DISCUSSION PAPER SERIES

DP13346
(v. 2)

MARKET STRUCTURE AND COMPETITION IN AIRLINE MARKETS

Federico Ciliberto, and

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

MARKET STRUCTURE AND COMPETITION IN AIRLINE MARKETS

Federico Ciliberto, and

Discussion Paper DP13346
First Published 30 November 2018
This Revision 29 February 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

MARKET STRUCTURE AND COMPETITION IN AIRLINE MARKETS

Abstract

We provide an econometric framework for estimating a game of simultaneous entry and pricing decisions while allowing for correlations between unobserved cost and demand shocks. We use our framework to account for selection in the pricing stage. We estimate the model using data from the US airline industry and find that not accounting for endogenous entry leads to biased estimation of demand elasticities. We simulate a merger between American and US Airways and find that product repositioning and post-merger outcomes depend on how we model the characteristics of the merged firm as a function of the pre-merger firms' characteristics.

JEL Classification: C35, C51, D43, L13, L41, L44

Keywords: multiple equilibria, market power, Self-selection, Entry, oligopoly, market structure, merger

Federico Ciliberto - ciliberto@virginia.edu
University of Virginia and CEPR

- murryc@bc.edu
Boston College

- elietamer@fas.harvard.edu
Harvard University

Market Structure and Competition in Airline Markets *

Federico Ciliberto† University of Virginia
Charles Murry‡ Boston College
Elie Tamer§ Harvard University

February 14, 2020

Abstract

We provide an econometric framework for estimating a game of simultaneous entry and pricing decisions while allowing for correlations between unobserved cost and demand shocks. We use our framework to account for selection in the pricing stage. We estimate the model using data from the US airline industry and find that not accounting for endogenous entry leads to biased estimation of demand elasticities. We simulate a merger between American and US Airways and find that product repositioning and post-merger outcomes depend on how we model the characteristics of the merged firm as a function of the pre-merger firms' characteristics.

*We thank Steve Berry, Timothy Bresnahan, Ambarish Chandra, Paul Grieco, John Panzar, Wei Tan, Randal Watson, and Jon Williams for insightful suggestions. We also thank participants at the Southern Economic Meetings in Washington (2005 and 2008), the 4th Annual CAPCP Conference at Penn State University, 2009, the JAE Conference at Yale in 2011, and the DC IO Conference in 2014, where early drafts of this paper were presented. Seminars participants at other institutions provided useful comments. Finally, we want to especially thank Ed Hall and the University of Virginia Alliance for Computational Science and Engineering, who have given us essential advice and guidance in solving many computational issues. We also acknowledge generous support of computational resources from XSEDE through the Campus Champions program (NSF-Xsede Grant SES150002).

†Department of Economics, University of Virginia, ciliberto@virginia.edu. Federico Ciliberto thanks the CSIO at Northwestern University for sponsoring his visit at Northwestern University. Research support from the Bankard Fund for Political Economy at the University of Virginia and from the Quantitative Collaborative of the College of Arts and Science at the University of Virginia is gratefully acknowledged.

‡Department of Economics, Boston College, Chestnut Hill, MA, murryc@bc.edu.

§Department of Economics, Harvard University, elietamer@fas.harvard.edu

1 Introduction

We estimate a simultaneous, static, complete information game where economic agents make both discrete and continuous choices. We study airlines that strategically decide whether to enter into a market *and* the prices they charge if they enter. Our aim is to provide a framework for combining both entry and pricing into one empirical model that allows us: i) to account for selection of firms into serving a market and, more importantly, ii) to allow for market structure to adjust as a response to counterfactuals, such as mergers.

Generally, firms self-select into markets that best match their observable and unobservable characteristics. For example, high quality products command higher prices, and it is natural to expect high quality firms to self-select into markets where there is a large fraction of consumers who value high-quality products. Previous work has taken the market structure of the industry, defined as the identity and number of its participants (be they firms or, more generally, products or product characteristics) as exogenous when estimating the parameters of the demand and supply relationships.[1] That is, firms, or products, are assumed to be randomly allocated into markets. This assumption has been necessary to simplify the empirical analysis, but it is not always realistic.

Non-random allocation of firms across markets can lead to self-selection bias in the estimation of the parameters of the demand and cost functions. Existing instrumental variables methods that account for endogeneity of prices do not resolve this selection problem in general.[2] Potentially biased estimates of the demand and cost functions can then lead to mis-measuring demand elasticities, and consequently market power. This is problematic because correctly measuring market power and welfare is crucial for the application of antitrust policies and for a full understanding of the competitiveness of an industry. For example, if the bias is such that we infer firms to have more market power than they actually have, the antitrust authorities may block the merger of two firms that would improve total welfare, possibly by reducing an excessive number of products in the market. Importantly, allowing

[1] See (Bresnahan, 1987; Berry, 1994; Berry, Levinsohn, and Pakes, 1995) and the large subsequent literature in IO that uses this methodology.

[2] This point was previously made by Olley and Pakes (1996) for the estimation of production functions.

for entry (or product variety) to change as a response to a merger is important. For example, when a firm (or product) exits due to consolidation from a merger, it is likely that other firm may now find it profitable to enter (or to offer new products in the market). Our empirical framework allows for such adjustments.

More generally, our model can also be viewed as a multi-agent version of the classic selection model (Gronau, 1974; Heckman, 1976, 1979). In the classic selection model, a decision maker decides whether to enter the market (e.g. work), and is paid a wage conditional on working. When estimating wage regressions, the selection problem deals with the fact that the sample is selected from a population of workers who found it "profitable to work." Here, firms (e.g. airlines) decide whether to enter a market and then, conditional on entry, they choose prices. Our econometric model accounts for this selection when estimating demand and supply equations, as in the single-agent selection model.

Our model consists of the following conditions: i) entry inequalities that require that, in equilibrium, a firm must be making non-negative profit in each market that it serves; ii) demand equations derived from a discrete choice model of consumer behavior; iii) pricing first-order-conditions, which can be formally derived under the postulated firm conduct. We allow for all firm decisions to depend upon market- and firm-specific random variables (structural errors) that are observed by firms but not the econometrician. In equilibrium firms make entry and pricing decisions such that all three sets of conditions are satisfied.

A set of econometric problems arises when estimating such a model. First, there are multiple equilibria associated with the entry game. Second, prices are endogenous as they are associated with the optimal behavior of firms, which is part of the equilibrium of the model. Finally, the model is nonlinear and so poses a heavy computational burden. We combine the methodology developed by Tamer (2003) and Ciliberto and Tamer (2009) (henceforth CT) for the estimation of complete information, static, discrete entry games with the widely used methods for the estimation of demand and supply relationships in differentiated product markets (see Berry, 1994; Berry, Levinsohn, and Pakes, 1995, henceforth BLP). We simultaneously estimate the parameters of the entry model (the observed fixed costs and the variances of the unobservable components of the fixed costs) and the parameters of the demand and supply relationships.

To estimate the model we use cross-sectional data on the US airline industry.[3] The data are from the second quarter of 2012's Airline Origin and Destination Survey (DB1B). We consider markets between US Metropolitan Statistical Areas (MSAs), which are served by American, Delta, United, USAir, Southwest, and low cost carriers (e.g. Jet Blue). We observe variation in the identity and number of potential entrants across markets.[4] Each firm decides whether or not to enter and chooses the price in that market.[5] The other endogenous variable is the number of passengers transported by each firm. The identification of the three conditions relies on variation in several exogenous explanatory variables, whose selection is supported by a rich and important literature, for example Rosse (1970), Panzar (1979), Bresnahan (1989), and Schmalensee (1989), Brueckner and Spiller (1994), Berry (1990), Ciliberto and Tamer (2009), Berry and Jia (2010), and Ciliberto and Williams (2014).

We begin our empirical analysis by running a standard GMM estimation (see Berry, 1994) on the demand and pricing first order conditions and comparing that to our proposed methodology with exogenous entry. Next, we estimate the model with endogenous entry using our methodology and compare the results with the exogenous entry results. We find that allowing for endogenous entry, the price coefficient in the demand function is estimated to be closer to zero than the case of exogenous entry, and markups are substantially larger.[6] Next, we use our estimated model to simulate the merger of two airlines in our data: American and US Airways.[7] Typical merger analysis involves predicting changes in market power and prices *given* a particular market structure using diversion ratios based on pre-merger

[3]We also illustrate our methodology by conducting a numerical exercise, see the Appendix E.

[4]A market is defined as a unidirectional pair of an origin and a destination airport, as in Borenstein (1989), Berry and Jia (2010), and Ciliberto and Williams (2014). An airline is considered a potential entrant if it is serving at least one market out of both of the endpoint airports. See the Appendix C for more details.

[5]In practice we use the median of the prices observed in a market in a quarter, where each individual price is weighted by the number of passengers on that ticket.

[6]The selection problem could lead to overestimation or underestimation of demand elasticities, and thus markups, depending on the covariance of demand, marginal cost, and fixed costs unobservables. We illustrate this dependence in the numerical exercise in Appendix E.

[7]The two firms merged in 2013 after settling with the Department of Justice.

market shares, or predictions from static models of product differentiation (see Nevo, 2000). Our methodology allows us to simulate a merger allowing for equilibrium changes to market structure after a merger, which in turn may affect equilibrium prices charged by firms.

There are several findings from the merger analysis, which depend, crucially, on on how we model the characteristics of the post-merger firm as a function of the pre-merger firms' characteristics. We consider four different scenarios. First, we assume that the merged firm takes on the best characteristics, both observed and unobserved, of the two pre-merger firms, and call this the *Best Case Scenario*. Then we simulate two sub-cases, one in which the merged firm only takes the best observable characteristics between the two pre-merger firms and keeps the surviving firm's unobservables, and another where we draw a new unobservable for the new merged firm. Lastly, we consider a case where the surviving firms inherits the average observed and unobserved characteristics between the two pre-merged firms, or what we call the *Average Case* scenario.

We find that under all four scenarios there is substantial post-merger entry and exit among the surviving airlines, especially for the surviving merged airline, American Airlines. For the scenario in which we assume the most merger efficiencies, the average price across all markets increases slightly, but consumer welfare also substantially rises due to post-merger entry from the new merged airline. Of course, there is a lot of heterogeneity across the types of markets, so we look at the effects of the merger on markets that share particular pre-merger market structures. For example, we find that the the merged airline would enter previously unserved markets with a likelihood between 49 and 53 percent, and prices would fall by between 8.4 and 6.0 percent in markets that were previously only served by an AA and US duopoly. In contrast, when we assume that the post-merger airline takes the average characteristics from AA and US (the *Average Case* scenario), total consumer welfare does not increase substantially, and may even fall. We find that the the merged airline would enter previously unserved markets with a likelihood of only between 16 and 18 percent and prices would rise by between 7.2 and 9.2 percent in markets that were previously only served by aAA and US duopoly. Clearly, assumptions about merger efficiencies matter – not just for pricing pressure, but also for post-merger entry/exit. We systematically document the these types of effects through many more pre-merger market structures.

Finally, we investigate the effects of the merger in markets originating or ending in DCA, which were of concern for antitrust authorities because both of the merging parties had a very strong incumbent presence. When we maintain that AA experiences large efficiencies, we predict that prices would decrease even though concentration decreases. In the other cases we find that prices would increase slightly along with concentration. In all cases, low-cost carriers are not likely to replace the exiting US Airways, which was a major concern for the DOJ and resulted in landing slot divestitures by the merging party.

There is other important work that has estimated static models of competition while allowing for market structure to be endogenous. Reiss and Spiller (1989) estimate a monoply model of airline competition. In contrast, we allow for multiple firms to choose whether or not to serve a market. Cohen and Mazzeo (2007) assume that firms are symmetric within types, as they do not include firm specific observable and unobservable variables. In contrast, we allow for very general forms of heterogeneity across firms. Berry (1999), Draganska, Mazzeo, and Seim (2009), Pakes et al. (2015) (PPHI), and Ho (2008) assume that firms self-select themselves into markets based on observable characteristics by imposing restrictions on information about the unobservables. In contrast, we focus on the case where firms self-select themselves into markets that better match their observable and *unobservable* characteristics. There are two recent papers that are closely related to ours. Eizenberg (2014) estimates a model of entry and competition in the personal computer industry. Estimation relies on a timing assumption (motivated by PPHI) requiring that firms do not know their own product quality or marginal costs before entry, which limits the amount of selection captured by the model.[8] Similar timing assumptions are made by other papers as well, such Sweeting (2013),

[8]If we are willing to make this timing assumption, there would not be a selection on *unobservables*, because the firm would only observe the demand and marginal cost shock after entering. In markets where there is a long lag between the entry/characteristic decision and the pricing decision, such as car manufacturing or computer manufacturing, such timing assumption would seem a reasonable assumption. In the airline industry, firms can enter and exit market quickly, as long as they have access to gates. So the timing assumption is less plausible. Generally, a prudent approach would be to allow for correlation in the unobservables, and if that is non zero, then we could conclude that the timing assumption would be less acceptable.

Lee (2013), Jeziorksi (2014b), Jeziorksi (2014a) in dynamic empirical games; and Fan (2013) and Fan and Yang (2017) in static games.[9] Another paper that is closely related to ours is Li et al. (2017), who estimate a model of service selection (nonstop vs connecting) and price competition in airline markets, but only consider sequential move equilibria. In addition, Li et al. (2017) do not allow for correlation in the unobservables, which is a key determinant of self-selection that we investigate in this paper.

The paper is organized as follows. Section 2 presents the methodology in detail in the context of a bivariate generalization of the classic selection model, providing the theoretical foundations for the empirical analysis. Section 3 introduces the economic model. Section 4 introduces the airline data, providing some preliminary evidence of self-selection of airlines into markets. Section 5 shows the estimation results and Section 6 presents results and discussion of the merger exercise. Section 7 concludes.

2 A Simple Model with Two Firms

We illustrate the inference problem with a simple model of strategic interaction between two firms that is an extension of the classic selection model. Two firms simultaneously make an entry/exit decision and, if active, realize some level of a continuous variable. Each firm has complete information about the problem facing the other firm. We first consider a stylized version of this game written in terms of linear link functions. This model is meant to be illustrative, in that it is deliberately parametrized to be close to the classic single agent selection model. This allows for a more transparent comparison between the single vs multi agent model. Section 3 analyzes a full model of entry and pricing.

Consider the following system of inequality conditions,

[9] There is also an empirical literature on auctions (Li and Zheng (2009), Gentry and Li (2014), Roberts and Sweeting (2013), Li and Zhang (2015)) that has relaxed, in static models, the assumption that unobservable payoff shocks are not known at the time entry decisions are taken. However, in contrast to this literature, we allow for multiple, possibly *correlated*, unobservables.

$$
\begin{aligned}
y_1 &= 1\left[\delta_2 y_2 + \gamma Z_1 + \nu_1 \geq 0\right], \\
y_2 &= 1\left[\delta_1 y_1 + \gamma Z_2 + \nu_2 \geq 0\right], \\
S_1 &= X_1\beta + \alpha_1 V_1 + \xi_1, \\
S_2 &= X_2\beta + \alpha_2 V_2 + \xi_2
\end{aligned} \tag{1}
$$

where $y_j = 1$ if firm j decides to enter a market, and $y_j = 0$ otherwise for $j \in \{1,2\}$. So $\{1,2\}$ is the set of *potential* entrants. The endogenous variables are $(y_1, y_2, S_1, S_2, V_1, V_2)$. We observe (S_1, V_1) if and only if $y_1 = 1$ and (S_2, V_2) if and only if $y_2 = 1$. The variables $\mathbf{Z} \equiv (Z_1, Z_2)$ and $\mathbf{X} \equiv (X_1, X_2)$ are exogenous where $(\nu_1, \nu_2, \xi_1, \xi_2)$ are unobserved and are independent of $(\mathbf{Z}, \mathbf{X})$ while the variables (V_1, V_2) are endogenous (such as prices or product characteristics).[10]

The above model is an extension of the classic selection model to cover cases with two decision makers and allows for the possibility of endogenous variables on the rhs (the V's). The key distinction is the presence of simultaneity in the 'participation stage' where decisions are interconnected.

We first make a parametric assumption on the joint distribution of the errors. Let the unobservables have a joint normal distribution,

$$
(\nu_1, \nu_2, \xi_1, \xi_2) \sim N\left(0, \Sigma\right),
$$

where Σ is the variance-covariance matrix to be estimated. The off-diagonal entries of the variance-covariance matrix are not generally equal to zero. Such correlation between the unobservables is the source of selectivity bias.

One reason why we would expect firms to self-select into markets is because the fixed costs of entry are related to the demand and the variable costs. One would expect products of higher quality to be, at the same prices, in higher demand than products of lower quality and also to be more costly to produce. For example, some unforeseen reason (unobserved to the researchers) why a luxury car is more attractive to consumers may also be the reason the car requires more up-front investment and requires greater costs to produce a single

[10] It is simple to allow β and γ to be different among players, but we maintain this homogeneity for exposition.

unit. This would introduce correlation in the unobservables of the demand, marginal, and fixed costs. Alternatively, the data could be generated by a process similar to the classic selection problem in labor markets: there could exist (unobservably) high ability firms who have lower costs and a more attractive product, just like there might be high ability workers who command higher wages and are more likely to receive offers.

In the structural model of the airline industry we present in Section 3, the unobservables that determine outcomes also enter directly into the selection equation (see equation 7 in section 3). So, even if the unobservables are mutually independent, the model would still lead to selection effects. Firms with higher unobserved demand or lower unobserved costs will be more likely to enter. This departs from the standard Heckman selection setup and its generalization to two firms above because the structural errors terms that appear in the outcome equations (the ξ_1 and ξ_2 in (1) above) do not enter the first two equations in (1) (the entry equations).

Given that the above model defined in equation (1) is parametric, the only non-standard complications that arise are multiplicity of equilibria in the underlying game and endogeneity of the V's. Generally, and given the simultaneous game structure, the system (1) has multiple Nash equilibria in the identity of firms entering into the market. This multiplicity leads to a lack of a well-defined "reduced form" which complicates the inference question. Also, we want to allow for the possibility that the V's are also choice variables (or variables determined in equilibrium such as prices).

