Market Structure and Competition in Airline Markets

Federico Ciliberto

University of Virginia

Charles Murry

Boston College

Elie Tamer

Harvard University

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 postmerger outcomes depend on how we model the characteristics of the merged firm as a function of the premerger firms' characteristics.

I. Introduction

We estimate a simultaneous, static complete information game where economic agents make both discrete and continuous choices. We study

We thank S. Berry, T. Bresnahan, A. Chandra, P. Grieco, J. Panzar, W. Tan, R. Watson, and J. Williams for insightful suggestions. We also thank participants at the SEA (Southern

Electronically published September 23, 2021

Journal of Political Economy, volume 129, number 11, November 2021.

https://doi.org/10.1086/715848

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 (i) accounts for selection of firms into serving a market and, more importantly, (ii) allows 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.

Nonrandom 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 variable 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 mismeasuring 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

Economic Association) Meetings, the 4th Annual CAPCP (Center for the Study of Auctions, Procurements and Competition Policy) Conference at Penn State University in 2009, the 2011 JAE (Journal of Applied Econometrics) Lecture at Yale, and the 2014 DC IO (Industrial Organization) Conference, where early drafts of this paper were presented. Seminar participants at many institutions provided useful comments. We want to especially thank Ed Hall and the University of Virginia Alliance for Computational Science and Engineering for essential advice and guidance in solving many computational issues. We also acknowledge the generous support of computational resources from XSEDE (Extreme Science and Engineering Discovery Environment) through the Campus Champions program (National Science Foundation–XSEDE grant SES150002). Federico Ciliberto thanks the Center for the Study of Industrial Organization at Northwestern University for sponsoring his visit. Research support from the Bankard Fund for Political Economy and Quantitative Collaborative of the College of Arts and Science at the University of Virginia is gratefully acknowledged. We also thank the editor and reviewers for comments that helped sharpen and improve this paper. Data are provided as supplementary material online. This paper was edited by Ali Hortaçsu.

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

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

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 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(s) 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 be viewed as a multiagent version of the classic selection model (Gronau 1974; Heckman 1976, 1979). In this 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." In our setting, 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 make nonnegative profit in each market that it serves; (ii) demand equations derived from a discrete choice model of consumer behavior; and (iii) pricing first-order conditions, which can be formally derived under the postulated firm conduct. We allow for all firm decisions to depend on 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) 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 (Berry 1994; Berry, Levinsohn, and Pakes 1995).

Our innovation on this front is to show how to estimate demand and supply equations in the presence of multiple equilibria in the entry stage by constructing moment inequities from conditional distributions of the residuals. 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. We consider markets between US Metropolitan Statistical Areas (MSAs), which are served by American Airlines (AA), Delta Air Lines (DL), United Airlines (UA), US Airways (US), Southwest (WN), and low-cost carriers (LCCs; e.g., JetBlue). We observe variation in the identity and number of potential entrants across markets.[4] Each firm decides whether to enter and chooses the price in that market. The other endogenous variable is the number of passengers transported by each firm. The identification of the three conditions listed above relies on variation in several exogenous explanatory variables, whose inclusion in the model is supported by a rich and important literature (e.g., Rosse 1970; Panzar 1979; Bresnahan 1989; Schmalensee 1989; Berry 1990; Brueckner and Spiller 1994; Ciliberto and Tamer 2009; Berry and Jia 2010; Ciliberto and Williams 2014).

We begin our empirical analysis by running a standard generalized method of moments (GMM) estimation (see Berry 1994) on the demand and pricing first-order conditions and comparing that with 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 by 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.[5] Next, we use our estimated model to simulate the merger of two airlines in our data: AA and US.[6] Typical merger analysis involves predicting changes in market power and prices given a particular market structure using diversion ratios based on premerger 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 how we model the characteristics of the postmerger firm as a function of the premerger firms' characteristics. We consider four different scenarios. First, we assume that the merged firm takes on the best

[3] We also illustrate our methodology by conducting a numerical exercise; see app. E.

[4] A market is defined as a unidirectional pair of an origin and a destination airport, as done so by 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 app. C for more details.

[5] 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 cost unobservables. We illustrate this dependence in the numerical exercise in app. E.

[6] The two firms merged in 2013 after settling with the DOJ.

characteristics, both observed and unobserved, of the two premerger firms, and call this the "best-case scenario." Then we simulate two subcases, one in which the merged firm takes the best observable characteristics between the two premerger firms and the average of AA's and US's premerger unobservables and another where we draw a new unobservable for the new merged firm. Last, we consider a case where the surviving firm inherits the average observed and unobserved characteristics between the two premerged firms, or what we call the "average-case scenario."

