deviation is 0.108). The density of the corresponding Gaussian distribution is depicted in dotted red in Figure 4. Similarly, the average spread in the high spread regime G_2 is 0.64 (standard deviation is 0.169), and the density of the corresponding Gaussian distribution is depicted in dotted blue in Figure 4. In the low (resp., high) spread regime the mean to standard deviation ratio of the spread equals 2.4 (resp., 3.79), which highlights a much greater concentration of the departures from the Nash equilibrium in the low spread regime.

The k-means clustering yields an average probability of $\pi_{st}(0) = 0.65$ that a specific spread drawn randomly in the sample originates from the low spread regime. Given this a priori probability and the spread mean and standard error in each regime, we can compute from Bayes's rule the probability $\pi_{st}(1)$ that Δ_{st} is actually drawn from the low spread regime for all s and t. Then, given these a posteriori probabilities, we can compute the maximum likelihood estimators for the means and standard errors of the two regimes. The new Gaussian distributions are used to revise $\pi_{st}(1)$ into $\pi_{st}(2)$ for every s and t according to Bayes's rule, which allows us to initiate another step of estimation for the two moments of the two Gaussian distributions. The EM algorithm repeats these steps until convergence.[10]

In Table 18 we report the moments of the two Gaussian distributions for two variants.[11] In both cases, the spread distribution for the subsample of routes with a stability index above 1, that we know are part of the high spread regime, is the same as the spread distribution in the high spread regime. The mean spread in the low spread regime is 0.25; it is twice as high in the high spread regime. The dispersion of the spread is also twice

---

[10]The EM algorithm stabilizes in a local maximum for the likelihood. In our setup it always converges to the same outcome.

[11]In the first variant, we a priori require that the spread mean and standard error in the high spread regime are respectively equal to the empirical mean and standard error of the spread among the subset of points Δ_{st} with a stability index S_{st} above 1. That is, we only apply the EM algorithm to estimate the first two moments of the Gaussian distribution of the spread in the low spread regime. This variant thus a priori imposes the theoretical consistency requirement that observations with a stability index above 1 are drawn from the same probability distribution as those falling in the high spread regime, even though the associated stability index is below 1. In the second variant, we impose no constraint on the moments of the Gaussian distribution of the spread in the high spread regime.
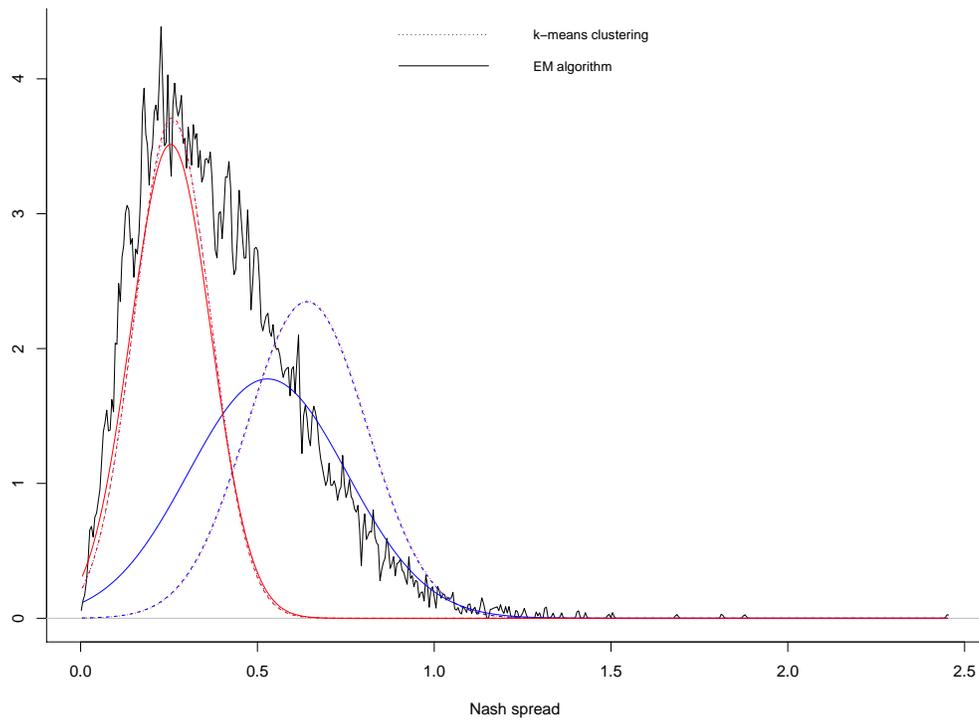
Figure 4: Spread Gaussian mixture from the EM algorithm