The data we observe are $(S_1 y_1, V_1 y_1, y_1, S_2 y_2, V_2 y_2, y_2, \mathbf{X}, \mathbf{Z})$ whereby, for example, S_1 is observed only when $y_1 = 1$. Given the normality assumption, we link the distribution of the unobservables conditional on the exogenous variables to the distribution of the outcomes to obtain the identified features of the model. The data allow us to estimate the distribution of $(S_1 y_1, V_1 y_1, y_1, S_2 y_2, V_2 y_2, y_2, \mathbf{X}, \mathbf{Z})$; the key to inference is to link this distribution to the one predicted by the model. To illustrate this, consider the observable $(y_1 = 1, y_2 = 0, V_1, S_1, \mathbf{X}, \mathbf{Z})$. For a given value of the parameters, the data allow us to identify

$$P(S_1 - \alpha_1 V_1 - X_1\beta \leq t_1;\, y_1 = 1, y_2 = 0 | X, Z) \tag{2}$$

for all[11] t_1. The particular form of the above probability is related to the residuals evaluated at t_1 and where we condition on all *exogenous variables* in the model. We elaborate further on this below.[12]

Remark 1 *It is possible to "ignore" the entry stage and consider only the linear regression parts in (1) above. Then, one could develop methods for dealing with distribution of* $(\xi_1, \xi_2 | Z, X, V)$*. For example, under mean independence assumptions, one would have*

$$E[S_1|Z, X, V] = X_1\beta + \alpha_1 V_1 + E[\xi_1|Z, X, V;\, y_1 = 1]$$

Here, it is possible to leave $E[\xi_1|Z, X, V;\, y_1 = 1]$ *as an unknown function of* (Z, X, V) *and then use a control function approach for example. In such a model, separating* (β, α_1) *from this unknown function (identification of* (β, α_1)*) requires extra assumptions that are hard to motivate economically (i.e., these assumptions necessarily make implicit restrictions on the entry model).*

To evaluate the probability in (2) above in terms of the model parameters, we first let $(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^U_{(1,0)})$ be the set of ξ_1 that are less than t_1 when the unobservables (ν_1, ν_2) belong to the set $A^U_{(1,0)}$. The set $A^U_{(1,0)}$ is the set where $(1, 0)$ is the unique (pure strategy) Nash equilibrium outcome of the model.

Next, let $\left(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^M_{(1,0)}, d_{(1,0)} = 1\right)$ be the set of ξ_1 that are less than t_1 when the unobservables (ν_1, ν_2) belong to the set $A^M_{(1,0)}$. The set $A^M_{(1,0)}$ is the set where $(1, 0)$ is one among the multiple equilibria outcomes of the model. Let $d_{(1,0)} = 1$ indicate that $(1, 0)$ was selected. The idea here is to try and "match" the distribution of residuals at a given parameter value predicted in the data, with its counterpart predicted by the model using

[11] Here we use the CDF, but we could also use probabilities of the form $P(t_0 \leq S_1 - \alpha_1 V_1 - X_1\beta \leq t_1;\, y_1 = 1, y_2 = 0|X, Z)$ for all $t_0 \leq t_1$. Bounding histogram like probabilities in some cases may be easier to compute.

[12] In the case where we have no endogeneity for example (α's equal to zero), then, one can use on the data side, $P(S_1 \leq t_1; y_1 = 1, y_2 = 0|\mathbf{X}, \mathbf{Z})$ which is equal to the model predicted probability $P(\xi_1 \leq -X_1\beta; y_1 = 1, y_2 = 0|\mathbf{X}, \mathbf{Z})$.

method of moments. By the law of total probability we have (suppressing the conditioning on $(\mathbf{X}, \mathbf{Z})$):

$$\begin{aligned} P(\xi_1 \leq t_1;\, y_1 = 1; y_2 = 0) \;&=\; P\left(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^U_{(1,0)}\right) \\ &+\; P(d_{1,0} = 1 \mid \xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^M_{(1,0)})\; P\left(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^M_{(1,0)}\right) \end{aligned} \tag{3}$$

The probability $P(d_{1,0} = 1 \mid \xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^M_{(1,0)})$ above is unknown and represents the equilibrium selection function. A feasible approach to inference, then, is to use the natural (or trivial) upper and lower bounds on this unknown function to get:

$$\begin{aligned} P\left(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^U_{(1,0)}\right) \leq P(\xi_1 \leq t_1;\, y_1 = 1; y_2 = 0) = P(S_1 + \alpha_1 V_1 - X_1\beta \leq t_1;\, y_1 = 1; y_2 = 0) \leq \\ P\left(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^U_{(1,0)}\right) + P\left(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^M_{(1,0)}\right) \end{aligned}$$

The middle part

$$P(S_1 - \alpha_1 V_1 - X_1\beta \leq t_1;\, y_1 = 1; y_2 = 0)$$

can be consistently estimated from the data given a value for (α_1, β, t_1). The LHS and RHS contain the following two probabilities

$$P\left(\xi_1 \leq t_1;\, (\nu_1, \nu_2) \in A^U_{(1,0)}\right),\; P\left(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^M_{(1,0)}\right).$$

These can be computed analytically (or via simulations) from the model for a given value of the parameter vector (that includes the covariance matrix of the errors) using the assumption that $(\xi_1, \xi_2, \nu_1, \nu_2)$ has a known distribution up to a finite dimensional parameter (we assume normal) and the fact that the sets $A^M_{(1,0)}$ and $A^U_{(1,0)}$, which depend on regressors and parameters, can be obtained by solving the game given a solution concept (See CT for examples of such sets). For example, for a given value of the unobservables, observables and parameter values, we can solve for the equilibria of the game which determines these sets.

Remark 2 *Note that we bound the distribution of the* residuals *as opposed to just the distribution of S_1 to allow some of the regressors to be endogenous. The conditioning sets in the LHS (and RHS) depend on exogenous covariates only, and hence these probabilities can be easily computed or simulated (for a given value of the parameters).*

The upper and lower bounds on the probability of the event $(S_2 - \alpha_2 V_2 - X_2\beta \leq t_2,\, y_1 = 0, y_2 = 1)$ can similarly be calculated. In addition, in the two player entry game (i.e. δ's are negative) above with pure strategies, the events $(1,1)$ and $(0,0)$ are uniquely determined, and so

$$P(S_1 - \alpha_1 V_1 - X_1\beta \leq t_1;\; S_2 - \alpha_2 V_2 - X_2\beta \leq t_2; y_1 = 1; y_2 = 1)$$

is equal to (moment equality)

$$P(\xi_1 \leq t_1, \xi_2 \leq t_2, \nu_1 \geq -\delta_2 - \gamma Z_1, \nu_2 \geq -\delta_1 - \gamma Z_2)$$

which can be easily calculated (via simulation for example). We also have:

$$P(y_1 = 0, y_2 = 0) = P(\nu_1 \leq -\gamma Z_1, \nu_2 \leq -\gamma Z_2)$$

To summarize, and for the two-equation selection models, the statistical moment inequality conditions implied by the model at the true parameters are:

$$m^l_{(1,0)}(t_1, \mathbf{Z}; \Sigma) \leq E\left(1\left[S_1 - \alpha_1 V_1 - X_1\beta \leq t_1; y_1 = 1; y_2 = 0\right]\right) \leq m^u_{(1,0)}(t_1, \mathbf{Z}; \Sigma)$$

$$m^l_{(0,1)}(t_2, \mathbf{Z}; \Sigma) \leq E\left(1\left[S_2 - \alpha_2 V_2 - X_2\beta \leq t_2; y_1 = 0; y_2 = 1\right]\right) \leq m^u_{(0,1)}(t_2, \mathbf{Z}; \Sigma)$$

$$E\left(1\left[S_1 - \alpha_1 V_1 - X_1\beta \leq t_1; S_2 - \alpha_2 V_2 - X_2\beta \leq t_2; y_1 = 1; y_2 = 1\right]\right) = m_{(1,1)}(t_1, t_2, \mathbf{Z}; \Sigma)$$

$$E\left(1\left[y_1 = 0; y_2 = 0\right]\right) = m_{(0,0)}(\mathbf{Z}; \Sigma)$$

where

$$\begin{aligned}
m^l_{(1,0)}(t_1, \mathbf{Z}; \Sigma) &= P\left(\xi_1 \leq t_1;\; (\nu_1, \nu_2) \in A^U_{(1,0)}\right) \\
m^u_{(1,0)}(t_1, \mathbf{Z}; \Sigma) &= m^l_{(1,0)}(t_1, \mathbf{Z}; \Sigma) + P\left(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^M_{(1,0)}\right) \\
m^l_{(0,1)}(t_2, \mathbf{Z}; \Sigma) &= P\left(\xi_2 \leq t_2;\; (\nu_2, \nu_2) \in A^U_{(0,1)}\right) \\
m^u_{(0,1)}(t_2, \mathbf{Z}; \Sigma) &= m^l_{(0,1)}(t_2, \mathbf{Z}; \Sigma) + P\left(\xi_2 \leq t_2; (\nu_1, \nu_2) \in A^M_{(0,1)}\right) \\
m_{(1,1)}(t_1, t_2, \mathbf{Z}; \Sigma) &= P(\xi_1 \leq t_1, \xi_2 \leq t_2, \nu_1 \geq -\delta_2 - \gamma Z_1, \nu_2 \geq -\delta_1 - \gamma Z_2) \\
m_{(0,0)}(\mathbf{Z}; \Sigma) &= P(\nu_1 \leq -\gamma Z_1, \nu_2 \leq -\gamma Z_2)
\end{aligned}$$

Hence, the above can be written as

$$E[\mathbf{G}(\theta, S_1y_1, S_2y_2, V_1y_1, V_2y_2, y_1, y_2; t_1, t_2)|\mathbf{Z}, X] \leq 0 \tag{4}$$

where $\mathbf{G}(.) \in \mathcal{R}^k$.

The last moment, $m_{(0,0)}(\mathbf{Z}; \Sigma)$, is the CT moment when no entrants are in the market. It is an important moment condition for the estimation of the fixed cost parameters. Observe that when $t_1, t_2 \to \infty$, the CMT moments collapse to the CT moments. Therefore, we also add the other CT moments to set of moment conditions that are used in estimation.

We use standard moment inequality methods to conduct inference on the identified parameters. In particular:[13]

Result 3 *Suppose the above parametric assumptions in model (1) are maintained. In addition, assume that* $(\mathbf{X}, \mathbf{Z}) \perp (\xi_1, \xi_2, \nu_2, \nu_2)$ *where the latter is normally distributed with mean zero and covariance matrix* Σ*. Then given a large iid data set on* $(y_1, y_2, S_1y_1, V_1y_1, S_2y_2, V_2y_2, \mathbf{X}, \mathbf{Z})$ *the true parameter vector* $\theta = (\delta_1, \delta_2, \alpha_1, \alpha_2, \beta, \gamma, \Sigma)$ *minimizes the nonnegative objective function below to zero:*

$$Q(\theta) = 0 = \int W(\mathbf{X}, \mathbf{Z}) \|\mathbf{G}(\theta, S_1y_1, S_2y_2, V_1y_1, V_2y_2, y_1, y_2)|\mathbf{Z}, X]\|_+ dF_{\mathbf{X},\mathbf{Z}} \tag{5}$$

for a strictly positive weight function $W(\mathbf{X}, \mathbf{Z})$.

It is simple to see that the above objective function is zero at the true parameter vector. In addition, if the model is partially identified, this objective function is also zero on all the parameters that belong to the identified set. The above is a standard conditional moment inequality model where we employ discrete valued variables in the conditioning set along with a finite (and small) set of t's.[14]

Clearly, the stylized model above provides intuition about the conceptual issues involved, but in the next section, we link this system to a model of behavior where the decision to enter (or to provide a product) is more explicitly linked to an economic condition of profits.

[13] See the Appendix A for more details. See CT for an analogous result and the proof therein.

[14] We discuss the selection of the t's in Appendix B.

This entails specification of costs, demand, and an equilibrium solution concept. This is the subject of the next Section, the main contribution of the paper.

3 A Model of Entry and Price Competition

3.1 The Structural Model

Above, we described our methodology using a linear outcome and selection equation for clarity and consistency with the literature on selection. In this section, we present a structural model of demand, pricing, and entry that we take to data from the airline industry. We consider the case of two potential entrants who decide, simultaneously, whether to serve a market and the price to charge in the market.

The profits of firm 1 if this firm decides to enter is

$$\pi_1 = (p_1 - c(W_1, \eta_1)) \mathcal{M} \cdot \tilde{s}_1 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) - F(Z_1, \nu_1),$$

where

$$\tilde{s}_1 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) = \overbrace{s_1 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi)}^{\text{duopoly demand}} y_2 + \overbrace{s_1 (p_1, X_1, \xi_1)}^{\text{monopoly demand}} (1 - y_2)$$

is the share of firm 1 which depends on whether firm 2 is in the market, $\mathcal{M}$ is the market size, $c(W_1, \eta_1)$ is the constant marginal cost for firm 1, $F(Z_1, \nu_1)$ is the fixed cost of firm 1, and prices $\mathbf{p} = (p_1, p_2)$. A profit function for firm 2 is specified in the same way.

In addition, we have equilibrium first order conditions that determine prices and shares,

$$\begin{cases} (p_1 - c(W_1, \eta_1)) \, \partial \tilde{s}_1 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) / \partial p_1 + \tilde{s}_1 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) = 0 \\ (p_2 - c(W_2, \eta_2)) \, \partial \tilde{s}_2 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) / \partial p_2 + \tilde{s}_2 (\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) = 0 \end{cases}, \tag{6}$$

which are the first order equilibrium conditions in a simultaneous Nash Bertrand pricing game.

In this model, $y_j = 1$ if firm j decides to enter a market, and $y_j = 0$ otherwise, where $j = 1, 2$ indexes the firms. We impose the following entry condition:

$$y_j = 1 \quad \text{if and only if} \quad \pi_j \geq 0 \quad j = 1, 2$$

There are six endogenous variables: p_1, p_2, S_1, S_2, y_1, and y_2. The observed exogenous variables are $\mathcal{M}$, $\mathbf{W} = (W_1, W_2)$, $\mathbf{Z} = (Z_1, Z_2)$, $\mathbf{X} = (X_1, X_2)$. So, putting these together, we get the following system:

$$\begin{cases} y_1 = 1 \Leftrightarrow \pi_1 = (p_1 - c(W_1, \eta_1))\, \mathcal{M} \cdot \tilde{s}_1(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) - F(Z_1, \nu_1) \geq 0, & \text{Entry Conditions} \\ y_2 = 1 \Leftrightarrow \pi_2 = (p_2 - c(W_2, \eta_2))\, \mathcal{M} \cdot \tilde{s}_2(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) - F(Z_2, \nu_2) \geq 0, & \\ & \\ S_1 = \tilde{s}_1(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi), & \text{Demand} \\ S_2 = \tilde{s}_2(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi), & \\ & \\ (p_1 - c(W_1, \eta_1))\, \partial \tilde{s}_1(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) / \partial p_1 + \tilde{s}_1(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) = 0, & \text{Equilibrium Pricing} \\ (p_2 - c(W_2, \eta_2))\, \partial \tilde{s}_2(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) / \partial p_2 + \tilde{s}_2(\mathbf{p}, \mathbf{X}, \mathbf{y}, \xi) = 0, & \end{cases} \tag{7}$$

The first two inequalities are entry conditions that require that in equilibrium a firm that serves a market must be making non-negative profits. The third and fourth equations are demand equations. The fifth and sixth equations are pricing first order conditions. An equilibrium of the model occurs when firms make entry and pricing decisions such that all the six conditions are satisfied. The firm level unobservables that enter into the fixed costs are denoted by ν_j, $j = 1, 2$. The unobservables that enter into the variable costs are denoted by η_j, $j = 1, 2$ while the unobservables that enter into the demand equations are denoted by ξ_j, $j = 1, 2$. The model represented by the set of equations above might have multiple equilibria in market structure. There are no multiple equilibria in the pricing game: Nocke and Schutz (2018) show that there is a unique pricing equilibrium in the case of single-product nested logit, which is what we consider in our application (See Appendix 3 of their On-line Appendix).

Even though the conceptual approach is the same, the inference procedure for this system is computationally more demanding for this model than the one we studied in Section 2. It is more complex because one needs to *solve for the equilibrium of the full model*, which has six (rather than just four) endogenous variables. On the other hand, one only had to solve for the equilibrium of the entry game in the model (1). The methodology presented in

Section (2) can be used to estimate model (7), but now there are *two* unobservables for each firm over which to integrate (the marginal cost and the demand unobservables).

To understand how the model relates to previous work, observe that if we were to estimate a reduced form version of the first two inequalities of the system (7), then that would be akin to the entry game literature (Bresnahan and Reiss, 1990, 1991; Berry, 1992; Mazzeo, 2002; Seim, 2006; Ciliberto and Tamer, 2009). If we were to estimate the third to sixth equation in the system (7), then that would be akin to the demand-supply literature (Bresnahan, 1987; Berry, 1994; Berry, Levinsohn, and Pakes, 1995), depending on the specification of the demand system. So, here we join a demand and entry model, while allowing the unobservables of the six conditions to be correlated with each other. This is important, as a model that combines both pricing and entry decisions is able to capture a richer picture of firms' response to policy. For example, the model allows for market structure to adjust optimally after a merger, which may in turn affect prices.

3.2 Parameterizing the model

To parametrize the various functions above, we follow Bresnahan (1987) and Berry, Levinsohn, and Pakes (1995), where the unit marginal cost can be written as:

$$\ln c\left(W_j, \eta_j\right) = \varphi_j W_j + \eta_j. \tag{8}$$

As in the entry game literature mentioned above, the fixed costs are

$$\ln F\left(Z_j, \nu_j\right) = \gamma_j Z_j + \nu_j. \tag{9}$$

We assume demand is derived from the canonical differentiated product discrete choice model (Bresnahan, 1987; Berry, 1994; Berry, Levinsohn, and Pakes, 1995). We include a product nest which allows for all of the inside products to share unobserved heterogeneity. Specifically, indirect utility for consumer i from choosing carrier j is

$$\begin{aligned} u_{ij} &= X_j'\beta + \alpha p_j + \xi_j + \upsilon_{ig} + (1-\lambda)\,\epsilon_{ij}, \qquad (10) \\ u_{i0} &= \epsilon_{i0}, \end{aligned}$$

where X_j is a vector of product characteristics, p_j is the price, (β, α) are the taste parameters, and ξ_j are product characteristics unobserved to the econometrician.

Following Berry (1994), carrier j's market share is

$$s_j(\mathbf{X}, \mathbf{p}, \xi, \beta_r, \alpha, \lambda) = \frac{e^{(X_j'\beta + \alpha p_j + \xi_j)/(1-\lambda)}}{D} \frac{D^{(1-\lambda)}}{1 + D^{(1-\lambda)}}, \tag{11}$$

where the D represents the sum of exponentiated utilities for all products

$$D = \sum_{j=1}^{J} e^{(X_j'\beta + \alpha p_j + \xi_j)/(1-\lambda)}.$$

Unlike in typical demand estimation, we need to compute shares for any given potential market structure. To do this, we introduce some notation. Let

$$E \equiv \{(y_1, .., y_j, .., y_K) : y_j = 1 \text{ or } y_j = 0, \forall 1 \leq j \leq K\}$$

denote the set of possible market structures, which contains 2^K elements. Let $e \in E$ be an element or a market structure. For example, in the model above where $K = 2$, the set of possible market structures is $E = \{(0,0), (0,1), (1,0), (1,1)\}$. Let $\mathbf{X}^e$, $\mathbf{p}^e$, and ξ^e, N^e denote the matrices of, respectively, the exogenous variables, prices, unobservable firm characteristics, and number of firms when the market structure is e.

We can express demand for any given market structure in the following way,

$$\ln s_j(\beta, \alpha, \mathbf{X}^e, \mathbf{p}^e, \xi^e) - \ln s_0(\beta, \alpha, \mathbf{X}^e, \mathbf{p}^e, \xi^e) = X_j\beta + \alpha p_j + \lambda \ln s_{j/g} + \xi_j, \tag{12}$$

where $s_{j/g}$ is share of carrier j among all other carriers in the market, excluding the outside option.