We find that under all four scenarios, there is substantial postmerger entry and exit among the surviving airlines, especially for the surviving merged airline, AA. 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 postmerger 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 premerger market structures. For example, we find that the merged airline would enter previously unserved markets with a likelihood of around 48% in the best-case scenario and that prices would increase by roughly 14% in markets previously only served by an AA and US duopoly. In contrast, when we assume that the postmerger airline takes the average characteristics from AA and US (the average-case scenario), we find that the merged airline would enter previously unserved markets with a likelihood of around 9% and that prices would rise by roughly 5% in markets previously only served by an AA and US duopoly. Clearly, assumptions about merger efficiencies matter—not just for pricing pressure but also for postmerger entry/exit. We systematically document these types of effects across many premerger market structures.

Finally, we investigate the effects of the merger in markets originating or ending at Reagan National Airport, which were of concern for antitrust authorities because both of the merging parties had a very strong incumbent presence. We find that prices would increase, though in different degrees that depend on the scenario under consideration. We also find that low-cost carriers are not likely to replace the previous US Airways routes, which was a major concern for the Department of Justice (DOJ) and resulted in slot divestitures by the merging party.

There is other important work related to estimating static models of competition while allowing for market structure to be endogenous. Reiss and Spiller (1989) estimate a monopoly model of airline competition and entry. In contrast, we allow for multiple firms to choose whether 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. Ellickson and Misra (2012) use a two-step method to estimate a static discrete game of incomplete information and correct for an outcome equation, in their case, revenues. Berry (1999), Draganska, Mazzeo, and Seim (2009), Ho (2009), and Pakes et al. (2015) assume that firms self-select 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 into markets that better match their observable and unobservable characteristics. There are three recent papers that are closely related to ours. First, Eizenberg (2014) estimates a model of entry and competition in the personal computer industry. Estimation relies on a timing assumption (motivated by Pakes et al. 2015) 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.[7] Similar timing assumptions are made by other papers as well, such as those of Lee (2013), Sweeting (2013), and Jeziorksi (2014a, 2014b) in dynamic empirical games and Fan (2013) and Fan and Yang (2020) in static games.[8] Second, Fan (2013) does allow for arbitrary correlation between unobservables, but her setting is one where firms choose a continuous product characteristic. Third, Li et al. (2021) estimate a model of service selection (nonstop vs. connecting) and price competition in airline markets but consider only sequential-move equilibria. In addition, Li et al. (2021) 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 II 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 III introduces the economic model. Section IV introduces the airline data, providing some preliminary evidence of self-selection of airlines into markets. Section V shows the estimation results, section VI presents results and discussion of the merger exercise, and section VII concludes.

[7] If we are willing to make this timing assumption, there would not be a selection on unobservables, because the firm would observe only 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 a timing assumption would seem reasonable. In the airline industry, firms can enter and exit the 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 nonzero, then we could conclude that the timing assumption would be less acceptable.

[8] There is also an empirical literature on auctions (Li and Zheng 2009; Roberts and Sweeting 2013; Gentry and Li 2014; Li and Zhang 2015) that, in static models, has relaxed 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.

II. 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- versus multiagent model. Section III analyzes a full model of entry and pricing.

Consider the following system of inequality conditions:

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

where, for $j \in \{1, 2\}$, $y_j = 1$ if firm j decides to enter a market and 0 otherwise. 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 (v_1, v_2, ξ_1, ξ_2) are unobserved and are independent of $(\mathbf{Z}, \mathbf{X})$, while the variables (V_1, V_2) are endogenous (e.g., prices or product characteristics).[9]

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 right-hand side (the Vs). 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,

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

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 a key source of selection bias, since correlation in the observables can be controlled for.

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

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, there could be an unobservable to the researcher variable that leads to a luxury car being more attractive to consumers, and at the same time this variable may be the reason the car requires more upfront investment and greater costs to produce a single 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 III, the unobservables that determine outcomes also enter directly into the selection equation (see eq. [7] in sec. III). 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 selection setup and its generalization to two firms above because the structural error 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 nonstandard 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, further difficulties arise because we want to allow for the possibility that the *V*s are also choice variables (or variables determined in equilibrium, e.g., prices).

The data we observe are $(S_1y_1, V_1y_1, y_1, S_2y_2, V_2y_2, y_2, \mathbf{X}, \mathbf{Z})$, 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 $(S_1y_1, V_1y_1, y_1, S_2y_2, V_2y_2, y_2, \mathbf{X}, \mathbf{Z})$; 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

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

for all t_1.[10] 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.[11]

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 or other semiparametric approaches, 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 $(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{U}}_{(1,0)})$ be the set of ξ_1 that are less than t_1 when the unobservables (v_1, v_2) belong to the set $A^{\mathrm{U}}_{(1,0)}$. The set $A^{\mathrm{U}}_{(1,0)}$ is the set where (1,0) is the unique (pure strategy) Nash equilibrium outcome of the model.