as high in the high spread regime (0.11 versus 0.22). The low spread regime thus will be characterized by dampened fare and volume of transported passengers fluctuations. The two Gaussian distributions are depicted in red (for the low spread regime) and blue (for the high spread regime) in Figure 4. It is clear from this figure that observations with a very low (resp. high) spread are almost surely allocated to the low (resp. high) spread regime. However the algorithm fails to identify clearly the regime of observations with an intermediate spread located around 0.4.

The final probability π_{st} that Δ_{st} is drawn from the low spread regime distribution ranges from 0 to 0.83. In Figure 5 we plot the within route average probability and its standard error for each of our 301 route sample. In the horizontal axis routes are ranked in the order of increasing within route average probabilities. The routes with low or high probabilities display the greatest spread concentration over time. The uncertainty about the relevant regime of airline routes with an average probability around 0.4 translates into quarters where the probability of a ruling low spread regime varies a lot over time.

Table 18: Nash spread Gaussian mixture

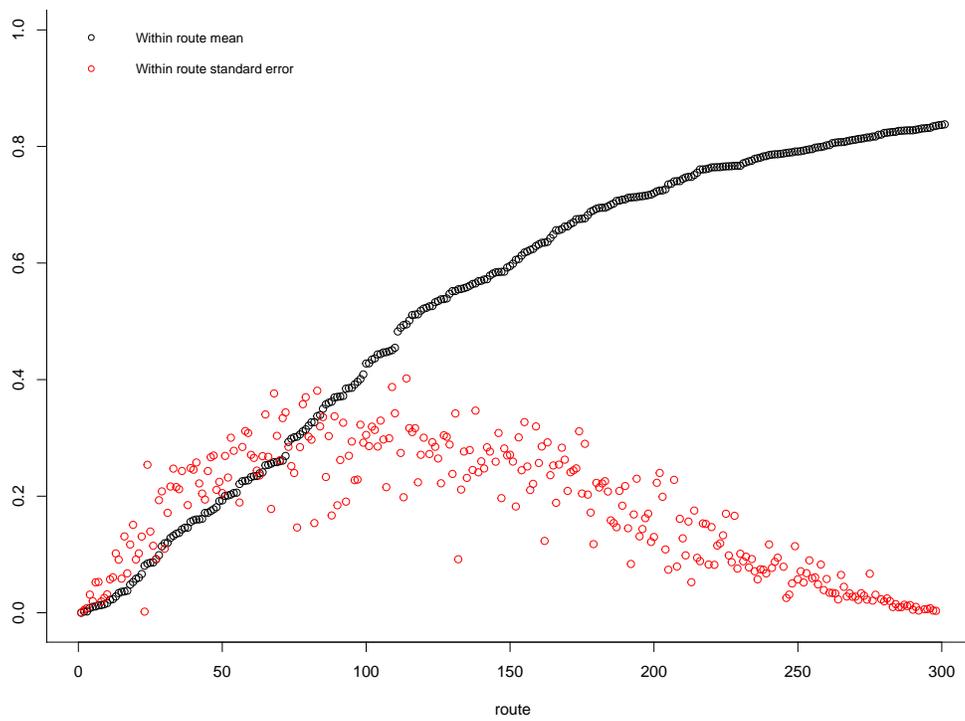| | Low spread regime | | High spread regime | |
Variant of the EM algorithm	Mean	Standard error	Mean	Standard error
Constrained	0.2615397	0.1162918	0.5156821	0.2419057
Unconstrained	0.2546883	0.1135389	0.5286944	0.2248107

Number of observations: 7422

Figure 5: Probability of the low spread regime