Lastly, unlike typical demand estimation but similar to the entry literature, we parameterize the joint distribution of unobservables. Following Berry (1992) and CT, we specify the unobservables that enter into the fixed cost inequality condition, η_{jm}, as including firm-specific unobserved heterogeneity, $\tilde{\eta}_{jm}$, as well as market specific unobserved heterogeneity, η_m. η_m are unobservables that are market specific and capture, for example, the fact that in market m there are cost shocks that are common across the potential entrants. Thus, we have $\eta_{jm} = \tilde{\eta}_{jm} + \eta_m$. Following Bresnahan [1987] and BLP [1995], the marginal cost and demand unobservables only includes firm-specific heterogeneity.

The unobservables have a joint normal distribution:

$$(\nu_1, \nu_2, \xi_1, \xi_2, \tilde{\eta}_{1m}, \tilde{\eta}_{2m}) \sim N(0, \Sigma) \tag{13}$$

where Σ is the variance-covariance matrix to be estimated. Notice that here we do not include η_m because we assume it is independent of other errors.[15]

The off-diagonal terms pick up the correlation between the unobservables that is part of the source of the selection bias in the model. In the empirical implementation of our model, we use the following variance-covariance matrix

$$\Sigma_m = \begin{bmatrix} \sigma_\xi^2 \cdot I_{Km} & \sigma_{\xi\eta} \cdot I_{Km} & \sigma_{\xi\nu} \cdot I_{Km} \\ \sigma_{\xi\eta} \cdot I_{Km} & \sigma_\eta^2 \cdot I_{Km} & \sigma_{\eta\nu} \cdot I_{Km} \\ \sigma_{\xi\nu} \cdot I_{Km} & \sigma_{\eta\nu} \cdot I_{Km} & \sigma_\nu^2 \cdot I_{Km} \end{bmatrix},$$

where I_{Km} is a $K_m \times K_m$ identity matrix. For computational simplicity, this specification restricts the correlations to be the same for each firm. It maintains that the correlation is non-zero only among the unobservables of a firm (within-firm correlation), and not between the unobservables of the K_m firms (between-firm correlation).

3.3 Simulation Algorithm

To estimate the parameters of the model we need to predict the market structures and derive distributions of demand and supply unobservables to construct the distance function. This requires the evaluation of a large multidimensional integral, therefore we have constructed an estimation routine that relies heavily on simulation. We solve directly for all equilibria at each iteration in the estimation routine.

The simulation algorithm is presented for the case when there are K potential entrants. We rewrite the model of price and entry competition using the parameterizations above.

[15] When we perform simulation, we draw $\tilde{\eta}_{jm}$ and η_m independently from two standard normal distributions. Then, we will apply the Cholesky decomposition to allow for correlations between the demand, marginal cost, and the firm specific fixed cost unobservables. Then, we add the market-specific fixed cost unobservable to the firm-specific fixed cost unobservable. See Online Appendix B for details.

$$\begin{cases} y_j = 1 \Leftrightarrow \pi_j \equiv (p_j - \exp(\varphi W_j + \eta_j))\, M s_j(\mathbf{X}^e, \mathbf{p}^e, \xi^e) - \exp(\gamma Z_j + \nu_j) \geq 0, \\ \ln s_j(\beta, \alpha, \mathbf{X}^e, \mathbf{p}^e, \xi^e) - \ln s_0(\beta, \alpha, \mathbf{X}^e, \mathbf{p}^e, \xi^e) = X_j'\beta + \alpha p_j + \lambda s_{j|g} + \xi_j \\ \ln[p_j - b_j(\mathbf{X}^e, \mathbf{p}^e, \xi^e)] = \varphi W_j + \eta_j, \end{cases}, \quad (14)$$

for $j = 1, ..., K$ and $e \in E$.

We present the simulation algorithm here and provide many more details, including computational guidance, in Appendix B.

First, we take ns pseudo-random independent draws from a $3 \times |K|$-variate joint standard normal distribution, where $|K|$ is the cardinality of K. Let $r = 1, ..., ns$ index pseudo-random draws. These draws remain unchanged during the minimization. Next, the algorithm uses three steps that we describe below.

Set the candidate parameter value to be $\Theta^0 = (\alpha^0, \beta^0, \varphi^0, \gamma^0, \Sigma^0)$.

1. We estimate the probability distributions of the residuals. The steps here do not involve any simulations.

   (a) Use α^0, β^0, φ^0 to compute the demand and first order condition residuals $\hat{\xi}_j^{\hat{e}}$ and $\hat{\eta}_j^{\hat{e}}$. These can be done easily using (14) above.

   (b) Construct $\Pr(\hat{\xi}^{\hat{e}} <= \mathbf{t_D}, \hat{\eta}^{\hat{e}} <= \mathbf{t_S} \mid \mathbf{X}, \mathbf{W}, \mathbf{Z})$, which are joint probability distributions of $\hat{\xi}^{\hat{e}}, \hat{\eta}^{\hat{e}}$ conditional on the values taken by the control variables. $\mathbf{t_D}$ are the t's for the demand residuals, while $\mathbf{t_S}$ are the t's for the supply residuals.

2. Next, we construct the probability distributions for the lower and upper bound of the "simulated errors"selected by the model for a guess of the parameters, Θ^0.

   (a) Simulate random vectors of unobservables (ν_r, ξ_r, η_r) from a multivariate normal density with a given covariance matrix, Σ^0, using the pseudo-random draws described above.

(b) For each potential market structure e of the $2^{|K|}-1$ possible ones (excluding the one where no firm enters), we solve the subsystem of the N^e demand equations and N^e first order conditions in (14) for the *equilibrium* prices $\bar{\mathbf{p}}_r^e$ and shares $\bar{\mathbf{s}}_r^e$.[16]

(c) Compute $2^{|K|}-1$ *total* profits.

(d) We use the total profits to determine which of the $2^{|K|}$ market structures are *predicted* as equilibria of the full model. If there is a unique equilibrium, say e^*, then we collect the simulated errors of the firms that are present in that equilibrium, $\xi_r^{e^*}$ and $\eta_r^{e^*}$. In addition, we collect $\nu_r^{e^*}$ and include them in $A_{e^*}^U$, which was defined in Section (2). If there are multiple equilibria, say e^* and e^{**}, then we collect the "simulated errors" of the firms that are present in those equilibria, respectively $\left(\xi_r^{e^*}, \eta_r^{e^*}\right)$ and $\left(\xi_r^{e^{**}}, \eta_r^{e^{**}}\right)$.[17] In addition, we collect $\nu_r^{e^*}$ and $\nu_r^{e^{**}}$ and include them, respectively, in $A_{e^*}^M$ and $A_{e^{**}}^M$, which were also defined in Section (2).[18]

(e) Construct
$\Pr\left(\xi_r^e <= \mathbf{t_D}, \eta_r^e <= \mathbf{t_S}; \nu \in A_e^M | \mathbf{X}, \mathbf{W}, \mathbf{Z}\right)$ and $\Pr\left(\xi_r^e, \eta_r^e; \nu \in A_e^U | \mathbf{X}, \mathbf{W}, \mathbf{Z}\right)$.[19]

3. We construct the distance function (5) in Section (2). The approach we use for inference follows the implementation of Chernozhukov, Hong, and Tamer (2007) in CT, where we use subsampling based methods to construct confidence regions.

Conceptually, the above is a minimum distance procedure that compares the distribution function from the data (constructed in Step 1 above) to the upper and lower bounds on this distribution predicted by the model (the upper and lower bounds are constructed in

[16]For example, if we look at a monopoly of firm j ($|e|=1$) then the demand $Q_j\left(p_{jr}, X_{jr}, \xi_{jr}; \beta\right)$ is readily computed, and the monopoly price, p_{jr}, as well. Given the parametric assumptions, there is a unique pure-strategy price equilibrium, conditional on the market structure. See Nocke and Schutz (2018) for uniqueness in the single product nested logit case considered in our empirical exercise.

[17]The set of firms in the two equilibria (if there are multiple equilibria) may not be the same.

[18]See Appendix B (page 4) for details, including how we handle situations where no pure-stategy equilibria exist.

[19]These CDFs in this setting with two unobservables for each firm are analogous to the ones with just one unobservable per firm on described in Section 2. We use the same t's that we used to construct the CDFs of the residuals.

Step 2. The upper and lower bounds in Step 2 are a result of multiple equilibria while the complication in Step 1 is due to endogeneity.

4 Data and Industry Description

We apply our methods to data from the airline industry. This industry is particularly interesting in our setting for two main reasons. First, there is considerable variation in prices and market structure across markets and across carriers, which we expect to be associated with self-selection of carriers into markets. Second, this is an industry where the study of market structure and market power are particularly meaningful because there have been several recent changes in the number and identity of the competitors, with recent mergers among the largest carriers (Delta with Northwest, United with Continental, and American with USAir). Our methods allow us to examine, within the context of our model, the implications of mergers on equilibrium prices and also on market structure. We start with an examination of our data, and then we provide our estimates.

4.1 Market and Carrier Definition

Data. We use data from several sources to construct a cross-sectional dataset, where the basic unit of observation is an airline in a market (a *market-carrier*). The main datasets are the second quarter of 2012's *Airline Origin and Destination Survey (DB1B)* and of the *T-100 Domestic Segment Dataset*, the *Aviation Support Tables*, available from the DOT's National Transportation Library. We also use the US Census for the demographic data.[20]

We define a market as a unidirectional trip between two airports, irrespective of intermediate transfer points.[21] The dataset includes the markets between the top 100 US Metropolitan Statistical Areas ranked by their population. We include markets that are not served by any carrier. There are $8,163$ unidirectional markets, and each one is denoted by $m = 1, ..., M$.

[20] See Section C of the Appendix for a detailed discussion on the data cleaning and construction.

[21] We do not model the decision of nonstop versus connecting flights. This is very difficult problem given the hub-network structure of airline markets. See Aguirregabiria and Ho (2012) for a treatment of hub-spoke networks using a dynamic game framework and Li et al. (2017) for a recent treatemnt in a static framework.

There are six carriers in the dataset: American, Delta, United, USAir, Southwest, and a low cost type, denoted by `LCC`. The *Low Cost Carrier* type includes Alaska, JetBlue, Frontier, Allegiant, Spirit, Sun Country, Virgin. These firms rarely compete in the same market. The subscript for carriers is j, $j \in \{AA, DL, UA, UA, LCC\}$. There are 23,155 market-carrier observations for which we observe prices and shares. There are 710 markets that are not served by any firm.

We denote the number of potential entrants in market m as K_m where $|K_m| \leq 6$. An airline is considered a potential entrant if it is serving at least one market out of both of the endpoint airports.[22]

Tables 1 and 2 present the summary statistics for the distribution of potential and actual entrants in the airline markets. Table 1 shows that American Airlines enters in 39 percent of the markets, although it is a potential entrant in 71 percent of markets. Southwest, on the other hand, is a potential entrant in 64 percent of markets, and enters in 46 percent of the time. So this already shows some interesting heterogeneity in the entry patterns across airlines. Table 2 shows the distribution in the number of potential entrants, and we observe that the large majority of markets have between four and six potential entrants, with less than 2 percent having just one potential entrant.

Table 1: *Entry Moments*

	Actual Entry	Potential Entry
AA	0.39	0.71
DL	0.73	0.95
LCC	0.18	0.46
UA	0.51	0.80
US	0.49	0.87
WN	0.46	0.64

Empirical entry probabilities and the percent of markets as a potential entrant, across airlines.

[22] See Goolsbee and Syverson (2008) for an analogous definition. Variation in the identity and number of potential entrants has been shown to help the identification of the parameters of the model (Ciliberto et al., 2010).

Table 2: *Distribution of Potential Entrants Across Markets*

	Number of Potential Entrants					
	1	2	3	4	5	6
Percent of Markets	1.74	10.61	14.58	16.57	28.13	28.37

Distribution of the fraction of markets by number of potential entrants.

For each firm in a market there are three endogenous variables: whether or not the firm is in the market, the price that the firm charges in that market, and the number of passengers transported. Following the notation used in the theoretical model, we indicate whether a firm is active in a market as $y_{jm} = 1$, and if it is not active as $y_{jm} = 0$. For example, we set $y_{LCC} = 1$ if at least one of the low cost carriers is active.

Table 3 presents the summary statistics for the variables used in our empirical analysis. For each variable we indicate in the last column whether the variable is used in the entry inequality conditions, demand and marginal cost equations. As in Berry, Carnall, and Spiller (2006), Berry and Jia (2010), and Ciliberto and Williams (2014), market size is the geometric mean of the MSA population of the end-point cities.

The top panel of Table 3 reports the summary statistics for the ticket prices and passengers transported in a quarter. For each airline that is actively serving the market we observe the quarterly mean ticket fare, p_{jm}, and the total number of passengers transported in the quarter, Q_{jm}. The average value of the mean ticket fare is 242.88 dollars and the average number of passengers transported is $2,602.79$.

Demand. Demand is here assumed to be a function of the number of *Origin Presence*, which is defined as the *number* of markets served by an airline out of the origin airport. We maintain that this variable is a proxy of frequent flyer programs: the larger the number of markets that an airline serves out of an airport, the easier is for a traveler to accumulate points, and the more attractive flying on that airline is, *ceteris paribus*. The *Distance* between the origin and destination airports is also a determinant of demand, as shown in previous studies (Berry, 1990; Berry and Jia, 2010; Ciliberto and Williams, 2014).

Table 3: *Summary Statistics*

	Mean	Std. Dev.	Min	Max	N	Equation
Endogenous Variables						
Price ($)	242.88	55.25	77.13	364.00	22,445	Entry, Utility, MC
Passengers	2602.79	7042.02	90	112,120	22,445	Entry, Utility, MC
All Markets						
Origin Presence	100.36	71.88	0	267	48,978	Utility, MC
Nonstop Origin	7.04	13.57	0	127	48,978	Entry
Nonstop Destin.	7.11	13.61	0	127	48,978	Entry
Distance (000)	1.11	0.58	0.15	2.72	48,978	Utility, MC
Markets Served						
Origin Presence	143.23	57.91	1	267	22,445	Utility, MC
Nonstop Origin	10.60	16.76	0	127	22,445	Entry
Nonstop Destin.	10.67	16.77	0	127	22,445	Entry
Distance (000)	1.17	0.56	0.20	2.72	22,445	Utility, MC

Summary statistics from sample dscribed in the text. Observations from 48,978 potential airline-markets from 8,163 distinct markets. 22,445 airline-markets are active.

The middle and bottom panels of Table 3 report the summary statistics for the exogenous explanatory variables. The middle panel computes the statistics on the whole sample, while the bottom panel computes the statistics only in the markets that are served by at least one airline.

There is clearly selection on observables in our setting. The mean value of *Origin Presence* is 100.36 across all markets, and it is up to 143.23 in markets that are actually served. The mean value of *Distance* is 1110 miles (one-way), which is slightly lower than the mean values for active airline-markets, 1170 miles.

Fixed and Marginal Costs in the Airline Industry.[23] The total costs of serving an

[23]We thank John Panzar for helpful discussions on how to model costs in the airline industry. See also Panzar (1979).

airline market consists of three components: airport, flight, and passenger costs.[24]

Airlines must lease gates and hire personnel to enplane and deplane aircrafts at the two endpoints. These *airport* costs do not change with an additional passenger flown on an aircraft, and thus we interpret them as fixed costs. We parameterize fixed costs as functions of *Nonstop Origin*, the number of non-stop routes that an airline serves out of the origin airport, and *Nonstop Destination*, the number of non-stop routes that an airline serves out of the destination airport, to capture economies of density (Brueckner and Spiller (1994)).

Next, a particular *flight's* costs also enter the marginal cost. This is because these costs depend on the number of flights serving a market, on the size of the planes used, on the fuel costs, and on the wages paid to the pilots and flight attendants. In our static model, the flight costs are variable in the number of passengers transported in a quarter. The *accounting* unit costs of transporting a passenger are those associated with issuing tickets, in-flight food and beverages, and insurance and other liability expenses. These costs are very small when compared to the airport and flight specific costs. We maintain that the flight and passenger costs enter the *economic* opportunity cost of flying a passenger.[25]

Returning to the middle and bottom panels of Table 3 we observe that there is selection on these observables as well. The mean value of *Nonstop Origin* is 7.04 in all markets, and 10.60 in markets that were actively served. The magnitudes are analogous for *Nonstop Destination*.

The economic marginal cost is not observable (Rosse, 1970; Bresnahan, 1989; Schmalensee, 1989). We parameterize it as a function of *Origin Presence*. The idea is that the opportunity cost is a function of i) the whole network of that carrier that can be reached out of that airport, and ii) of the degree of competition that the carrier faces out of that airport, which is here captured by the size of the network that other airlines have at the origin airport. Given our interpretation of flight costs as entering the variable costs, we also allow the marginal

[24]Other costs are incurred at the aggregate, national, level, and we do not estimate them here (advertising expenditures, for example, are rarely market specific).

[25]This can be interpreted as the highest profit that the airline could make off of an alternative trip that uses the same seat on the same plane, possibly as part of a flight connecting two different airports (Elzinga and Mills, 2009).

cost to be a function of the non-stop distance, *Distance*, between two airports.

4.2 Identification

We begin by discussing the source of exogenous variation in our estimation and how the parameters of the model are identified. Several variables are omitted in the demand estimation, and their omission could bias the estimation of the price coefficient. For example, we do not include frequency of flights or whether an airline provides connecting or nonstop service between two airports. As mentioned before, quality of airline service is also omitted. All these variables enter in ξ. We instrument for price using the exogenous variables for *all potential* rivals. These instruments are different than the "BLP instruments" widely used in the literature (Berry, Levinsohn, and Pakes, 1995). The aggregation typically used in the form of the BLP instruments has been shown to be problematic (see Gandhi and Houde, 2016). [26] Our approach is slightly different from the standard one and capture greater variation in competitive environments because: i) we include every potential entrants' characteristics separately instead of summing or averaging the characteristics in a market; ii) we consider the characteristics of all potential entrants, and not just those of the actual entrants. In addition, the exogenous variables that affect fixed costs, which correlate with equilibrium prices through the entry conditions in our model, also enter as instruments for the demand estimation.

The fixed cost parameters in the entry inequalities are identified if there is a variable that shifts the fixed cost of one firm without changing the fixed costs of the competitors. This condition is also required to identify the parameters in Ciliberto and Tamer (2009), but in our case this variable should also be excluded from demand and marginal cost. First, we use the carrier's *Nonstop Destination*, the number of nonstop flights from the destination airport. Our choice of this variable as our exclusion restriction is motivated by the observation that passengers only care about the network out of the origin airport when they select and airline, for example because of their ability to accumulate frequent flyer miles over time.[27] In our

[26] For example, this approach is also used by Berry and Jia (2010).