Next, let $(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}, d_{(1,0)} = 1)$ be the set of ξ_1 that are less than t_1 when the unobservables (ν_1, ν_2) belong to the set $A^{\mathrm{M}}_{(1,0)}$. The set $A^{\mathrm{M}}_{(1,0)}$ 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 method of moments. By the law of total probability, we have (suppressing the conditioning on $(\mathbf{X}, \mathbf{Z})$)

$$\begin{aligned} P(\xi_1 \leq t_1; y_1 = 1; y_2 = 0) &= P\big(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{U}}_{(1,0)}\big) \\ &+ P(d_{1,0} = 1 \mid \xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}) P\big(\xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}\big). \end{aligned} \tag{3}$$

The probability $P(d_{1,0} = 1 \mid \xi_1 \leq t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)})$ 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

[10] Here we use the cumulative distribution function (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 $t_0 \leq t_1$. Working with histogram-like or cell probabilities can have some computational advantages.

[11] In the case where we have no endogeneity, e.g., (α 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 $P(\xi_1 \leq -X_1\beta;\ \ y_1 = 1,\ \ y_2 = 0 | \mathbf{X}, \mathbf{Z})$.

$$\begin{aligned}P\left(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{U}}_{(1,0)}\right) &\le P(\xi_1 \le t_1; y_1 = 1; y_2 = 0)\\ &= P(S_1 + \alpha_1 V_1 - X_1\beta \le t_1; y_1 = 1; y_2 = 0)\\ &\le P\left(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{U}}_{(1,0)}\right)\\ &\quad + P\left(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}\right).\end{aligned}$$

The middle part,

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

can be consistently estimated from the data given a value for $(\alpha_1, \beta t_1)$. The left-hand and right-hand sides contain the following two probabilities:

$$P\left(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{U}}_{(1,0)}\right), P\left(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}\right).$$

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 (ξ_1, ξ_2, v_1, v_2) has a known distribution up to a finite dimensional parameter (we assume normal) and the fact that the sets $A^{\mathrm{M}}_{(1,0)}$ and $A^{\mathrm{U}}_{(1,0)}$, which depend on regressors and parameters, can be obtained by solving the game given a solution concept (for examples of such sets, see Ciliberto and Tamer 2009). For example, for a given value of the unobservables, observables, and parameter values, we can solve for the equilibria of the game that 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 on the left-hand side (and right-hand side) 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 $(S_2 - \alpha_2 V_2 - X_2\beta \le t_2, y_1 = 0, y_2 = 1)$ can similarly be calculated. In addition, in the two-player entry game (i.e., δs are negative) above with pure strategies, the events (1,1) and (0,0) are uniquely determined, and so

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

is equal to

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

which can be easily calculated (e.g., via simulation). We also have

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

For the two-equation selection model we describe, the statistical moment inequality conditions implied by the model at the true parameters are

$$\begin{aligned}
m^{l}_{(1,0)}(t_1, \mathbf{Z}; \Sigma) &\le E(1[S_1 - \alpha_1 V_1 - X_1\beta \le t_1; y_1 = 1; y_2 = 0]) \\
&\le m^{u}_{(1,0)}(t_1, \mathbf{Z}; \Sigma), \\
m^{l}_{(0,1)}(t_2, \mathbf{Z}; \Sigma) &\le E(1[S_2 - \alpha_2 V_2 - X_2\beta \le t_2; y_1 = 0; y_2 = 1]) \\
&\le m^{u}_{(0,1)}(t_2, \mathbf{Z}; \Sigma), \\
E(1[S_1 - \alpha_1 V_1 - X_1\beta \le t_1; S_2 - \alpha_2 V_2 - X_2\beta &\le t_2; y_1 = 1; y_2 = 1]) \\
&= m_{(1,1)}(t_1, t_2, \mathbf{Z}; \Sigma), \\
E(1[y_1 = 0; y_2 = 0]) &= m_{(0,0)}(\mathbf{Z}; \Sigma),
\end{aligned}$$