[27] Berry and Jia (2010) also assume that the variable *Nonstop Destination* is excluded from the demand.

robustness analysis we have determined that we can also include the carrier's *Nonstop Origin*. Notice that the origin-specific variable, *Nonstop Origin* is the same across markets from the same origin airport. In contrast, the destination variable, *Nonstop Destination*, is not, and this allows for the fixed costs to change across markets from the same airport.

A crucial source of exogenous variation across markets, which reinforces the identification power of the instruments discussed above, is given by the variation in the identity and number of potential entrants across markets, as in Berry (1992). First, the parameters of the exogenous variables in the entry inequalities are *point identified* when there is only one potential entrant because the model would collapse to a classic discrete choice model. Second, the exogenous variables shifting the demand function vary across markets from the same airport. If the exogenous variables in the demand function were the same across all markets from the same airport, then the differences in prices and shares that we observe in those markets would have to be fully explained by the random variables. Instead, the variation is also explained by the variation in the identity of the potential entrants and, consequently, by variation in the attributes of rival products.

Next, we discuss the variation in the data that identifies the variance-covariance matrix. The variance of the unobservable entering the demand function is identified by the variance in (the logarithms of) the odds, which, in turn, are functions of the shares of passengers transported by the airlines. The variance of the unobservables entering in the marginal cost is identified by the variance in the markups charged by the firm, which in turn are functions of the observed prices. The variance in the unobservables entering the entry inequality is identified by the variance in the *variable profits*, which in turn are functions of the observed revenues. Notice that variable profits are expressed in monetary terms, and therefore the fixed cost parameters do not suffer from the standard caveat that they are identified up to

However, they assume that this variable enters the marginal cost equation. Earlier, we discussed our assumptions about marginal and fixed costs in our context (whereas Berry and Jia, 2010, do not model fixed costs). We think of marginal costs as the opportunity cost of serving other passengers from the origin airport, and so should include variables relating to the origin airports. As discussed in the text, we think of fixed costs as relating to economies of density (see Brueckner and Spiller, 1994). One way to capture the network density is to consider how many connections happen at the destination airport.

a scale.

Next, we describe how the correlations between the unobservables are identified.[28] The two most important correlations are those that govern the unobserved selection: the correlations of the unobserved fixed cost with the unobserved component of marginal cost and demand. For example, suppose there is a set of firms that share the same observable attributes (i.e., same market type) which implies we predict them to have the same exact revenue conditional on entering the market. If, among this set of firms, we observe in the data that firms that enter are more likely to have a lower price (again, holding revenues constant), then we would infer that there is a positive correlation between marginal costs (the reason for the low price) and fixed costs (the reason for entering, holding revenue fixed). If among this group of firms we observe firms that enter are more likely to have higher market shares, then we would infer that there is a negative correlation between unobserved demand (the reason why demand is high) and unobserved fixed costs (low fixed costs being the reason for entering conditional on revenues). More generally, we observe three things in the data: demand, prices, and entry. We use the averages, variances, and covariances between these variables to identify features of the utility function, cost functions (marginal and fixed), and covariances between utility and costs.

5 Results

We organize the discussion of the results in two steps. First, we present the results when we estimate demand and supply using the standard GMM method (i.e. Berry, 1994). Next, we estimate demand and supply using our method, but assume that entry is exogenous. Lastly, we present results using our methodology that accounts for firms' entry decisions. To facilitate the comparison across model specifications and methodologies, in all columns of Table 4 we report the confidence region that is defined as the set that contains the parameters

[28] Given our assumptions (or lack thereof) on equilibria selection in our model, we do not claim that the parameters of interest are point identified. However, it is useful to generally understand what covariation in the data informs us about the identified set.

that cannot be rejected as the truth with at least 95% probability.[29]

5.1 Results with Exogenous Market Structure

In Column 1 of Table 4, we display the results from GMM estimation of a model where the inverted demand is given by a nested logit regression, as in Equation 12.[30]

In order to limit the space over which to draw for the minimization procedure, we standardize all the exogenous variables.[31] All the results are as expected and resemble those in previous work, for example Berry and Jia (2010) and Ciliberto and Williams (2014).[32] Starting from the demand estimates, we find the price coefficient to be negative, and included in [-2.394,-2.192] and λ, the nesting parameter, to be between 0 and 1.[33] The corresponding median elasticity is included in [-8.170,-8.096], and the confidence interval for the median markup is [30.312,30.383]. A larger presence at the origin airport is associated with more demand as in (Berry, 1990), and longer route distance is associated with stronger demand as well. The marginal cost estimates show that it is increasing in distance, and decreasing in presence.

Next, we estimate the same exogenous entry model using our methodology. We do this because our methodology requires additional assumptions to those of GMM, such as maintaining the assumption that the unobservables are normally distributed. Estimating the exogenous version using our methodology allows us to (1) examine how close the estimates using these additional assumption are to the standard GMM approach and (2) compare the endogenous market structure version of the model more directly with the exogenous market structure version.

We present the results of this estimation in Column 2 of Table 4. We observe that all of

[29]This is the approach that was used in CT. See the On-line Supplement to CT and Chernozhukov, Hong, and Tamer (2007) for details. Notice that there are no multiple equilibria in Columns 1 and 2.

[30]We instrument for price and the nest shares using the value of the exogenous data for every firm, regardless of whether they are in the market, including fixed costs which are excluded from supply and demand. So, for example, there are six instruments for every element in X, W, and Z.

[31]See Section C in the Appendix for more details.

[32]We also have estimated the GMM model only with the demand moments, and the results were very similar. See Section D in the Appendix.

[33]We denote fares in $100s for readability of the estimates.

the cost estimates in Column 2 overlap those in Column 1. Most of the demand estimates in Column 2 overlap with those in Column 1, and the ones that do not overlap are very close. The estimate of the median elasticity of demand and of the markup are also overlapping the ones in Column 1. In future work we hope to relax some of the distributional assumptions made in our current work.

Table 4: *Parameter Estimates*

	GMM	Exogenous Entry	Endogenous Entry
Demand			
Price	[-2.394, -2.192]	[-2.450, -2.290]	[-1.992, -1.956]
λ	[0.318, 0.518]	[0.300, 0.566]	[0.116, 0.144]
Distance	[0.309, 0.365]	[0.155, 0.394]	[1.085, 1.273]
Origin Presence	[0.293, 0.340]	[-0.128, 0.191]	[-0.613, -0.426]
LCC	[-0.334, -0.143]	[-2.028, -0.991]	[0.353, 0.853]
WN	[0.217, 0.336]	[-0.115, 0.575]	[0.904, 1.409]
Constant	[-1.588, -1.105]	[-2.588, -2.180]	[-5.191, -4.878]
Marginal Cost			
Distance	[0.118, 0.124]	[0.076, 0.128]	[0.111, 0.158]
Origin Presence	[-0.030, -0.020]	[-0.050, 0.014]	[-0.649, -0.626]
Cons LCC	[-0.349, -0.320]	[-0.506, -0.281]	[-0.002, 0.108]
Cons WN	[-0.154, -0.137]	[-0.165, -0.038]	[0.203, 0.330]
Constant	[5.360, 5.370]	[5.341, 5.394]	[5.267, 5.290]
Fixed Cost			
Nonstop Origin	–	–	[-0.452 -0.264]
Nonstop Dest.	–	–	[-2.260 -1.885]
Constant	–	–	[-1.657 -1.288]
Variance-Covariance*			
Variance Demand	1.523	[1.336 2.853]	[5.142, 5.826]
Variance Marg. Cost	0.060	[0.018 0.055]	[0.334,0.373]
Variance Fixed Cost	–	–	[3.721 5.5332]
Demand-MC Covariance	0.185	[0.112 0.256]	[-0.099,0.243]
Demand-FC Covariance	–	–	[0.631,0.786]
MC-FC Covariance	–	–	[1.119,1.215]
Market Power			
Median Elasticity	[-8.170, -8.096]	[-9.429, -7.020]	[-4.031, -3.864]
Median Markup	[28.129, 28.259]	[22.219, 33.006]	[44.416, 46.083]

Results from estimation of the model presented in Section 3. Column 1: Standard GMM estimation. Column 2: Estimation using the methodology described in Section 2, but holding market structure exogenous. Column 3: Estimation using the methodology described in Section 2. Column 1 presents the standard 95% confidence intervals. Columns 2 and 3 contain 95% confidence bounds constructed using the method in Chernozhukov, Hong, and Tamer (2007). Price coefficient multiplied by 100.

5.2 Results with Endogenous Market Structure

Column 3 of Table 4 displays the estimates from our model using the methodology developed in Section 2.

We estimate the coefficient of price to be included in [-1.992, -1.956] with a 95 percent probability, which is statistically smaller than the estimate from the model with exogenous market structure in Column 2 of Table 4.

We estimate λ for the exogenous entry case to be in the interval [0.300, 0.566] (Column 2 of Table 4), while in the endogenous entry case we estimate λ to be included in [0.116, 0.144]. Thus, we find that the within group correlation in unobservable demand is also estimated with a bias when we do not account for the endogenous market structure.

Overall, these sets of results lead us to over-estimate the elasticity of demand and under-estimate the market power of airline firms when we maintain that market structure is exogenous. To see this, we compare the implied mean elasticities in the bottom panel of Table 4. The mean elasticity for the exogenous market structure case is [-9.429, -7.020], while the we estimate the mean elasticity to be [-4.031, -3.864] when we allow for endogenous market structure. This leads to a difference in estimated markups: [22.219, 33.006] in the exogenous case compared with [44.416, 46.083] in the endogenous market structure case.

Next, we show the results for the estimates of the fixed cost parameters. Clearly, these are not comparable to the results from the previous model where market structure is assumed to be exogenous and fixed cost estimates are not recoverable. Column 3 of Table 4 shows the constant in the fixed cost inequality condition to be included in [-1.657, -1.288], and greater values of the variables *Nonstop Origin* and *Nonstop Destination* lead to lower fixed costs as one would expect if there were economies of density.

We compute the confidence interval for mean fixed costs to be [3931.25, 10949.74] dollars. To put these numbers in perspective, we need to recall that these are *market* fixed costs, and they are not the fixed costs paid to serve one of the legs of that market. Compared to the number of (uni-directional) non-stop segments served by an airline, the number of (uni-directional) markets served by that airline is many times larger. That is, a single non-stop leg will be part of the service on many markets, and we cannot infer the cost of serving the single non-stop leg, which is bound to be much larger, from the fixed costs of serving the markets.[34] The confidence interval of the ratio of the fixed costs over the variable profit is [0.110, 0.124], which means that the fixed costs are approximately 10 percent of the variable profits for the average carrier-market.

Next, we investigate the estimation results for the variance-covariance matrix. The variance of demand error is included in [1.336, 2.853] in Column 2 (exogenous market structure) and in [5.142, 5.826] in Column 3 (endogenous market structure). The variance of the marginal cost unobservables is estimated in [0.018, 0.055] in Column 2 and [0.334, 0.373] in Column 3. The larger values are explained in part by the fact that in the exogenous case, the distribution represents a *selected distribution* whereas in the endogenous case our estimates represent the full unselected distribution of the errors.

The covariance between the demand and marginal cost is positive in all three columns, with the caveat that it is not statistically different from zero in Column 3, although most of the confidence intervals is to the right of the zero value.

The covariance of the demand and fixed cost unobservables is estimated to be included in [0.631, 0.786] and the covariance between fixed and marginal costs unobservables is [1.119,1.215]. Carriers with unexpectedly (not predicted by observables in the model) high demand also have unexpectedly high fixed costs. Firms with unexpectedly high fixed costs have unexpectedly high marginal costs.

The variance covariance matrix implies that unobservables that lead to high demand correlate with higher fixed and marginal costs. This is intuitive if unobservables represent quality and the cost of quality – higher quality increases demand but it comes at some cost to the airline that we do not capture in the covariates. This is in contrast to an alternative story that is more akin to the selection on ability in labor markets where high demand firms are also low-cost producers.

Finally, we discuss the fit of the model. This consists of comparing the equilibrium market

[34]See Aguirregabiria and Ho (2012) for a rigorous discussion of this point.

structures, prices, and shares predicted by the model with those observed in the data. The particular way we think about model fit is necessitated by the fact that the model does not make unique predictions and that, if we were to compare aggregate statisitics, we would be comparing samples with different market structures. We compare model predictions to the data simulation-by-simulation and market-by-market and then tally up the number of times the model predictions are consistent with the data. For the model prediction to be consistent with the data, the data (e.g. a price) must lie in the 95 percent confidence interval.[35]

Specifically, we draw 100 parameters from the identified set, and simulate the model 200 times. For any given market structure in any given market, we construct the confidence interval for prices by taking the 2.5 and 97.5 percentile across parameter vectors. Then we compare the price for each airline for that market in the data to the confidence error for the predicted price. We do this again for product shares.[36]

The data lie within the confidence interval for prices 41.14 percent of the time and our model fits the shares 36.76 percent of the times. The model replicates the entry patterns well. In Table 5 we display the empirical entry probabilities for each airline along with the confidence intervals for entry probabilities predicted by the model. Additionally, the model fits the exact market structure 32.02 percent of the time (meaning all six carriers have the correct participation in the market) and the model predicts a given airline's entry correctly 74.48 percent of the time.[37] In our sample, 8.7 percent of markets are not served by any carrier while our model predicts this outcome in between 3.6 percent and 4.9 percent of markets.

[35] We construct the confidence interval for the prediction for an individual market in the same way we compute confidence intervals elsewhere, by sampling parameter vectors in the identified set.

[36] Note that in the typical econometric procedures used to estimate logit and random coefficient demand systems, shares and prices fit the data perfectly by construction. Our econometric procedure differs in that we do not have a completely flexible product characteristic residual that is allowed to adjust to exactly fit the data.

[37] These four numbers are not included in Table 4 for sake of brevity.

Table 5: Aggregate Entry Probabilities

	AA	DL	LCC	UA	US	WN	No Entry
Data	0.390	0.727	0.175	0.513	0.488	0.457	0.087
Model Prediction	[0.403, 0.436]	[0.737, 0.784]	[0.178, 0.215]	[0.521, 0.560]	[0.492, 0.533]	[0.439, 0.488]	[0.036, 0.049]

Note: Entry probabilities across all markets in the sample described in the text. Intervals for the model are constructed using the sub-sampling routine described in the text.

6 The Economics of Mergers When Market Structure is Endogenous

We present results from counterfactual exercises where we allow a merger between two firms, American Airlines and US Airways. A crucial concern of a merger from the point of view of a competition authority is the change in prices after the merger. It is typically thought that mergers imply greater concentration in a market, which, in turn, implies an increase in prices. However, in reality changes in the potential set of entrants along with changes in costs and demand after a merger may lead firms to optimally enter or exit markets. For example, cost synergies for the merged firm may cause entry into a new market to be profitable. Or, after the merger of the two firms, there might be room in the market for another entrant. Or, if demand is greater for the new merged firm, it may be able to steal market share from a rival such that the rival can not profitably operate.

Our methodology is ideally suited to evaluate both the endogenous price responses and the endogenous market structure responses as a consequence of a merger. Importantly, as we discuss below, changes in market structure imply changes in prices, and vice versa, so incorporating optimal entry decisions into a merger analysis is crucial for understanding the total effect of mergers on market outcomes. Section 9 of the Horizontal Merger Guidelines (08/19/2010) of the Department of Justice states that entry alleviates concerns about the adverse competitive effects of mergers. In contrast, the canonical model of competition among differentiated products takes as exogenous the set of competing products (eg BLP and Nevo, 2001), and thus the post-merger and pre-merger market structures are the same, except that the products are now owned by a single firm.[38]

[38]Mazzeo et. al. (2014) make a similar argument. They quantify the welfare effects of merger with

6.1 The Price and Market Structure Effects of the AA-US Merger

To simulate the effects of the AA-US Airways merger for a particular market, we use the following procedure. If US Airways (US) was a potential entrant we delete them and consider American (AA) the surviving firm. If American is a potential entrant before the merger, they continue to be a potential entrant after the merger. If American (AA) was not a potential entrant and US Air was a potential entrant before the merger, we assume that after the merger American is now a potential entrant. If neither firm was a potential entrant before the merger this continues after the merger.

We consider four different assumptions about what it means for AA and US to merge. The four assumptions underscore the key observation that post-merger efficiencies could come from both observed and unobserved features of the carriers. Thus, the different assumptions that we discuss next have to do with potential efficiencies from the merger, and have two aims: to check the robustness of the results of the counterfactual exercise, and, to help with the interpretation of those empirical results.

First, we consider a case where the surviving firm, AA, takes on the best observed and unobserved characteristics of both pre-merger carriers and call this the "Best Case" scenario.[39] More specifically, we combine the characteristics of both firms and assign the "best" characteristic between AA and US to the new merged firm. For example, in the consumer utility function, our estimate of *Origin Presence* is positive, so, after the merger, we assign the maximum of *Origin Presence* between AA and US to the post-merger AA. For marginal costs, we assign the highest level of *Origin Presence* between AA and US to the post-merger AA. And for fixed costs, we assign the highest level of *Nonstop Origin* and *Nonstop Dest*

endogenous enty/exit in a computational exercise using a stylized model that is similar to our model. In contrast, we provide a methodology to *estimate* an industry model and perform a merger analysis using those estimates. Also, we allow for multiple equilibiria in both estimation and the merger analysis, whereas Mazzeo et. al. (2014) assume a unique outcome from a selection rule based on ex ante firm profitability.

[39]This is the "best case" scenario that the firms would be able to present in court to make the strongest case that the merger is pro-competitive. Our reasoning for choosing to look at the "best case" scenario from the merging parties' viewpoint is that a merger should definitively not be allowed if there are no gains even under such scenario. However, this case may cause the exit of some firms or prices to rise in some markets, so this might not be, ex ante, the best case from the point of view of the regulator.

between AA and US to the post-merger AA. We implement the same procedure for the unobserved shocks. We use the same simulation draws from estimation for the merger scenario, and we assign the “best” simulation draw (for utility the highest and for costs the lowest) between AA and US to the post-merger AA. Our second scenario closely follows the “best case” scenario, but AA inherits only the best observable characteristics, and we assume the new firm inherits AA’s previous unobservables. The results are presented as a subcase called “only observables.” Our third scenario assumes that the new firm inherits the best observables and gets a new draw for the unobservables, which we term “new unobservables.” We simulate these two sub-cases to help us quantify the relative importance to the merger simulations of the efficiency in observables versus unobservables.

Lastly, we consider a scenario where the surviving firm takes on the mean values of the observed and unobserved characteristics from the two pre-merger firms, and call this the “Average Case Scenario.”

Table 6: Aggregate Effects of Merger, per market ($)

	Mean Fare	Mean Consumer Welfare	Total AA+US Profit*	Total DL+UA Profit*
Pre-merger	[197.72, 205.30]	[171813, 294556]	[427, 721]	[941, 1,580]
Post-merger				
Best Case	[199.81, 207.84]	[302562, 504337]	[1639, 2812]	[906, 1,527]
... only observables	[195.55, 202.90]	[193839, 322185]	[617, 1048]	[937, 1,577]
... new unobservables	[195.56, 202.89]	[193394, 321417]	[614, 1039]	[937, 1,577]
Average Case	[202.96, 210.66]	[156677, 268422]	[284, 489]	[948, 1590]

Note: Confidence intervals are constructed using the sub-sampling routine described in the text. * In millions of USD – sum of profit across all markets. *Mean Consumer Welfare* is the average market-level consumer welfare across all markets.