where

$$\begin{aligned}
m^{l}_{(1,0)}(t_1, \mathbf{Z}; \Sigma) &= P\big(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{U}_{(1,0)}\big), \\
m^{u}_{(1,0)}(t_1, \mathbf{Z}; \Sigma) &= m^{l}_{(1,0)}(t_1, \mathbf{Z}; \Sigma) + P\big(\xi_1 \le t_1; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(1,0)}\big), \\
m^{l}_{(0,1)}(t_2, \mathbf{Z}; \Sigma) &= P\big(\xi_2 \le t_2; (\nu_2, \nu_2) \in A^{\mathrm{U}}_{(0,1)}\big), \\
m^{u}_{(0,1)}(t_2, \mathbf{Z}; \Sigma) &= m^{l}_{(0,1)}(t_2, \mathbf{Z}; \Sigma) + P\big(\xi_2 \le t_2; (\nu_1, \nu_2) \in A^{\mathrm{M}}_{(0,1)}\big), \\
m_{(1,1)}(t_1, t_2, \mathbf{Z}; \Sigma) &= P(\xi_1 \le t_1, \xi_2 \le t_2, \nu_1 \ge -\delta_2 - \gamma Z_1, \nu_2 \ge -\delta_1 - \gamma Z_2), \\
m_{(0,0)}(\mathbf{Z}; \Sigma) &= P(\nu_1 \le -\gamma Z_1, \nu_2 \le -\gamma Z_2).
\end{aligned}$$

Hence, the above can be written as

$$E[\mathbf{G}(\theta, S_1 y_1, S_2 y_2, V_1 y_1, V_2 y_2, y_1, y_2; t_1, t_2) | \mathbf{Z}, X] \le 0, \tag{4}$$

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

The last moment, $m_{(0,0)}(\mathbf{Z}; \Sigma)$, is the Ciliberto and Tamer (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 $t_1, t_2 \to \infty$, our moments collapse to the CT moments. The superscripts l and u stand for lower and upper bounds, respectively.

We use standard moment inequality methods to conduct inference on the identified parameters. In particular, we note the following.

RESULT 1. 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 independent and identically distributed 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,

$$Q(\theta) = 0 = \int W(\mathbf{X}, \mathbf{Z}) \, \| \mathbf{G}(\theta, S_1 y_1, S_2 y_2, V_1 y_1, V_2 y_2, y_1, y_2) | \mathbf{Z}, X] \|_+ \, dF_{\mathbf{X},\mathbf{Z}}, \quad (5)$$

for a strictly positive weight function $W(\mathbf{X}, \mathbf{Z})$.[12]

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.[13]

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

III. A Model of Entry and Price Competition

A. *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

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

where

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

is the share of firm 1, which depends on whether firm 2 is in the market, M is the market size, $c(W_1, \eta_1)$ is the constant marginal cost for firm 1, $F(Z_1, v_1)$ 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.

[12] See app. A for more details. See Ciliberto and Tamer (2009) for an analogous result and the proof therein. Note here that if θ is partially identified, the objective function yields an outer set on the identified set. Sharp sets, though easy to define, are harder to compute in this model.

[13] We discuss the selection of the t's in app. B.

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

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

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:

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

There are six endogenous variables: p_1, p_2, s_1, s_2, y_1, and y_2. The observed exogenous variables are M, $\mathbf{W} = (W_1, W_2)$, $\mathbf{Z} = (Z_1, Z_2)$, and $\mathbf{X} = (X_1, X_2)$. So, putting these together, we get the following system:

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

The first two inequalities are entry conditions that require that, in equilibrium, a firm that serves a market must be making nonnegative 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 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 with nested logit demand, a result dating back to at least Mizuno (2003) or, more recently, Nocke and Schutz (2018).

Even though the conceptual approach is the same, the inference procedure is computationally more demanding for this model than the one we studied in section II. 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 model (1). The methodology presented in section II 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 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 equations in 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 one another. 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.

B. 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 that allows for all of the inside products to share unobserved heterogeneity. Specifically, indirect utility for consumer i from choosing carrier j is

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

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 by the econometrician.

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

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

where 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

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

denote the set of possible market structures that 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:

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

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

Last, unlike typical demand estimation but similar to the entry literature, we parameterize the joint distribution of unobservables. Following Berry (1992) and Ciliberto and Tamer (2009), 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. Here, 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 Berry, Levinsohn, and Pakes (1995), the marginal cost and demand unobservables include only 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.[14]

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

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

where I_{K_m} is a 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 nonzero among only the unobservables of a firm (within-firm correlation) and not between the unobservables of the K_m firms (between-firm correlation).

C. *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 as

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

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

[14] 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 app. B for details.

The algorithm simulates profits for every $e \in E$ and determines which market structures are in equilibrium. We outline the key steps of the algorithm next. More details, including computational guidance, can be found in appendix B (apps. A–E are available online).

First, we take *ns* pseudo-random independent draws from a $3 \times K$-variate joint standard normal distribution. Let $r = 1, \ldots, ns$ 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 $\Theta^0 = (\alpha^0, \beta^0, \varphi^0, \gamma^0, \Sigma^0)$.