In Table 6, we present confidence intervals for *aggregate* statistics to provide an industry wide analysis of how an hypothetical merger would impact market structure, prices and consumer and producer surplus. The rows in Table 6 represent the pre-merger predictions of the model (first row) and the four scenarios we consider after the merger. The first column is the 95% confidence interval for the median fare (share weighted across markets). The second column is the median consumer welfare across markets, the third column is the total profit for AA and US (summed over all markets) and the fourth column is the sum of total profit for Delta and United.[40]

Under the “Best Case” scenario, the confidence interval for average prices is slightly greater than the baseline, although the two intervals overlap, [197.72, 205.30] versus [199.81, 207.84]. Consumer welfare would increase substantially from [171813, 294556] to [302562, 504337], as would the the profit of the new merged firm compared with the sum of the pre-merger AA and US profit, [427, 721] to [1639, 2812]. This welfare increase is likely unreasonably large, but highlights the importance of merger efficiency assumptions, as our other estimates of consumer welfare are much more moderate. In the sub-case where only the best observable characteristics are inherited by the merged firm, consumer welfare goes up by less than the best case, as does the merged firm’s profit. The “new unobservables” case is similar. In the case where the new firm inherits the average pre-merger characteristics, the results are much different. The confidence interval for consumer surplus is shifted lower, as is the merged-firm’s proft. In this case, inheriting the average characteristic is not enough efficiency to overcome the merger paradox. The sum of DL and UA profit remains roughly unchanged. We will explore the mechanisms for these changes below, which are partially due to the fact that our model endogenizes product market structure changes due to a merger. Overall, there are substantial potential efficiency gains from the merger, but this crucially dependent on assumptions about cost synergies due to the merger.

In Table 7 we report changes in predicted entry probabilities after the merger for all four cases. Specifically, we display 95% confidence interval for entry probabilities for each of the airlines for the baseline and all four merger scenarios. After the merger, AA’s likelihood of entry increases substantially in the best case scenario, from entry in [0.403, 0.436] to [0.807, 0.832] of markets. The increase in entry is not surprising given that AA inherits all of USAir’s potential markets. This happens at the expense of the other airlines, who see slight decreases in entry probabilities, even though they face one fewer potential entrant. In the other cases, AA sees a modest increase in the number of markets served, and the other airlines realize very slight increases in aggregate entry probabilities. In the remaining

[40]To compute changes in welfare we consider the log-sum logit compensating variation formula, see Train (2009).

discussion in this section, we go deeper into the mechanisms that explain these aggregate changes by considering changes in particular types of markets.

Table 7: Entry Probabilities, Post-merger

	AA	DL	LCC	UA	US	WN
Pre-merger	[0.403, 0.436]	[0.737, 0.784]	[0.178, 0.215]	[0.521, 0.560]	[0.492, 0.533]	[0.439, 0.488]
Post-merger						
Best Case	[0.807, 0.832]	[0.730, 0.778]	[0.175, 0.212]	[0.514, 0.554]	–	[0.436, 0.485]
... only observables	[0.631 0.672]	[0.736 0.784]	[0.178 0.215]	[0.521, 0.560]	–	[0.439, 0.487]
... new unobservables	[0.631 0.672]	[0.736, 0.783]	[0.178, 0.214]	[0.520, 0.559]	–	[0.439 0.487]
Average Case	[0.632, 0.675]	[0.738, 0.785]	[0.179, 0.215]	[0.523, 0.561]	–	[0.440, 0.488]

Note: Entry probabilities across all markets in the sample described in the text. Confidence intervals are constructed using the sub-sampling routine described in the text.

We begin our detailed analysis by looking at two sets of markets that are at the polar opposites in terms of post-merger effects: markets that were not served by any airline before the merger; and markets that were served by American and USAir as a duopoly before the merger. These are natural starting points because we want to ask whether new markets could be profitably served as a consequence of the merge of American and USAir, which is clearly a strong reason for the antitrust authorities to allow for a merger to proceed. We also want to examine pre-merger duopolies, which are markets that are most likely to see high prices and larg welfare losses post-merger.

In the following tables we report the likelihood of observing particular market structures and expected percentage change in prices conditional on a particular market structure transition. Table 8 is a simple "transition matrix" that relates the probability of observing a market structure post-merger (columns) conditional on observing a market structure pre-merger (rows).[41] The 2 x 2 table consists of the two pre-merger market structures, with no firm in the market and with a duopoly of US and AA. The post-merger market structures are those markets with no firm in the market and with a monopoly of AA/US.[42]

Table 8 shows that under the *Best Case Scenario* the probability that the merged firm AA/US will enter a market that was not previously being served is between 48.9 and 52.9

[41] Although our model is static, we use the terminology "transition" in order to convey predicted changes pre-merger to post-merger.

[42] The complete transition table would be of dimension 64 x 32 for each pre-merger market structure, which we do not present for practical purposes. Instead, we take slices of these tables.

percent, which is a substantial and positive effect of the merger that would be ignored by the standard economic analysis with exogenous market structure. We also find that there is a probability between 99.4 and 100 percent that the merged firm would serve a market as a monopolist that both independent firms were serving pre-merger. In those two-to-one cases the merged firm would charge a *lower* price (between -8.4 and -6.0 percent) due to the efficiency gains from the merger.

The predictions from the other scenarios are remarkably different, which illustrates the importance of the assumptions we make on the observed and unobserved characteristics of the merged firm. More specifically, under the *Average Case* we find that the probability that the merged firm AA/US will enter a market that was not previously being served is between 16.3 and 18.4 percent, much lower than the *Best Case*. The prices would now *increase* by 7.2 to 9.2 percent for those markets.

Table 8: Market Structures in AA and US Monopoly and Duopoly Markets

	Post-merger Entry		Post-merger %Δ Price
Pre-merger	No Firms	AA Monopoly	AA Monopoly
Best Case Scenario			
No Firms	[0.471, 0.511]	[0.489, 0.529]	–
AA/US Duopoly	[0.000, 0.000]	[0.994, 1.000]	[-8.4, -6.0]
...only observables			
No Firms	[0.677, 0.711]	[0.289, 0.323]	–
AA/US Duopoly	[0.000, 0.000]	[0.952, 0.969]	[-13.5, -12.7]
...new unobservables			
No Firms	[0.504, 0.539]	[0.461, 0.496]	–
AA/US Duopoly	[0.255, 0.297]	[0.604, 0.632]	[2.4, 5.6]
Average Case Scenario			
No Firms	[0.816, 0.837]	[0.163, 0.184]	–
AA/US Duopoly	[0.000, 0.003]	[0.948, 0.965]	[7.2, 9.2]

Next, we can investigate how the entry of the other potential entrants would change the prices in those markets where AA and US were a duopoly before the merger. Table 9 shows the probability that one of the other four competitors would enter, and the corresponding change in AA's price, in markets where there was a duopoly of American and USAir pre-merger.

Under the *Best Case Scenario* (top panel of Table 9) we find very little evidence that other competitors would enter. In fact, in the other three cases, there is only a small chance that a carrier would replace US in a previous AA-US duopoly. The most likely carriers to replace US are Delta and United, the two other major airlines. In those cases, we would expect prices to change by between -0.1 and 14.3 percent (average case, Delta) or between -0.7 and 18.1 percent (average case, United).

Table 9: Entry in former AA and US Duopoly Markets

Best Case Scenario	Duopoly AA/US & DL	Duopoly AA/US & LCC	Duopoly AA/US & UA	Duopoly AA/US & WN
Prob mkt structure	[0.000, 0.003]	[0.000, 0.001]	[0.000, 0.002]	[0.000, 0.001]
Percent Change in price of AA	[-0.098, 0.251]	[-0.249, 0.491]	[-0.046, 0.812]	[-0.113, 0.771]
... only observables				
Prob mkt structure	[0.013, 0.023]	[0.001, 0.004]	[0.008 0.016]	[0.003, 0.010]
Percent Change in price of AA	[-12.4 -7.0]	[-10.9, 3.3]	[-11.4, -3.7]	[-17.1, -2.0]
...new unobservables				
Prob mkt structure	[0.013, 0.024]	[0.001, 0.005]	[0.013, 0.021]	[0.003, 0.010]
Percent Change in price of AA	[-14.1, 9.1]	[-11.6, 35.6]	[-6.3, 11.8]	[-14.4, 30.6]
Average Case Scenario				
Prob mkt structure	[0.014, 0.023]	[0.000, 0.003]	[0.009, 0.018]	[0.003, 0.009]
Percent Change in price of AA	[-0.1, 14.3]	[-18.9, 46.8]	[-0.7, 18.1]	[-8.2, 26.9]

We now take a different direction of investigation. Instead of focusing on markets where there would be an ex-ante concern that prices increase after the merger, we explore in more depth the possible benefits of a merger, which could allow a new, possibly more efficient, firm to enter into markets that were monopolies pre-merger.

In Table 10 we consider the likelihood that after its merger with US, AA enters a market where it was *not present* before the merger. In this table we only consider those markets that were monopolies before the merger. In the first column we display the likelihood that AA replaces the monopolist after the merger, and in the second column we display the likelihood that AA joins the monopolist and forms a duopoly after the merger. For example, AA would replace DL as a monopolist with a probability between 1.6% and 2.3%, for the "Best Case

Scenario." It is much more likely that AA enters to form a duopoly, between 50% and 53.6%, and the DL prices would fall by roughly 2% in that case. AA is more likely to replace an LCC than other airlines, and in all cases of duopoly we should expect lower prices on the order of one to two percent. Under "Average Case Scenario" the likelihood of entry is much less than in the "Best Case Scenario." These results highlight the potential benefits of the merger. They also highlight, again, that the merged firm faces a stronger competition in entry from the other major carriers.

The intuition for the new market entry by AA/US and the corresponding changes in prices is straightforward. Under our assumptions about the merger, the new firm will typically generate higher utility and/or have lower costs in any given market than each of AA and US did separately before the merger. Low costs will promote entry of AA and lower prices for rivals after entry (in our model prices are strategic complements) and higher utility will promote entry by AA and upward price pressure, or even lead to exit by incumbents, as we predict in those monopoly markets where AA/US replaces the incumbent.

In Table 11, we focus on markets where AA is already present in the market and another incumbent duopolist *exits* after the merger. There are two reasons why a competitor would drop out of a market after a merger. First, after the merger AA might become more efficient in terms of costs, therefore lowering price and making it difficult for the rival to earn enough variable profit to cover fixed costs.[43] Second, AA might become more attractive to consumers after the merger and steal business from rivals. For ease of exposition we only consider markets where AA and other incumbents were in the market, and we do not report the results for the other merging firm, USAir.

The first row of Column 1 in Table 11 shows that, for the *Best Case Scenario*, there is a probability between 1.3 and 2.1 percent that DL will leave the duopoly market with AA after the merger. In such cases, AA's price will be between 3.1% and 15.2% higher. Overall the greatest likelihood of exit is for the LCC airline, along with the highest expected change

[43]AA could either experience a decrease in marginal costs, or a decrease in fixed costs. For the fixed costs case, AA could have been a low marginal costs firm before the merger, but high fixed costs prevented entry. After the merger, a decrease in fixed costs could lead to entry with the already low marginal costs.

Table 10: Post-merger Entry of AA in Former Monopolies

	AA Replacement	AA Entry	
Pre-merger Firm	Entry Probability	Entry Probability	Price Change (%)
Best Case Scenario			
DL	[0.016, 0.023]	[0.500, 0.536]	[-2.6, -2.0]
LCC	[0.038, 0.062]	[0.420, 0.472]	[-2.2, -1.7]
UA	[0.024, 0.037]	[0.494, 0.533]	[-2.5, -2.0]
WN	[0.019, 0.030]	[0.439, 0.475]	[-2.0, -1.6]
...only observables			
DL	[0.007, 0.011]	[0.303, 0.333]	[-2.4, -1.8]
LCC	[0.015, 0.034]	[0.237, 0.271]	[-2.0, -1.5]
UA	[0.010, 0.018]	[0.270, 0.297]	[-2.2, -1.8]
WN	[0.010, 0.016]	[0.248, 0.277]	[-1.8, -1.4]
...new unobservables			
DL	[0.013, 0.020]	[0.503, 0.534]	[-2.5, -2.0]
LCC	[0.037, 0.065]	[0.415, 0.460]	[-2.1, -1.7]
UA	[0.023, 0.035]	[0.502, 0.540]	[-2.5, -2.0]
WN	[0.017, 0.027]	[0.435, 0.465]	[-1.9, -1.5]
Average Case Scenario			
DL	[0.002, 0.004]	[0.166, 0.186]	[-1.5, -1.1]
LCC	[0.005, 0.018]	[0.151, 0.189]	[-1.7, -1.2]
UA	[0.003, 0.006]	[0.177, 0.201]	[-1.6, -1.2]
WN	[0.003, 0.006]	[0.150, 0.168]	[-1.3, -0.9]

in price. In the cases where AA's new price will be slightly lower than before the merger, this suggests that the cost efficiency effects dominate the business stealing effects. In contrast, when there are not sizable efficiencies, the likelihood of rival exit is small, only up to 0.6% for LCC in the *Average Case* scenario, but as low as 0.[44]

Table 11: Likelihood of Exit by Duopoly Competitors after AA-US Merger

Pre-merger Firm	Probability of Exit	AA Price Change (%)
Best Case Scenario		
DL	[0.013, 0.021]	[3.1, 15.2]
LCC	[0.020, 0.054]	[-2.2, 27.5]
UA	[0.022, 0.033]	[-1.5, 5.2]
WN	[0.013, 0.027]	[1.7, 19.2]
...only observables		
DL	[0.007, 0.013]	[-15.1 -9.4]
LCC	[0.008, 0.030]	[-21.0 -3.3]
UA	[0.013, 0.021]	[-15.6 -9.8]
WN	[0.007, 0.015]	[-21.7 -12.2]
...new unobservables		
DL	[0.006 0.011]	[12.1 29.7]
LCC	[0.008 0.029]	[3.1 46.7]
UA	[0.008 0.016]	[11.6 36.0]
WN	[0.006 0.013]	[7.8 50.3]
Average Case Scenario		
DL	[0.001, 0.002]	[31.6, 78.9]
LCC	[0.000, 0.006]	[-58.3, 95.6]
UA	[0.001, 0.004]	[8.2, 46.7]
WN	[0.000, 0.003]	[16.2, 142.8]

[44]Even in the *Average Case Scenario* AA may experience some efficiencies. Although AA could also have higher costs after the merger, which is likely in cases where AA was already an entrant because they were probably more efficient than US in such markets.

6.2 The Economics of Mergers at a Concentrated Airport: Reagan National Airport

The Department of Justice reached a settlement with American and USAir to drop its antitrust challenge if American and USAir were to divest assets (landing slots and gates) at Reagan National (DCA), La Guardia (LGA), Boston Logan (BOS), Chicago O'Hare (ORD), Dallas Love Field (DAL), Los Angeles (LAX), and Miami International (MIA) airports. The basic tenet behind this settlement was that new competitors would be able to enter and compete with AA and US, should the new merged airline significantly raise prices.

We conduct a counterfactual exercise on the effect of the merger in markets originating or ending at DCA. These markets were of the highest competitive concern for antitrust authorities because both merging parties had a very strong incumbent presence.[45]

Table 12 reports the results of a counterfactual exercise that looks at the exit of competitors and changes in price in markets with DCA as an endpoint that were served by both AA and US before the merger.[46]

Let us begin with the triopoly AA/US/DL. We find that there is a significant likelihood that the market becomes more concentrated. The AA/US/DL market turns into a AA/DL market with probability [0.964, 1.000] for the "Best Case" scenario [0.964, 1.000] for the "Average Case" scenario, for example. We find that in all cases, this would not result in a significant rise in price.

In none of the pre-merger markets where AA and US were both present, LCC or WN are likely to replace US. This finding confirms that DL and UA offer a service that is a closer substitute to the one provided by AA and US than WN and LCC do. This also justifies the DOJs concern that airport slots go to Southwest or Jet Blue instead of incumbent majors.

For market with four firms,, the most likely outcome across all cases is a consolidation to AA/DL/UA. For the *Best Case Scenario*, this is accompanied by lower prices. For the

[45]Although we do not model slot constraints, our model would provide crucial information on which airports would be the ones where anticompetitive concerns would be the most relevant and the results suggest DCA was indeed one where there should have been competitive concerns regarding AA/US.

[46]None of the DCA markets in our sample were a AA/US duopoly before the merger, so we look at other market structures that involve both airlines.

Average Case Scenario, we find evidence of higher prices in several markets. Assumptions about merger efficiencies are crucial for determining the effects of the merger.

Overall, our results suggest that the decisions made by the Department of Justice to facilitate the access to airport facilities to new entrants were justified under certain assumptions about merger efficiencies (our "Average Case"), and should help control the post-merger increase in prices and promote low-cost carrier coverage at DCA.

Table 12: Post-merger entry and pricing Reagan National Airport

	Post-merger Market Structure				
Pre-merger Markets	AA/DL	AA/UA	AA/DL/LCC	AA/DL/UA	AA/DL/WN
Best Case Scenario					
AA,US,DL Markets					
Mkt Struct. Transitions	[0.964, 1.000]	[0.000, 0.000]	[0.000, 0.000]	[0.000, 0.000]	[0.000, 0.000]
%Δ Shares Weighted Price	[-13.2, -9.9]	[n.a.]	[n.a.]	[n.a.]	[n.a.]
AA,US,DL,UA Markets					
Mkt Struct. Transitions	[0.002, 0.020]	[0.002, 0.020]	[0.000, 0.000]	[0.968, 0.993]	[0.000, 0.000]
%Δ Shares Weighted Price	[-28.3, -3.4]	[-34.3, -6.3]	[n.a.]	[-12.5, -11.0]	[n.a.]
Average Case Scenario					
AA,US,DL Markets					
Mkt Struct. Transitions	[0.959, 1.000]	[0.000, 0.000]	[0.000, 0.019]	[0.000, 0.030]	[0.000, 0.027]
%Δ Shares Weighted Price	[-0.8, 1.7]	[n.a.]	[-29.9, 21.4]	[-10.0, 47.9]	[-30.8, 33.2]
AA,US,DL,UA Markets					
Mkt Struct. Transitions	[0.000, 0.002]	[0.000, 0.004]	[0.000, 0.000]	[0.975, 0.999]	[0.000, 0.000]
%Δ Shares Weighted Price	[-12.4, 11.8]	[-41.1, 24.3]	[n.a.]	[0.8, 1.8]	[n.a.]

Note: Counterfactual predictions for markets with DCA as one endpoint. Pre-merger market structure of AA/DL/UA and AA/US/DL/UA.