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

   a) Use $\alpha_0, \beta_0, \varphi_0$ 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 $\Pr(\hat{\boldsymbol{\xi}}^{\hat{e}} <= \mathbf{t_D}, \hat{\boldsymbol{\eta}}^{\hat{e}} <= \mathbf{t_S} \mid \mathbf{X}, \mathbf{W}, \mathbf{Z})$, which are joint probability distributions of $\hat{\boldsymbol{\xi}}^{\hat{e}}$, $\hat{\boldsymbol{\eta}}^{\hat{e}}$ 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 predicted by the model to match those of the residuals in step 1 above. Due to multiplicity, the model instead predicts lower and upper bounds using the simulated errors given Θ^0. In particular, we take the following steps.

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

   b) For each potential market structure e of the $2^{|K|} - 1$ possibilities (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 $\bar{\mathbf{s}}_r^e$.[15]

   c) We 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, for example, e^*, then we collect the simulated errors of the firms that are present in that equilibrium, $\boldsymbol{\xi}_r^{e^*}$ and $\boldsymbol{\eta}_r^{e^*}$. In addition, we collect $\mathbf{v}_r^{e^*}$ and include them in $A_{e^*}^{\mathrm{U}}$,

[15] For example, if we look at a monopoly of firm j ($|e| = 1$), then the demand $Q_j(p_{jr}, X_{jr}, \xi_{jr}; \beta)$ is readily computed and is the monopoly price, p_{jr}. Given the parametric assumptions, there is a unique pure-strategy price equilibrium, conditional on the market structure.

which was defined in section II. If there are multiple equilibria, for example, e^{*} and e^{**}, then we collect the simulated errors of the firms that are present in those equilibria, respectively ($\boldsymbol{\xi}_r^{e^*}$, $\boldsymbol{\eta}_r^{e^*}$) and ($\boldsymbol{\xi}_r^{e^{**}}$, $\boldsymbol{\eta}_r^{e^{**}}$).[16] In addition, we collect $\boldsymbol{\nu}_r^{e^*}$ and $\boldsymbol{\nu}_r^{e^{**}}$ and include them, respectively, in $A^{\mathrm{M}}_{e^*}$ and $A^{\mathrm{M}}_{e^{**}}$, which were also defined in section II.[17]

e) We construct $\Pr(\boldsymbol{\xi}_r^e <= \mathbf{t_D}, \boldsymbol{\eta}_r^e <= \mathbf{t_S}; \boldsymbol{\nu} \in A_e^{\mathrm{M}} | \mathbf{X}, \mathbf{W}, \mathbf{Z})$ and $\Pr(\boldsymbol{\xi}_r^e <= \mathbf{t_D}, \boldsymbol{\eta}_r^e <= \mathbf{t_S}; \boldsymbol{\nu} \in A_e^{\mathrm{U}} | \mathbf{X}, \mathbf{W}, \mathbf{Z})$.[18]

3. We construct the distance function (5) in section II. The approach we use for inference follows the implementation of Chernozhukov, Hong, and Tamer (2007) in Ciliberto and Tamer (2009), 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 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.

IV. 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 US Airways). Our methods allow us to examine, within the context of our model, the implications of mergers on equilibrium prices and on market structure. We start with an examination of our data, and then we provide our estimates.

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

[17] See app. B (p. 4) for details, including how we handle situations where no pure-strategy equilibria exist.

[18] These CDFs in this setting with two unobservables for each firm are analogous to the ones with just one unobservable per firm, as described in sec. II. We use the same *t*'s that we used to construct the CDFs of the residuals.

A. *Market and Carrier Definition*

1. Data

We use data from several sources to construct a cross-sectional data set, where the basic unit of observation is an airline in a market (a market carrier). The main data sets are the second quarters of 2012's Airline Origin and Destination Survey and of the T-100 Domestic Segment Data Set's Aviation Support Tables, available from the Department of Transportation's National Transportation Library. We also use the US Census for demographic data.[19]

We define a market as a unidirectional trip between two airports, irrespective of intermediate transfer points.[20] The data set includes the markets between the top-100 US MSAs ranked by their population. We include markets that are not served by any carrier. There are 8,163 unidirectional markets, and each is denoted $m = 1, \dots, M$. There are six carriers in the data set: AA, DL, UA, US, WN, and a low-cost carrier denoted LCC. The LCCs include Alaska, JetBlue, Frontier, Allegiant, Spirit, Sun Country, and Virgin. These firms rarely compete in the same market. The subscript for carriers is j, $j \in \{\text{AA, DL, UA, US, LCC}\}$. There are 22,445 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.[21]

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 enters in 39% of the markets, although it is a potential entrant in 71% of markets. Southwest, on the other hand, is a potential entrant in 64% of markets and enters in 46% 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% having just one potential entrant.

For each firm in a market, there are three endogenous variables: whether 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

[19] See app. C for a detailed discussion on the data cleaning and construction.

[20] We do not model the decision of nonstop versus connecting flights. This is a very difficult problem given the hub-network structure of airline markets. Aguirregabiria and Ho (2012) describe a treatment of hub-spoke networks using a dynamic game framework and Li et al. (2021) a recent treatment in a static framework.

[21] 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. 2016).

TABLE 1
Entry Moments

	Actual Entry	Potential Entry
AA	.39	.71
DL	.73	.95
LCC	.18	.46
UA	.51	.80
US	.49	.87
WN	.46	.64

Note.—Values indicate empirical entry probabilities and percentage of markets as a potential entrant, across airlines.

a market as $y_{jm} = 1$ and if it is not active as $y_{jm} = 0$. For example, we set $y_{\mathrm{LCC}} = 1$ 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 column 6 whether the variable is used in the entry inequality conditions, demand, and marginal cost equations. Like Berry, Carnall, and Spiller (2006), Berry and Jia (2010), and Ciliberto and Williams (2014), we set market size as the geometric mean of the MSA population of the end-point cities.

Panel A 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, and the average number of passengers transported is 2,602.79.

2. Demand

Demand is assumed to be a function 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 it is for a traveler to accumulate points and thus 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 2
Distribution of Potential Entrants Across Markets

	1	2	3	4	5	6
Percentage of markets	1.74	10.61	14.58	16.57	28.13	28.37

Note.—Values indicate the distribution of the fraction of markets by number of potential entrants.

TABLE 3
SUMMARY STATISTICS

	Mean (1)	Standard Deviation (2)	Minimum (3)	Maximum (4)	N (5)	Equation (6)
			A. Endogenous Variables			
Price ($)	242.88	55.25	77.13	364.00	22,445	Entry, utility, marginal cost
Passengers	2,602.79	7,042.02	90	112,120	22,445	Entry, utility, marginal cost
			B. All Markets			
Origin presence	100.36	71.88	0	267	48,978	Utility, marginal cost
Nonstop origin	7.04	13.57	0	127	48,978	Entry
Nonstop destination	7.11	13.61	0	127	48,978	Entry
Distance (000)	1.11	.58	.15	2.72	48,978	Utility, marginal cost
			C. Markets Served			
Origin presence	143.23	57.91	1	267	22,445	Utility, marginal cost
Nonstop origin	10.60	16.76	0	127	22,445	Entry
Nonstop destination	10.67	16.77	0	127	22,445	Entry
Distance (000)	1.17	.56	.20	2.72	22,445	Utility, marginal cost

NOTE.—Summary statistics from a sample described in the text. Observations are of 48,978 potential airline markets from 8,163 distinct markets; 22,445 airline markets are active.

Panels B and C of table 3 report the summary statistics for the exogenous explanatory variables. Panel B 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 1,110 miles (one-way), which is slightly lower than the mean values for active airline markets, 1,170 miles.

3. Fixed and Marginal Costs in the Airline Industry

The total costs of serving an airline market consists of three components: airport, flight, and passenger costs.[22] Airlines must lease gates and hire personnel to enplane and deplane aircraft at the two endpoints. These

[22] We thank John Panzar for helpful discussions on how to model costs in the airline industry. See also Panzar (1979). Other costs are incurred at the aggregate, national level, and we do not estimate them here (advertising expenditures, e.g., are rarely market specific).

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 nonstop routes that an airline serves out of the origin airport) and nonstop destination (the number of nonstop 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 with the airport and flight-specific costs. We maintain that the flight and passenger costs enter the economic opportunity cost of flying a passenger.[23]

Returning to panels B and C 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 the nonstop distance between two airports. We also allow the marginal cost to be different for LCCs and Southwest through the use of dummy variables.

B. 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 from the BLP (Berry, Levinsohn, and Pakes) 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 2019).[24] Our approach is slightly different from the standard one and captures greater variation

[23] 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).

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

in competitive environments because we (i) include every potential entrants' characteristics separately instead of summing or averaging the characteristics in a market and (ii) 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, that is, 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 care only about the network out of the origin airport when they select an airline, for example, because of their ability to accumulate frequent flyer miles over time. In our 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 (i.e., 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, following 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 a scale.

Next, we describe how the correlations between the unobservables are identified.[25] 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 components 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. As a caveat to this discussion, our inference methods do not rely on point identification and hold whether or not the model point identifies the parameters.

V. 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. Last, 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 that cannot be rejected as the truth with at least 95% probability.[26]

[25] 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.