7 Conclusions

We provide an empirical framework for studying the quantitative effect of self-selection of firms into markets and its effect on market power in static models of competition. The counterfactual exercise consists of a merger simulation that allows for changes in market structures, and not just in prices. The main takeaways are: i) allowing for the selection of firms into markets based on unobservables can lead to different estimates of price elasticities and markups than those that we find when we assume that market structure is exogenous to the pricing decision; ii) this in turn leads to potentially important differences from exogenous entry models in the predicted response to policy counterfactuals, such as merger simulations.

More generally, this paper contributes to the literature that studies the effects that mergers or other policy changes have on the prices and structure of markets, and consequently the welfare of consumers and firms. These questions are of primary interest for academics and researchers involved in antitrust and policy activities.

One extension of our model is to a context where firms can change the characteristics of the products they offer. To illustrate, consider Sovinsky Goeree (2008) who investigates the role of informative advertising in a market with limited consumer information. Sovinsky Goeree (2008) shows that the prices charged by producers of personal computers would be higher if firms did not advertise their products, because consumers would be unaware of all the potential choices available to them, thus granting greater market power to each firm. However, this presumes that the producers would continue to optimally produce the same varieties if consumers were less aware, while in fact one would expect them to change the varieties available if consumers had less information, for example by offering less differentiated products. It is possible to extend our framework to investigate questions like this where firms choose product characteristics.

Also, the proposed methodology can be applied in all economic contexts where agents interact strategically and make both discrete and continuous decisions. For example, it can be applied to estimate a model of household behavior where a husband and a wife must decide whether to work and how many hours.

We also show that our results depend, as one would expect, on the assumptions that we make on the efficiency gains from a merger. First, quantifying the efficiency gains from a merger is a difficult empirical exercise that is at the center of all merger investigations by the federal agencies, and which is often based on confidential *accounting* cost data. Second, even if current and past *accounting* cost data are available, normally it takes time for the efficiencies to be fully realized. We believe that our approach, which is based on being upfront and clear about the efficiency gains, provides a promising path for future research in antitrust merger research. More generally, determining the efficiency gains from a merger is a difficult empirical exercise that is at the center of all merger investigations by the federal agencies. In some cases it takes a long time for the efficiencies to be fully realized, and it is not always possible to identify their magnitude. Our approach shows how we can quantify these efficiencies under various plausible assumptions. We hope our approach provides a promising approach for future research in antitrust merger research.

To conclude, we summarize some of the limitations of our approach. There are several components/variables in the classical model (Bresnahan, 1987; or Berry, 1994) that are taken as exogenous. More specifically, the classical model takes as exogenous: the entry decision; the location decision in the space of the observed characteristics; and the location decision in the space of the unobserved characteristics. Our goal is to relax one of those – the decision to participate in the market, and continue to assume that the location in the space of the observed and unobserved characteristics is exogenous. We leave to future work the next step, which is to relax those assumptions as well. Some recent important work in that direction is in Li et al. (2017). Also, Petrin and Seo (2017) propose an interesting approach for the problem of endogenous product characteristics (conditional on entry) by using information from the firms' necessary optimality conditions for the choice of product characteristics.

References

Aguirregabiria, Victor and Chun-Yu Ho. 2012. "A dynamic oligopoly game of the US airline industry: Estimation and policy experiments." *Journal of Econometrics* 168 (1):156 – 173. URL http://www.sciencedirect.com/science/article/pii/S030440761100176X. The Econometrics of Auctions and Games.

Berry, Steven and Panle Jia. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics* 2 (3):1–43.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society* :841–890.

Berry, Steven T. 1990. "Airport presence as product differentiation." *The American Economic Review* :394–399.

———. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica: Journal of the Econometric Society* :889–917.

———. 1994. "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics* :242–262.

———. 1999. "Free entry and social inefficiency in radio broadcasting." *The RAND Journal of Economics* 30 (3):397–420.

Bresnahan, Timothy F. 1987. "Competition and collusion in the American automobile industry: The 1955 price war." *The Journal of Industrial Economics* :457–482.

———. 1989. "Empirical studies of industries with market power." *Handbook of industrial organization* 2:1011–1057.

Bresnahan, Timothy F and Peter C Reiss. 1990. "Entry in monopoly market." *The Review of Economic Studies* 57 (4):531–553.

———. 1991. "Entry and competition in concentrated markets." *Journal of Political Economy* :977–1009.

Brueckner, Jan K and Pablo T Spiller. 1994. "Economies of traffic density in the deregulated airline industry." *Journal of Law and Economics* :379–415.

Chernozhukov, Victor, Han Hong, and Elie Tamer. 2007. "Estimation and confidence regions for parameter sets in econometric models1." *Econometrica* 75 (5):1243–1284.

Ciliberto, Federico, Amalia Miller, Helena Skyt Nielsen, and Marianne Simonsen. 2010. "Playing the fertility game at work." *Unpublished manuscript* .

Ciliberto, Federico and Elie Tamer. 2009. "Market structure and multiple equilibria in airline markets." *Econometrica* 77 (6):1791–1828.

Ciliberto, Federico and Jonathan W Williams. 2014. "Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry." *Inference on Conduct Parameters in the Airline Industry (February 12, 2014)* .

Cohen, Andrew M and Michael J Mazzeo. 2007. "Market structure and competition among retail depository institutions." *The Review of Economics and Statistics* 89 (1):60–74.

Draganska, Michaela, Michael Mazzeo, and Katja Seim. 2009. "Beyond plain vanilla: Modeling joint product assortment and pricing decisions." *QME* 7 (2):105–146.

Eizenberg, Alon. 2014. "Upstream Innovation and Product Variety in the US Home PC Market*." *The Review of Economic Studies* :rdu004.

Elzinga, Kenneth G and David E Mills. 2009. "Predatory Pricing in the Airline Industry: Spirit Airlines v. Northwest Airlines (2005)." *The Antitrust Revolution, 5th Edition, edited by J. Kwoka and L. White, Forthcoming* .

Fan, Ying. 2013. "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market." *American Economic Review* 103 (5):1598–1628.

Fan, Ying and Chenyu Yang. 2017. "Competition, Product Proliferation and Welfare: A Study of the U.S. Smartphone Market." Tech. rep., mimeo, University of Michigan.

Gandhi, Amit and Jean-François Houde. 2016. "Measuring substitution patterns in differentiated products industries." *University of Wisconsin-Madison and Wharton School* .

Gentry, Matthew and Tong Li. 2014. "Identification in Auctions With Selective Entry." *Econometrica* 82 (1):315–344.

Goolsbee, Austan and Chad Syverson. 2008. "How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines." *The Quarterly Journal of Economics* 123 (4):1611–1633.

Gronau, Reuben. 1974. "Wage Comparisons–A Selectivity Bias." *The Journal of Political Economy* :1119–1143.

Heckman, James J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." In *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER, 475–492.

———. 1979. "Sample selection bias as a specification error." *Econometrica: Journal of the econometric society* :153–161.

Ho, Katherine. 2008. "Insurer-Provider Networks in the Medical Care Market." *American Economic* .

Jeziorksi, Przemyslaw. 2014a. "Effects of mergers in two-sided markets: The U.S. radio industry." *American Economic Journal: Microeconomics* 6 (4):35–73.

———. 2014b. "Estimation of cost synergies from mergers: The U.S. radio industry." *RAND Journal of Economics* 45 (4):816–846.

Lee, Robin. 2013. "Vertical Integration and Exclusivity in Platform and Two-Sided Markets." *American Economic Review* 103 (7):2960–3000.

Li, Sophia, Joe Mazur, Yongjoon Park, James Roberts, and Andrew Sweeting. 2017. "Endogenous and Selective Service Choices After Airline Mergers." *working paper, University of Maryland* .

Li, Tong and Bingyu Zhang. 2015. "Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects." *American Economic Journal: Microeconomics* 7 (2):188–214.

Li, Tong and Xiaoyong Zheng. 2009. "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions." *The Review of Economic Studies* 76 (4):1397–1429.

Mazzeo, Michael J. 2002. "Product choice and oligopoly market structure." *RAND Journal of Economics* :221–242.

Nevo, Aviv. 2000. "Mergers with differentiated products: The case of the ready-to-eat cereal industry." *The RAND Journal of Economics* :395–421.

Nocke, Volker and Nicolas Schutz. 2018. "Multiproduct-Firm Oligopoly: An Aggregative Games Approach." *Econometrica* 86 (2):523–557.

Olley, G Steven and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica: Journal of the Econometric Society* :1263–1297.

Pakes, Ariel, J Porter, Joy Ishii, and Kate Ho. 2015. "Moment Inequalities and Their Application." *Econometrica* 83:315–334.

Panzar, John C. 1979. "Equilibrium and welfare in unregulated airline markets." *The American Economic Review* :92–95.

Petrin, Amil and Boyoung Seo. 2017. "Identification and Estimation of Discrete Choice Demand Models when Observed and Unobserved Product Characteristics are Correlated." .

Reiss, Peter C and Pablo T Spiller. 1989. "Competition and entry in small airline markets." *Journal of Law and Economics* :S179–S202.

Roberts, James and Andrew Sweeting. 2013. "When Should Sellers Use Auctions?" *American Economic Review* 103 (5).

Rosse, James N. 1970. "Estimating cost function parameters without using cost data: Illustrated methodology." *Econometrica: Journal of the Econometric Society* :256–275.

Schmalensee, Richard. 1989. "Inter-industry studies of structure and performance." *Handbook of industrial organization 2* 2.

Seim, Katja. 2006. "An empirical model of firm entry with endogenous product-type choices." *RAND Journal of Economics* :619–640.

Sovinsky Goeree, Michelle. 2008. "Limited information and advertising in the US personal computer industry." *Econometrica* 76 (5):1017–1074.

Sweeting, Andrew. 2013. "Dynamic Product Positioning in Differentiated Product Industries: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry." *Econometrica* 81 (5).

Tamer, Elie. 2003. "Incomplete simultaneous discrete response model with multiple equilibria." *The Review of Economic Studies* 70 (1):147–165.

Train, Kenneth E. 2009. *Discrete choice methods with simulation.* Cambridge university press.

Appendix: Market Structure and Competition in Airline Markets

Federico Ciliberto* Charles Murry† Elie Tamer‡

February 10, 2020

A Identification Details

This section closely follows Ciliberto and Tamer (2009) (henceforth CT), and we refer to that paper for further reading.

We provide a set of sufficient conditions that guarantee point identification of the model parameters in equation (1) in the main text. These conditions are natural in this context and rely on large support regressors. Our inference methods do not require that these conditions be satisfied as the moment inequalities adapt to partial identification, but we give them here to give intuition as to what exogenous variation might be helpful for gaining identification.

Theorem 1 *Suppose* $\mathbf{Z} = (z_1, z_2)$ *is such that* $z_1|z_2, \mathbf{X}$ *has continuous support over the real line and that* $\gamma \neq 0$. *In addition, assume that* $E\big([X_i \vdots X_{3-i}][X_i \vdots V_i]' \mid z_i\big)$ *has full column rank for* $i = 1, 2$. *Suppose that there is Nash equilibrium play (possibly in mixed strategies) and that* $(\nu_1, \nu_2, \xi_1, \xi_2) \perp (\mathbf{X}, \mathbf{Z})$. *Then,*

1. *The parameters of the first two inequalities in (1) are identified as* $z_1, z_2 \to \infty$.

2. *In addition,* $(\beta, \alpha_1, \alpha_2)$ *are also point identified as* $z_1, z_2 \to \infty$.

The intuition for the above result is simple. Large support conditions are sufficient for point identification of the entry model (see Tamer, 2003). Now, for the outcome equation, we can do 2SLS *at infinity* as follows. For large values of z_1 (large negative or positive values depend on the sign of γ which can be learned fast by looking at whether large positive values of z_1 say correspond to higher likelihood of seeing a player 1 in the market) for example, player 1 is in the market with probability 1. Hence, we can use $\mathbf{X}_2$ as an instrument for V_1

*University of Virginia, ciliberto@virginia.edu.
†Boston College, charles.murry@bc.edu.
‡Harvard University, elietamer@fas.harvard.edu.

and do 2sls on the first outcome equation conditional on the event that $z_1 \to \infty$. Driving player 1 to enter with probability 1 eliminates the correlation between ξ_1 and $y_1 = 1$ which allows us to use "standard methods" to estimate the first outcome equation. These methods would be based on the moment condition

$$E[(X_1', X_2')'\xi_1 | z_1 \to \infty] = 0$$

Hence, what is needed for the identification of outcome equation 1 for example (arguments for the second outcome equation are similar) is two excluded variables: a standard instrument X_2 and an excluded variable from the outcome equation, z_1 in this case, that takes large values and can influence the entry of player 1. Such a variable can be one that affects fixed costs only, but not variable costs and can be exogenously moved. In the standard case, the only needed condition is an instrument X_2. So, to control for the first stage, we are required to have another instrument that can take large values. Note that the identification results in the Theorem above do NOT require that 1) the joint distribution of the unobservables be known, but requires that those be independent of the exogenous regressors, and 2) that the players play pure strategies (also here, the results in the Theorem do not require that the sign of the Δ's be known but we maintain here that the sign of these is strictly negative). On the negative side, these point identification results based on large supports lead to slow rates of convergence which makes it hard to be used with standard data sets.

Without such large support conditions, it is unclear whether we get point identification and hence it is crucial that any inference methods used is robust to failure of point identification. Basing our inference on the derived *moment inequalities* does not require that the parameter is point identified. The confidence regions that these methods use are based on inverting test statistics like the following ones.

So, under the null that $\theta = \theta^*$, we have

$$H_0: \quad E[\mathbf{G}(\theta^*, S_1 y_1, S_2 y_2, V_1 y_1, V_2 y_2, y_1, y_2) | \mathbf{Z}, X] \leq 0 \quad \text{for all} \, (\mathbf{X}, \mathbf{Z}, t_1, t_2)$$

The next theorem provides the objective function that we use to define our test statistic.

Theorem 2 *Suppose the above parametric assumptions in model (1) are maintained. In addition, assume that* $(\mathbf{X}, \mathbf{Z}) \perp (\xi_1, \xi_2, \nu_2, \nu_2)$ *where the latter is normally distributed with mean zero and covariance matrix* Σ. *Then given a large data set on* $(y_1, y_2, S_1 y_1, V_1 y_1, S_2 y_2, V_2 y_2, \mathbf{X}, \mathbf{Z})$ *the true parameter vector* $\theta = (\delta_1, \delta_2, \alpha_1, \alpha_2, \beta, \gamma, \Sigma)$ *minimizes the nonnegative objective function below to zero:*

$$Q(\theta) = 0 = \int W(\mathbf{X}, \mathbf{Z}) \| \mathbf{G}(\theta, S_1 y_1, S_2 y_2, V_1 y_1, V_2 y_2, y_1, y_2) | \mathbf{Z}, X] \|_+ dF_{\mathbf{X}, \mathbf{Z}} \tag{A.1}$$

for a strictly positive weight function $(\mathbf{X}, \mathbf{Z})$.

The above is a standard conditional moment inequality model where we employ discrete valued variables in the conditioning set along with a finite (and small) set of t's.[1]

[1]It is possible to use recent advances in inference methods in moment inequality models with a continuum

B Computational Guide

The estimation algorithm compares moments from the data (which themselves depend on parameters) to model predicted analogues. The entry model generally predicts multiple equilibria, so for a given set of parameters and a given draw from the joint distribution of model errors, there may be multiple predictions from the model. Below, we detail the steps we use to estimate the model: (i) Evaluating the moments from the data, (ii) simulating the bounds on the moments from the model, (iii) comparing the data to the model, (iv) computation of optimization and inference, and (v) counterfactuals.

(i) Moments from the data

For a given guess of the parameters $\Theta^0 = (\alpha^0, \beta^0, \varphi^0, \gamma^0, \Sigma^0)$ we estimate the probability distribution functions for the residuals from the demand and supply equations. In the data, each market has an observed market structure, $\hat{e}_m \in E$. For all markets and all active carriers, we compute the following two residuals:

$$\hat{\xi}_{jm}^{\hat{e}_m} = log(s_{jm}) - log(s_0) - X'\beta^0 - \alpha^0 p_{jm} - \lambda ln(s_{jm|g}) \tag{A.2}$$

$$\hat{\eta}_{jm}^{\hat{e}_m} = ln(p_{jm} - [\frac{1-\lambda}{\alpha(1-\lambda s_{j|g} - (1-\lambda)s_j)}]) - \varphi W_{jm}, \tag{A.3}$$

where we are clear that the residuals are specific to a particular market structure, $\hat{e}_m$, j indexes carriers, and m indexes markets.

The moments we use in estimation are the joint distribution of these residuals. In practice, we compute the joint probability distribution (joint between supply, demand, and across all firms) function,

$$\Pr(\hat{\xi}^{\hat{e}} <= \mathbf{t_D}, \hat{\eta}^{\hat{e}} <= \mathbf{t_S} \mid \mathbf{X}, \mathbf{W}, \mathbf{Z}) \tag{A.4}$$

by constructing a histogram by binning up the domain of the residuals and counting the frequency of residuals in each bin.[2] As we describe next, the moments are constructed by taking differences between the bin-counts of the distribution of residuals from the data with the distribution of selected errors predicted by the model.

(ii) Model predictions

We construct the distribution of structural errors predicted to be selected by the model using simulation. For the same guess fo parameters, Θ^0, we make 100 draws from the joint

of moments, but these again will present computational difficulties especially in the empirical model we consider below. We detail in the next Section the exact computational steps that we use to ensure well behavior (and correct coverage) of our procedures.

[2] We estimate the CDF using histograms (using Matlab's HISTCOUNTS function). This takes much less computer memory (at least the way we are thinking about the problem). The dimensionality of the array that defines the histogram can be as small as 2-dimensional (a matrix) for a market with a single entrant and as large as 12-dimensional array for a market with six entrants (each firm has a demand and supply residual).

distribution of demand, marginal cost, and fixed cost errors, $\begin{pmatrix} \xi \\ \eta \\ \nu \end{pmatrix} \sim MVN(0, \Sigma^0)$ for every market. We solve for all possible equilibria in each market for each simulation draw. Solving for all possible equilibria involves finding a vector of profits (defined by the three vector-valued equations in Equation 16 in the main text) that are consistent with pure strategy Nash behavior for *every potential market structure*. Finding a vector of profits involves finding the Nash Equilibrium of prices for any particular market structure, which is itself the solution to a system of implicit non-linear equations in prices defined by the pricing first-order-conditions.[3] Because we have six potential entrants, we have 2^6 market structures to solve for profits for every simulation draw in each market.[4]

When there are no pure-strategy equilibria in the entry game, we know that there exists at least one equilibrium in mixed-strategies. In that case, which happens *very* rarely in our empirical analysis, we proceed as follows. First, we determine the firms for which it is a dominant strategy not to enter. Then, we know that there will be at least one mixed strategy equilibrium where one of the remaining firms assigns a positive probability to the entry decision. Finally, we count this observation-simulation as contributing to the upper bound of the CDF of the simulated errors for all those firms.[5]

We collect all of the simulated errors that are part of equilibrium play. For example, if in one market for one simulation AA and DL are the active firms, the structural errors associated with those two carriers are the model selected errors, $(\xi^{e^*}_{AA,mr}, \xi^{e^*}_{DL,mr}, \eta^{e^*}_{AA,mr}, \eta^{e^*}_{DL,mr})$, where r indexes simulation draws. To construct the joint distribution of selected errors, we follow the same procedure of binning up the domain of the errors (using the same bin cutoffs) as we used for the computation of the distribution of the residuals. However, because of multiple equilibria, there will be an upper limit to the distribution and a lower limit. The upper limit is defined when we do not include any errors from those simulation-markets with multiple equilibria, as we are agnostic about equilibrium selection. The lower limit includes all errors from markets with multiple equilibria. This procedure is analogous to the procedure in Ciliberto and Tamer (2009), and yields the upper and lower joint distributions of model selected errors:

$$\Pr\left(\xi^e_r <= \mathbf{t_D}, \eta^e_r <= \mathbf{t_S}; \begin{pmatrix} \xi \\ \eta \\ \nu \end{pmatrix} \in A^M_e | \mathbf{X}, \mathbf{W}, \mathbf{Z}\right) \tag{A.5}$$

[3]We employ a multi-method strategy for finding equilibrium prices. For a vast majority of the cases, iterating on the markup equation solves for the vector of equilibrium prices quickly. However, we also employ quasi-Newton root-finding when function iteration fails or moves slowly.