[26] This is the approach that was used by Ciliberto and Tamer (2009). For details, see Ciliberto and Tamer (2009, online supplement) and Chernozhukov, Hong, and Tamer (2007). Notice that there are no multiple equilibria in cols. 1 and 2. In col. 3, multiple equilibria are allowed to occur, but in practice, we did not find that multiple equilibria were as common in our estimates as in Ciliberto and Tamer (2009).

TABLE 4
PARAMETER ESTIMATES

	GMM (1)	Exogenous Entry (2)	Endogenous Entry (3)
		A. Demand	
Price ($100)	[−2.385, −2.185]	[−2.315, −2.282]	[−1.557, −1.488]
λ	[.320, .519]	[.294, .366]	[.186, .206]
Distance	[.308, .364]	[.394, .461]	[.724, .793]
Origin presence	[.291, .339]	[.102, .169]	[1.688, 1.752]
LCC	[−.333, −.143]	[−1.078, −.486]	[.080, .273]
WN	[.216, .335]	[−.077, .206]	[−.029, .128]
Constant	[−2.299, −1.817]	[−2.961, −2.851]	[−4.683, −4.587]
		B. Marginal Cost	
Distance	[.118, .124]	[.112, .130]	[.083, .094]
LCC	[−.313, −.287]	[−.419, −.288]	[−.027, .054]
WN	[−.144, −.127]	[−.247, −.080]	[−.079, −.017]
Constant	[5.343, 5.351]	[5.339, 5.348]	[5.132, 5.179]
		C. Fixed Cost	
Nonstop origin			[−.387, −.327]
Nonstop destination			[−1.538, −1.473]
Constant			[1.227, 1.315]
		D. Variance-Covariance	
Variance demand	1.514	[2.354, 3.425]	[1.736, 1.876]
Variance marginal cost	.059	[.072, .132]	[.330, .353]
Variance fixed cost			[14.640, 15.636]
Demand–marginal cost covariance	.184	[.278, .504]	[.470, .512]
Demand–fixed cost covariance			[.674, .829]
Marginal cost–fixed cost covariance			[−.709, −.659]
		E. Market Power	
Median elasticity	[−8.163, −8.091]	[−7.281, −7.063]	[−4.105, −4.007]
Median markup	[28.146, 28.274]	[30.366, 31.564]	[53.617, 56.051]

NOTE.—Results from estimation of the model presented in sec. III. Column 1 presents the standard GMM estimation, col. 2 estimation using the methodology described in sec. II but holding market structure exogenous, and col. 3 estimation using the methodology described in sec. II. Column 1 presents the standard 95% confidence intervals. Columns 2 and 3 contain 95% confidence bounds constructed using the method of Chernozhukov, Hong, and Tamer (2007). The price coefficient is multiplied by 100.

A. *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).[27]

[27] 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.

In order to limit the space from which to draw for the minimization procedure, we standardize all the exogenous variables.[28] All the results are as expected and resemble those in previous work, for example, that of Berry and Jia (2010) and Ciliberto and Williams (2014).[29] Starting from the demand estimates, we find the price coefficient to be negative and included in $[-2.385, -2.185]$ and λ, the nesting parameter, to be between zero and one.[30] The corresponding median elasticity is included in $[-8.163, -8.091]$, and the confidence interval for the median markup is $[28.146, 28.274]$. 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.

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 assumptions 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 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 close to the ones in column 1.

B. *Results with Endogenous Market Structure*

Column 3 of table 4 displays the estimates from our model using the methodology developed in section II. We estimate the coefficient of price to be included in $[-1.557, -1.488]$ with a 95% 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 $[0.294, 0.366]$ (table 4, col. 2), while in the endogenous entry case, we estimate λ to be included in $[0.186, 0.206]$. 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. We also find that the coefficient of the market distance is larger, suggesting that self-selection is associated with market distance.

[28] See app. C for more details.

[29] We also have estimated the GMM model only with the demand moments, and the results were very similar. See app. D.

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

Overall, these sets of results lead us to overestimate the elasticity of demand and underestimate the market power of airline firms when we maintain that market structure is exogenous. To see this, we compare the implied mean elasticities in panel 5 of table 4. The mean elasticity for the exogenous market structure case is [−7.281, −7.063], while we estimate the mean elasticity is [−4.105, −4.007] when we allow for endogenous market structure. This leads to a difference in estimated markups: [30.366, 31.564] in the exogenous case compared with [53.617, 56.051] 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 [1.227,1.315], 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, not shown in the table, to be [$52,990, $59,275]. 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 with the number of (unidirectional) nonstop segments served by an airline, the number of (unidirectional) markets served by that airline is many times larger. That is, a single nonstop leg will be part of the service on many markets, and we cannot infer the cost of serving the single nonstop leg, which is bound to be much larger, from the fixed costs of serving the markets.[31]

Next, we investigate the estimation results for the variance-covariance matrix. The variance of demand error is included in [2.354, 3.425] in column 2 (exogenous market structure) and in [1.736, 1.876] in column 3 (endogenous market structure). The variance of the marginal cost unobservables is estimated in [0.072, 0.132] in column 2 and [0.330, 0.353] 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.[32] The variance of the fixed costs is included in [14.460,15.636].