[4]We parallelize our code across markets. The 1 percentile of time that it takes to solve everything one time for all simulations for all markets is 10.2 seconds; the 99 percentile that it takes is 24.6 seconds. The median time that it takes is 12 seconds.

[5]For example, suppose that Firm 1 and Firm 2 are the only firms in a market, for a given simulation, for which entry is not a dominated strategy. Then, we maintain that the simulated errors for those two firms, for that simulation in that market, contribute to the upper bound of the CDF.

and

$$\Pr\left(\xi_r^e <= \mathbf{t_D}, \eta_r^e <= \mathbf{t_S}; \begin{pmatrix}\xi\\ \eta\\ \nu\end{pmatrix} \in A_e^U | \mathbf{X}, \mathbf{W}, \mathbf{Z}\right), \tag{A.6}$$

where (from Section 2) $\begin{pmatrix}\xi\\ \eta\\ \nu\end{pmatrix} \in A_e^U$ denotes the realizations of the errors that imply unique equilibria and $\begin{pmatrix}\xi\\ \eta\\ \nu\end{pmatrix} \in A_e^M$ denotes realization of the fixed cost error that imply multiple equilibria. To construct the density function from the simulations, we take a simple average over the psuedo-random Monte Carlo draws, indexed by r.

(iii) Constructing Moments

The moments we use for estimation involve comparing the distribution of residuals to the upper and lower bounds of the distribution of model selected errors. We take the squared difference of the bin counts that define the conditional joint distribution of residuals and model selected errors. We sum these squared differences across bin counts and across the different conditional distributions. Because of multiple equilibria, we only penalize the function if the cdf of the residuals is greater than the upper limit for the selected errors or less than the lower limit for the selected errors. Notice that there are essentially two types of ways the model will not fit the data: (1) conditional on a market structure and (X, Z, W), the residuals have a different distribution than the selected errors, and (2) the model predicts different market structures than the data.

To choose the t's in the grid, we proceed in two steps. First, we determine the distributions of the demand and marginal cost residuals when we estimate the model with GMM without selection. We use this to learn over what support the residuals are defined, in terms of their max and min value. Using this approach, we selected the following values for the t of the demand: [-10;-7.5;-5;-2.5;0;2.5;5;7.5;10]. And we chose the following values for the t of the marginal cost: [-2;-1.5;-1;-0.5;0;0.5;1;1.5;2], or one-fifth the scale of the demand errors. In our experimentation when estimating Column 1 of Table 4 (the exogenous case) we found that this proportional relationship is important. Ideally, one would want to have a very fine grid for the t but there is a trade-off because of memory limitations (explained above) associated with storing so many cells used to construct the histogram estimate.

(iv) Objective Function Minimization, Computational Details, and Inference

The minimization of the distance function given by Equation (A.1) in the main text and described above is computationally intensive because we have to use simulation methods to integrate two multi-dimensional distribution functions and then compare them. In addition to constructing the distribution functions, we need to solve for Nash equilibria in many markets, and for many possible combinations of firms in each market. We need to do these things many times because the objective function may be non-smooth and non-convex, so finding a set of parameters that minimize the objective function may be taxing.

The keys to finding the global minimum are:

1. Parallel computing;

2. Good initial guesses on as many parameters as possible;

3. Using many different starting values;

4. Using flexible minimization routines that mix different built-in algorithms.

Each one of these ingredients is important in finding a global minimum. Overall, we reach the area of the global minimum in approximately three days and the optimization is completed in approximately seven days. In practice, we have very good starting values for many of the parameters from IV regressions that do not account for endogenous entry.

For inference purposes, we continue the minimization longer in order to collect as many parameters close to the *argmin*. We use Matlab's optimization algorithms to sample the objective function and we save the results to get a snapshot of the surface of the function. In addition, we randomly and non-randomly sample parameters close to the minimum to achieve good coverage around the minimum in order to construct confidence regions.

Good Initial Guess on as many parameters as possible: The GMM estimation that assumes exogenous market structure provides us with natural starting values for the parameters of the utility and marginal cost functions.

To get starting values for the parameters in the fixed cost function, and for the remaining parameters in the variance covariance matrix, we proceeded as follows. We compute the total revenues (observed prices times observed quantities) minus the inferred variable costs (GMM inferred marginal costs times quantities). This difference is equal to the profit of the firm, plus the fixed costs. Therefore, this difference should be thought of as the upper bound on the fixed costs. We regress this difference on the exogenous variables that enter into the entry condition and saved these parameter estimates for our next step.

Many Starting Values: We start with multiple initial values, which are derived as follows. For each of the initial guesses above, we find a reasonable interval around those guesses in the sense that the intervals are on the same scale as the standard errors from the GMM estimation and the interval implies sensible economic predictions, for example positive marginal costs and markups that are not near zero. For example, for the price parameter estimate, which is equal to -0.0229 in the GMM, we prepare an interval equal to [-0.035, -0.010]. We repeat this exercise for all the parameters. An important remark: Recall that in order to limit the space over which to draw for the argmin, we have standardized all the exogenous variables. This stabilizes the search and allows us to limit the parameter search within the intervals [-2,2] for all the exogenous variables while running PATTERNSEARCH.[6]

[6]We did not find any of the parameters getting close to the bounds during the minimization process. Otherwise, we would have restarted the minimization with wider bounds for that parameter.

Next, we draw up to *50,000* independent random draws from these intervals. Out of these 50,000 starting values, we select the 10 that are associated with the lowest distance function values. This first step takes approximately one day of time, but we save these function evaluations, so this is itself part of the minimization and confidence function construction processes. Our next step is to use canned algorithms in the GLOBAL OPTIMIZATION TOOLBOX in Matlab to minimize the function starting from the first round of 10 lowest values.

Multiple Iterations of Flexible Minimization Routines: In our experimentation we have used different combinations of three canned algorithms in the GLOBAL OPTIMIZATION TOOLBOX in Matlab: SIMULANNEALBND, PATTERNSEARCH, and FMINSEARCHOS. We have found that PATTERNSEARCH provides the best minimization results after we draw the 50,000 initial values, as described above. Therefore, we take the best 10 values out of the 50,000, and run PATTERNSEARCH.

We have found that after 1 day, patternsearch converges to a new parameter value. We found that the new parameter value depends on the starting values, and that is why it is crucial to draw as many starting values as we do. This is because the distance function is highly nonlinear and the minimization problem is complex.

At this point we take the 10 local minima after running PATTERNSEARCH, and reiterate the process described in the Section above (Many Starting Value), but choosing tighter bounds. We draw another 50,000 independent random draws. We run patternsearch again on the 10 that are associated with the lowest distance function.

In our work, we have run few iterations of this two-step process of i) drawing randomly, then ii) using PATTERNSEARCH this process this time. We have determined that this process is the one that reaches the global minimum in the most efficient way.

We finish the estimation with fminsearchOS, a more flexible implementation of Matlab's fminsearch, which can be found on Matlab's FileExchange platform. There are no bounds on the parameters when we run FMINSEARCHOS. FMINSEARCHOS only takes few hours to converge.

Overall, for the estimation of Column 3 in Table 4 of the paper we end up with 589,083 iterations. The minimization process takes one week of time.

Plots Around the Minimum Following a Referee's suggestion, we show plots of the objective function around the minimum as we vary each parameter, one by one. The plots are in Figure A1. Although this is not a formal proof of optimality, the plots lend credibility to the outcome of our estimation routine.[7]

Inference The construction of confidence regions follows Chernozhukov, Hong, and Tamer (2007), which involves obtaining critical values via subsampling. In practice, we use 100 subsamples of one-quarter the size of the original dataset and start the subsampling routine

[7]The domain of each plot was chosen so that the minimum can be visually identified. In all cases, if we extend the domain the function value in those regions increases considreably and the plot becomes difficult to visuall inspect.

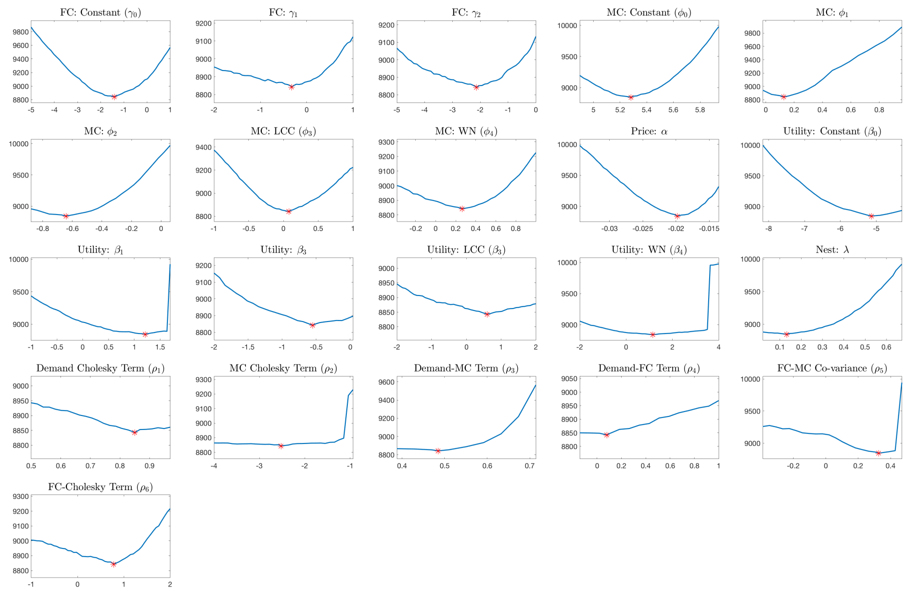


Figure A1: Objective Function in Each Parameter Dimension

from the *argmin* we found in the initial estimation. We compute the confidence region for the point using the procedure outlined in the On-line Appendix of CT, page 5.

Note on normalizing the objective function $Q(\theta)$: after finding a minimizer of the objective function, and so obtaining a value $\eta_n = \min_\theta Q_n(\theta)$ we normalize the *sample* objective function and use $Q_n^*(\theta) = Q_n(\theta) - \eta_n$ to construct confidence regions (as suggested in CHT). This is similar to constructing constructing χ^2 confidence regions in GMM and plays a role in that it guarantees that in finite samples, the confidence intervals are non-empty. Under a well specified model, this normalization plays no role asymptotically. If the model is misspecified (i.e., $\min Q(\theta) > 0$), then this normalization will guarantee that we are constructing confidence regions for the $arg\min$ of $Q(\theta)$.

(v) Details of Counterfactuals

We predict the effects of a AA-US merger for four different assumptions about the new merged firms:

1. the surviving firm, AA, takes the best observed and unobserved characteristics from the pre-merger AA and US,

2. the surviving firm inherits the best observed characteristics, but we take AA's unobservable,

3. the surviving firm inherits the best observed characteristics, but we re-draw errors for the firm,
4. the surviving firm takes the mean values of the observed and unobserved characteristics from the pre-merger AA and US, and

To construct confidence intervals for each of these counterfactual scenarios, we draw 100 parameter vectors from the confidence set and evaluate the counterfactual equilibrium market outcomes for each parameter vector. For a specific table, for example Table 8, starting from all of the counterfactual simulations we condition on a particular pre-merger market structure. For a particular parameter vector, we compute the upper and lower bounds for the number of times we observe the market structure changing to each possible market structures, post merger. So for Table 8, we took all the simulations that (for a particular parameter vector) predicted AA/US Duopoly before the merger and then computed then counted the number of times we observed an AA monopoly post-merger. Then, to get the 95% confidence sets, we take the 2.5 percentile and 97.5 percentile of the probability of observing each market structure, across the 100 parameter vectors we drew from the original confidence set. We follow the same procedure for the prices – always conditioning on the pre-merger market structure.

(vi) Timing and Acknowledgments

At the beginning of our computational work, in 2014, we ran testing for the code on multiple systems, including the XSEDE resources Gordon and Trestles at the San Diego Supercomputing Center. Performance and scaling tests on Gordon indicated at most 32 workers (cores) provided the shortest execution time before communication overhead to the workers becomes significant. The computationally intense estimation of our models at the time in a relatively short period of time was made feasible because of the use of XSEDE resources.[8] In our experimentation we found that having 200 or 500 simulations did not make a difference in our results, while the time taken was much larger. Thus, the costs of more simulations outweighed the benefits.

More recently, We have used two other resources to implement the optimization routine, the HPC system at the University of Virginia known as Rivanna, and the HPC system at Penn State University known as ACI. Rivanna is a 4800-core, high-speed interconnect cluster, with 1.4 PBs of storage available in a fast Lustre filesystem. ACI is a 23,000 core high-speed cluster with 20 PBs of storage and 640 teraflops of total peak performance.

We gratefully acknowledge the use of both the XSEDE resources and those at the University of Virginia and Penn State University.

[8]John Towns, Timothy Cockerill, Maytal Dahan, Ian Foster, Kelly Gaither, Andrew Grimshaw, Victor Hazlewood, Scott Lathrop, Dave Lifka, Gregory D. Peterson, Ralph Roskies, J. Ray Scott, Nancy Wilkens-Diehr, "XSEDE: Accelerating Scientific Discovery", Computing in Science & Engineering, vol.16, no. 5, pp. 62-74, Sept.-Oct. 2014.

C Data Construction

The main data are from the domestic *Origin and Destination Survey (DB1B)*, the *Form 41 Traffic T-100 Domestic Segment (U.S. Carriers)*, and the *Aviation Support Tables : Carrier Decode*, all available from the Department of Transportation's National Transportation Library. We also use the US Census for the demographic data, specifically to get the total population in each Metropolitan Statistical Area. The *Origin and Destination Survey (DB1B)* is a 10 percent sample of airline tickets from reporting carriers. The dataset includes information on the origin, destination, and other itinerary details of passengers transported, most importantly the fare. The *Form 41 Traffic T-100 Domestic Segment (U.S. Carriers)* contains domestic non-stop segment data reported by US carriers, including carrier, origin, destination of the trip. The dataset *Aviation Support Tables : Carrier Decode* is used to clean the information on carriers, more specifically to determine which carriers exit the industry over time, and which one merge, or are owned by another carrier.

We define a market as a unidirectional trip between two airports, irrespective of intermediate transfer points. For example, we will assume that the nonstop service between Chicago O'Hare (ORD) and New York La Guardia (LGA) is in the same market as the connecting service through Cleveland (CLE) from ORD to LGA. The market ORDLGA is a different market from LGAORD.

We follow Borenstein (1989) and assume that flights to different airports in the same metropolitan area are in separate markets. To select the markets, we merge this dataset with demographic information on population from the U.S. Census Bureau for all the Metropolitan Statistical Areas of the United States. We then construct a ranking of airports by the MSA's market size. Our final dataset includes a sample of markets between the top 100 Metropolitan Statistical Areas, ranked by the population size in 2012. We exclude the Youngstown-Warren Regional Airport, Toledo Express Airport, St. Pete-Clearwater International Airport, Muskegon County Airport, and Lansing Capital Region International Airport because there are too few markets between these airports and the remaining airports.

Then, we proceed to further clean the data as follows. We drop: 1) Tickets with more than 6 coupons overall, or more than 3 coupons in either direction if a round-trip ticket; 2) Tickets involving US-nonreporting carrier flying within North America (small airlines serving big airlines) and foreign carrier flying between two US points; 3) Tickets that are part of international travel; 4) Tickets involving non-contiguous domestic travel (Hawaii, Alaska, and Territories) as these flights are subsidized by the US mail service; 5) Tickets whose fare credibility is questioned by the DOT or for which the bulk fare indicator was equal to 1 ; 6) Tickets that are neither one-way nor round-trip travel; 7) Tickets including travel on more than one airline on a directional trip (known as interline tickets), here identified by whether there was a change in the ticket carrier for the ticket.

Next, we follow the approach in Borenstein (1989) and Ciliberto and Williams (2014) and consider a round-trip ticket as two directional trips on the market, and the fare paid on each directional trip is equal to half of the round-trip fare. A one-way ticket is one directional trip.

Moreover, as in Berry and Jia (2010) and Ciliberto and Williams (2014), tickets sold under a code-share agreement (for example, a ticket sold by USAir on a United operated flight) are allocated to the airlines that sold the tickets (so, in the example, to USAir). This is consistent with the notion that the ticketing carrier has access to the "metal" (the seats) of the operating carrier. Notice that this implies that there can be observations where the airline does not have any nonstop routes out of an airport, but the airline can sell tickets for flights out of that airport.

We then drop: 1) Tickets with fares less than 20 dollars; 2) Tickets in the top and bottom one percentiles of the year-quarter fare distribution, and tickets for which the fare per mile (the yield) was in the top and bottom one percentiles of the year-quarter yield distribution.

We then aggregate the ticket data by ticketing carrier and thus the unit of observation is market-carrier-year-quarter specific.

Next, we drop markets whose distance is less than 150 miles. We also drop airlines that served fewer than 90 passengers in a quarter. Finally, we determine the markets that are not served by any airline, but that could be potentially served by one. These are the markets that were served at least 80 percent of all quarters between the first quarter in 1994 and the first quarter in 2017.

The airlines in the initial dataset are: American, Alaska, JetBlue, Delta, Frontier, Allegiant, Spirit, Sun Country, United, USAir, Virgin, Southwest. By the second quarter of 2012, Southwest had completed the acquisition of AirTran, although the two carriers were still issuing tickets with different code (FL vs WN). As in Ciliberto and Tamer (2009), we deal with how to treat regional airlines that operate through code-sharing with national airlines as follows. We assume that the decision to serve a spoke is made by the regional carrier, which then signs code-share agreements with the national airlines. As long as the regional airline is independently owned and issues tickets, we treat it separately from the national airline.

The low cost type is composed of: Alaska, JetBlue, Frontier, Allegiant, Spirit, Sun Country, Virgin. We re-elaborate their data as follows. The LCC's number of passengers is the sum of the passengers over all the LCCs that serve a market. The LCC's price is the passenger weighted mean of the prices charged by all the LCC airlines in a market. For the explanatory variables we take the maximum value among the low cost carriers serving a market of the variables *Origin Presence*, *Destination Presence*, *Nonstop Network Origin*, *Nonstop Network Destination*. We also take the maximum of the categorical variables that indicate whether a firm is a potential entrant in a market.

After this preliminary cleaning, we compute the 95 percentile of the mean prices and yield per mile, and we drop markets where prices and yields above these values were observed.

In order to compute the confidence intervals as in Chernozhukov, Hong, and Tamer (2007) we discretize the exogenous variables. The discretization is done as follows. First, we standardize the continuous variables. Then, we construct intervals where the thresholds are given by -1, -0.5, 0, 0.5, 1, as well as integers such as -2, 2, -3, 3. The discretization affects the variables in both the solving the model as well as the values of the instruments. When also estimate the exogenous-entry GMM specification with these discretized variables.

D Robustness Analysis

This Section investigates how demand estimates change with changes in the modeling of the demand and in the nature of the exogenous variation that identifies the demand coefficients. The results are before the discretization of the variables.

Column 1 of Table A2 presents the baseline results from running a standard OLS regression. The price coefficient is estimated equal to -0.004, and it implies a median elasticity of -0.902, which is inconsistent with a model of profit maximization. There are 15,100 observations out of 22,445 for which the elasticity is larger than -1.

Column 2 of Table A2 presents the baseline results from running a standard two stage least squares nested logit regression, when we use *Nonstop Destination* and *Nonstop Origin* as instrumental variables. We use both the values of the firm associated with the observation as well as the values of the potential competitors. This is analogous to the identification strategy in Bresnahan (1987). The coefficient estimate of the price is now equal to -0.012, and the median elasticity is -3.005.

Table A2: *Parameter Estimates with Exogenous Market Structure*

	OLS Logit	Simple Logit IV	Nested Logit IV	Nested Logit IV	Nested Logit IV
Demand					
Price	-0.004	-0.012	-0.028	-0.020	-0.026
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
σ	–	–	0.529	0.361	0.420
	–	–	(0.016)	(0.012)	(0.016)
Distance	-0.241	0.161	0.923	0.518	0.790
	(0.016)	(0.027)	(0.038)	(0.028)	(0.037)
Origin Presence	0.007	0.006	0.004	0.022	0.005
	(0.000)	(0.000)	(0.000)	(0.048)	(0.000)
LCC	1.004	0.458	-0.505	0.082	-0.338
	(0.037)	(0.049)	(0.062)	(0.041)	(0.060)
WN	1.201	0.957	0.062	0.445	0.232
	(0.022)	(0.027)	(0.040)	(0.031)	(0.039)
Constant	-8.148	-6.338	-1.936	-4.066	-2.736
	(0.046)	(0.108)	(0.179)	(0.127)	(0.175)
Elasticities and Percentage Contribution Margins					
Median Elasticity	-0.901	-3.005	-11.975	-6.578	-9.1220
Elasticities $>= -1$	15,100	7	0	0	0
Nonstop Destination IV	No	Yes	Yes	Yes	Yes
Nonstop Origin IV	No	Yes	Yes	Yes	No
Potential Entrants IV	No	No	No	Yes	No

Column 3 presents the results when we estimate a nested logit as in Berry (1994) using the same instrumental variables that we used in Column 2. We find the coefficient of price equal to -0.028, the coefficient of the nesting parameter equal to 0.528, and the corresponding median elasticity equal to -11.975.

Column 4 presents the results when we include the information on the potential entrants as instrumental variables. In practice, we add six variables as instrumental variables, one for each of the six firms (AA, DL, UA, LCC, WN, US). The coefficient estimate of price is now -0.020, and the nesting parameter is estimated equal to 0.361. These values are very similar to those in our GMM estimates in Table 4 of the paper. The corresponding median elasticity is equal to -6.578.

Finally, Column 5 of Table A2 shows the results if we maintain that only *Nonstop Destination* can be used as instrumental variables. This is a key mantained assumption in the identification strategy in Berry and Jia (2010). We find now that the price coefficient is estimated equal to -0.026 and the nesting parameter is 0.420. The corresponding median elasticity is -9.122.

Table A3: *Parameter Estimates, Alternative Specifications*

	Main Specification (Table 4 in paper)	Fixed Variance	Segment Effects
Demand			
Price	[-1.992, -1.956]	[-1.650, -1.633]	[-1.360, -1.304]
λ	[0.116, 0.144]	[0.226, 0.238]	[0.157, 0.189]
Distance	[1.085, 1.273]	[0.079, 0.148]	[0.243, 0.300]
Origin Presence	[-0.613, -0.426]	[-0.459, -0.411]	[-0.609, -0.505]
LCC	[0.353, 0.853]	[-1.558, -1.166]	[-1.053, -0.787]
WN	[0.904, 1.409]	[0.299, 0.486]	[-0.351, -0.210]
Constant	[-5.191, -4.878]	[-5.638, -5.560]	[-6.473, -5.929]
Marginal Cost			
Distance	[0.111, 0.158]	[-0.087, -0.071]	[-0.036, -0.029]
Origin Presence	[-0.649, -0.626]	[-0.701, -0.692]	[-0.641, -0.600]
Cons LCC	[-0.002, 0.108]	[-0.426, -0.371]	[-0.365, -0.286]
Cons WN	[0.203, 0.330]	[0.215, 0.281]	[-0.120, -0.073]
Constant	[5.267, 5.2901]	[5.261, 5.277]	[5.792, 5.804]
Fixed Cost			
Constant	[-1.657, -1.288]	[-1.353, -1.291]	[-2.606, -1.992]
Nonstop Origin	[-0.452, -0.264]	[-0.393, -0.323]	[-0.143, -0.041]
Nonstop Dest.	[-2.260, -1.885]	[-2.497, -2.434]	[-2.032, -1.857]
Variance-Covariance*			
Variance of Demand	[5.142, 5.826]	[3.939, 4.123]	[4.494,4.766]
Variance Marg. Cost	[0.334,0.373]	[0.299,0.312]	[0.322,0.338]
Variance of Fixed Cost	[3.721 5.5332]	0.500*	1/3
Demand-MC Covariance	[-0.099,0.243]	[0.429,0.467]	[0.541,579]
Demand-FC Covariance	[0.631,0.786]	[-0.246,-0.202]	[-0.598,-0.537]
MC-FC Covariance	[1.119,1.215]	[0.246,0.255]	[0.164,0.173]
Objective Function Value	0.885	0.890	0.886

Price coefficient multiplied by 100. * denotes a value that is fixed in estiamtion. The function value has been divided by 10000.

Overall, Table A2 shows that the parameter estimates of the price coefficients are stable across Columns 3-5, and show that the information on the potential entrants, as well as the inclusion of *Nonstop Origin* as instrumental variables delivers estimates of the median elasticity that are closer to previous work. More specifically, Berry and Jia (2010) use data from 1996 to 2006 and estimate it between -2 and -3 in 2006 and trending upward from 1999. Ciliberto and Williams (2014) use data from 2006 to 2008 and estimate the aggregate price elasticity to be equal to -4.320 in their model that does not allow for collusive behavior. Berry and Jia (2010) and, later on, Ciliberto and Williams (2014), use a two-type model of demand, where they distinguish between two types, a coach type whose elasticity both papers estimate to be between -6 and -6.5; and a business type, whose elasticity of demand both papers estimate to be around -0.5. The estimates of the aggregate price elasticity differs in Berry and Jia (2010) and Ciliberto and Williams (2014) because they estimate different fractions of coach and business travelers. Berry and Jia (2010) estimate the share of business passengers between 41 and 49 percent. Ciliberto and Williams estimate the share of business passengers to be 34 percent. We infer that the average elasticity of demand doubled between the period analysed by Berry and Jia and the one analysed by Ciliberto and Williams, because of an increase in the share of economy passengers. Our dataset is from 2012, well after the ones used by Berry and Jia (2010) and by Ciliberto and Williams (2014), and therefore the increase in price elasticity is consistent with an increase in the share of economy passengers.

In Table A3, we display estimation results from two alternative specifications of our model, to show that freeing up the variance of the fixed cost firm unobservable is possible in our context, in contrast to the CT (and any other discrete choice model, where the variance needs to be normalized). In Column 1, we display our main results from Table 4 in the main text, simply replicating them. In Column 2, we display results from a specification where we fix the variance of fixed costs to 0.5 and the variance of the market specific error to 0.5 (so that the sum is still equal to 1). In Column 3, we display the results from a specification where we include an additional random effect at the bidirectional level. In practice, we maintain that there are three errors in the entry equation: one that is firm specific; one that is uni-directional specific (thus, LGAJFK has a different market specific error than JFKLGA), and one that is bi-directional specific (thus, it is the same in the market LGAJFK and in the market JFKLGA). Each one of these three errors is maintained to have a variance equal to 1/3. We observe that the results in Columns 2 and 3 are quite close to each other, with the most important difference being that the price and nesting parameters are even smaller in Column 3 than in Column 2. The comparison of the results in Column 1 vs 2 and 3 shows that freeing up the variance of the fixed costs leads to a lower function value of the objective functions, as one would expect since there is one more free parameter. In general, the magnitudes of the parameters are close to each other in the three specifications, which assures that the model is robust to letting the variance of the firm fixed cost free.

E Numerical Exercise

We run a series of numerical exercises to show that GMM estimates of markups are biased if the true model has endogenous entry and to show that our estimation methodology works well when we know the true parameters.

First, we present a slightly simplified version of our model. The simplifications include fewer demand and cost variables. In practice, the model includes the minimum number of parameters to make it comparable to our empirical analysis. The model is represented by the following system of conditions:

$$
\begin{aligned}
&Demand: && ln(s_{jm}) = \alpha p_{jm} + c_1 + X_{jm} b_1 + \lambda ln(s_{jgm}) + \xi_{jm} && \text{(A.7)}\\
&Supply: && ln(c_{jm}) = c_2 + b_2 X_{jm} + \eta_{jm} && \text{(A.8)}\\
&Entry: && y_j = 1 \Leftrightarrow \pi_j \equiv (p_{jm} - c_{jm}) M_m s_{jm} - exp(c_3 + b_4 Z + \nu) FC_{jm} \geq 0, && \text{(A.9)}\\
& && && \text{(A.10)}
\end{aligned}
$$

where the epxression for demand and marginal costs are the following,

$$
s_{jm} = \frac{exp(\alpha p_{jm} + X_{jm}\beta + \xi_{jm})}{1 + \sum_k exp(\alpha p_{jm} + X_{km}\beta + \xi_{km})} \tag{A.11}
$$

$$
c_{jm} = p_{jm} - \frac{\alpha(1-\rho)}{1 - \rho s_{jgm} - (1-\rho) s_{jm}}. \tag{A.12}
$$

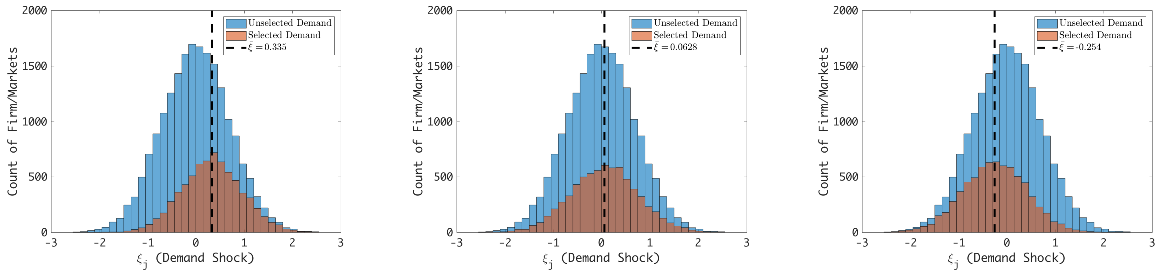
We assume the following variance covariance matrix for a particular airline:

$$
\Sigma_{jm} = \begin{bmatrix} \sigma_\xi^2 & \sigma_{\eta\nu} & \sigma_{\xi\nu} \\ \sigma_{\eta\nu} & \sigma_\eta^2 & \sigma_{\eta\nu} \\ \sigma_{\xi\nu} & \sigma_{\eta\nu} & \sigma_\nu^2 \end{bmatrix}.
$$

As in the main text, we assume that the correlation is only among the unobservables within a firm, and not between the unobservables of the K_m firms. This specification also restricts the correlations to be the same for each firm and clearly reduces the parameters to be estimated. However, the specification is rich compared to existing methods.

We generate covariates from a standard normal, one covariate for demand and one for fixed costs. We also randomly generate market sizes for each market.

First, we can describe the role of selection in the model by displaying the distribution of errors pre-selection and post-selection for various values associated with the covariance matrix of the unobserved terms. In Figure A3 we display three graphs of histograms of errors. In each graph, the larger histogram represents the pre-selected distributions of demand errors for all of the potential firms in all of the simulated markets. This distribution is drawn from an underlying joint normal with a mean of zero and the covariance matrix parameters displayed in Column 1 of Table A4, except that we vary the correlation between demand and fixed costs. In each graph, the smaller histogram represents those demand errors from firms in markets that the model predicts to enter, or in other words, the selected errors. It



(a) $cov(\xi, \nu) = -0.4$ (b) $cov(\xi, \nu) = 0$ (c) $cov(\xi, \nu) = 0.4$

Figure A3: Distribution of Demand Errors (ξ) For Different Covariances

is clear that the distribution of selected demand errors changes as the covariance between demand and fixed costs changes. In the model, negative correlation between demand and fixed cost shocks implies positively selected firms, which is intuitive and can be seen in the first panel of Figure A3. The middle panel shows that without any correlation, the model induces only a slightly shifted distribution of demand errors.[9] When demand and fixed cost shocks are positively correlated, the distribution of selected demand shocks is shifted to the left. We would expect the corresponding bias in elasticity estiamtes to vary based on the values of the covariance matrix as well.

Monte Carlo Simulation: Bias in "Standard" Model

We document the bias from estimating a standard model that does not account for selection. To do this, we solve the model 1000 times for different random draws of the covariates, errors, price parameter (α), nest parameter (λ), covariance matrix, and sets of potential entrants. For each of the 1000 generated data sets, we estimate demand using the method suggested in Berry (1994) and compute the implied markups. In Figure A3, we graphically compare the implied markups from GMM to the true markups used to generate the 500 different draws of data, varying the price sensitivity across datasets as well. It is clear that the estimates are systematically different than the true values.

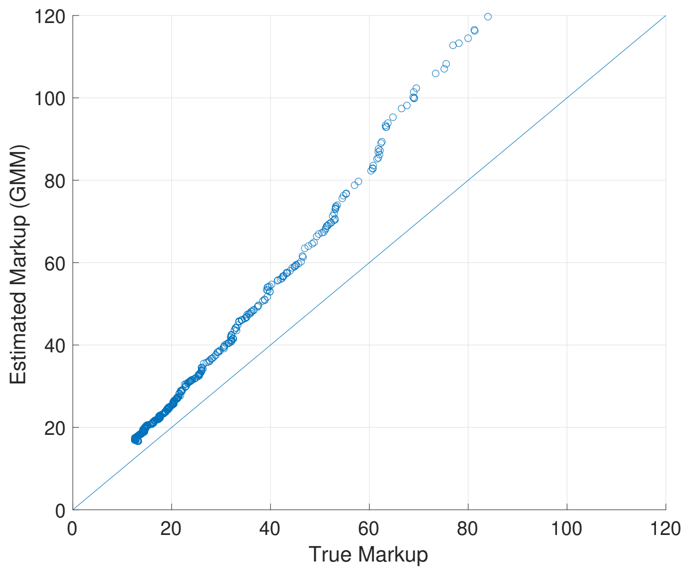
Model Estimation with Endogenous Entry

Next, we estimate the model using simulated data, employing the methodology we present in Sections 2 and 3 of the main text. The true parameters are in Column 1 of Table A4. In Column 2 we present the estimates using GMM, not accounting for selection a la Berry (1994). In the Column 3, we present the 95% confidence intervals using our methodology.

Our methodology does quite well. Most of the true parameters lie within their associated confidence intervals using our methodology, and in many cases the confidence intervals are

[9] We computed this numerical exercise for many different parameters and have found positive and negative selection in the case where $cov(\xi, \eta) = 0$. We chose to display this particular case because these are the parameters we use for the estimation exercise below.

Figure A3: GMM Bias in Markups Across Different Parameter Values



Note: Plot of true markups versus estimated markups using GMM that does not account for endogenous selection/entry. Each point represents a different draw of data, errors, price parameter, and nesting parameter.

tight.[10] In particular, our methodology does a much better job at estimating the price coefficient than GMM.

It is not surprising that the price parameter, in particular, suffers from bias in the GMM estimation, because it links all three model conditions through its role in determining markups (and, thus, the entry profit threshold condition as well).

Table A4: Parameter Estimates Using Simulated Data

	True	GMM	Endogenous Entry
Demand			
Price	-0.02	-0.034 (0.005)	[-0.029, -0.021]
Constant	-3	0.077 (0.995)	[-3.124, -1.842]
X	0.5	1.284 (0.278)	[-0.819, 0.903]
Nest (λ)	0.30	0.307 (0.166)	[0.273, 0.360]
Marginal Cost			
Constant	5	5.067 (0.003)	[4.670, 5.272]
X	0.5	0.375 (0.003)	[0.107, 0.610]
Fixed Cost			
Constant	3	–	[1.071, 3.318]
Z	-0.5	–	[-0.594, -0.231]
Variance-Covariance			
Marg. Cost Variance	0.10	0.074	[0.181, 0.263]
Demand Variance	2	3.254	[1.345, 3.498]
Demand-FC Covariance	-0.10	–	[-0.027, 0.147]
Demand-MC Covariance	0.20	–	[0.551, 1.200]
MC-FC Correlation	0.10	0.3945	[0.286, 0.607]

Column 1: parameter values used to create simualted data. Column 2: Standard GMM estimation. Column 3: Estimation using the methodology described in Section 2. Standard errors in parentheses in Column 2. Columns 3 contain 95% confidence bounds constructed using the method in Chernozhukov, Hong, and Tamer (2007). The dataset includes 5000 markets with up to four potential entrants. We use 100 draws to simulate the joint distribution of errors. As in the empirical application, we fix the fixed cost variance at 0.5.

[10] The confidence intervals here are larger than in our empirical exercise. One reason is that our real data looks irregular in the sense that it does not look normal like the fake data – in this sense the real data might better satisfy large support conditions. Second, we use more bins to discretize the real data. Third, there is more variation in potential entrants in the real data.

References

Berry, Steven T. 1994. “Estimating discrete-choice models of product differentiation.” *The RAND Journal of Economics* :242–262.

Chernozhukov, Victor, Han Hong, and Elie Tamer. 2007. “Estimation and confidence regions for parameter sets in econometric models1.” *Econometrica* 75 (5):1243–1284.

Ciliberto, Federico and Elie Tamer. 2009. “Market structure and multiple equilibria in airline markets.” *Econometrica* 77 (6):1791–1828.

Tamer, Elie. 2003. “Incomplete simultaneous discrete response model with multiple equilibria.” *The Review of Economic Studies* 70 (1):147–165.