The covariance between the demand and marginal cost is positive in all three columns. The covariance of the demand and fixed cost unobservables is estimated to be included in [0.674, 0.829], and the covariance between fixed and marginal costs unobservables is [−0.709, −0.659].

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

[32] See app. E for further discussion and comparisons of selected and unselected distribution of errors.

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 low 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 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 statistics, 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., price) must lie in the 95% confidence interval.[33]

Specifically, we draw 100 parameters from the identified set and simulate the model 200 times, using a new set of simulated unobservables. For any given market structure in any given market, we construct the confidence interval for prices by taking the 2.5 and 97.5 percentiles across parameter vectors. Then we compare the price for each airline for that market in the data with the confidence error for the predicted price. We do this again for product shares.[34]

The data lie within the confidence interval for prices 45.51% of the time, and our model fits the shares 39.77% of the time. 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 31.26% of the time (i.e., all six carriers have the correct participation in the market), and the model predicts a given airline's entry correctly 73.74% of the time.[35] In our sample, 8.7% of markets are

[33] 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.

[34] 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.

[35] These four numbers are not included in table 4 for the sake of brevity.

TABLE 5
AGGREGATE ENTRY PROBABILITIES

	AA	DL	LCC	UA	US	WN
Data	.390	.727	.175	.513	.488	.457
Model prediction	[.391, .395]	[.742, .745]	[.185, .189]	[.514, .518]	[.485, .490]	[.459, .464]

NOTE.—Entry probabilities across all markets in the sample described in the text. Intervals for the model are constructed using the subsampling routine described in the text.

not served by any carrier, while our model predicts this outcome in between 4.4% and 4.6% of markets.

VI. 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 cannot 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 DOJ 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 (e.g., Berry, Levinsohn, and Pakes 1995; Nevo 2001), and thus the postmerger and premerger market structures are the same, except that the products are now owned by a single firm.[36]

[36] Mazzeo, Seim, and Varela (2018) make a similar argument. They quantify the welfare effects of merger with endogenous entry/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 equilibria in both estimation and the merger analysis, whereas Mazzeo, Seim, and Varela (2018) assume a unique outcome from a selection rule based on ex ante firm profitability.

A. *The Price and Market Structure Effects of the AA-US Merger*

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

We consider four different scenarios about what it means for AA and US to merge. The four scenarios underscore the key observation that postmerger efficiencies could come from both observed and unobserved features of the carriers. The different assumptions that we discuss next allow us to check the robustness of the results of the counterfactual exercise and 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 premerger carriers and call this the best-case scenario.[37] 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 postmerger AA. For the fixed costs, we assign the highest level of nonstop origin and nonstop destination between AA and US to the postmerger AA. We implement the same procedure for the unobserved shocks. We use a new set of simulated unobservables (the same ones we used to determine the fit of our model), and we assign the best simulation draw (for utility, the highest; for costs, the lowest) between AA and US to the postmerger 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 the average of AA's and US's premerger unobservables. The results are presented as a subcase called "mean unobservables."

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 subcases to help us quantify the relative importance to the merger simulations of the efficiency in observables

[37] 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 procompetitive. 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 a 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.

TABLE 6
Aggregate Effects of Merger, per Market ($)

	Mean Fare	Consumer Welfare	Total AA + US Profit
Premerger	[229.32, 239.50]	[8,969, 10,063]	[2,403, 2,711]
Postmerger:			
Best case	[232.06, 242.38]	[9,509, 10,733]	[3,303, 3,796]
Mean unobservables	[229.91, 240.09]	[8,367, 9,400]	[1,798, 2,034]
New unobservables	[230.71, 240.95]	[8,663, 9,721]	[2,119, 2,390]
Average case	[231.74, 242.03]	[7,501, 8,411]	[1,868, 2,104]

Note.—Confidence intervals are constructed using the subsampling routine described in the text. Mean fares are in US dollars. Consumer welfare is the total compensating variation of the observed product offering in millions of US dollars. Total AA + US profit is the sum of profit across all markets in millions of US dollars.

versus unobservables. Last, we consider a scenario where the surviving firm takes on the mean values of the observed and unobserved characteristics from the two premerger firms and call this the "average-case scenario."

In table 6, we present confidence intervals for aggregate statistics to provide an industry-wide analysis of how a hypothetical merger would impact market structure, prices, consumer welfare, and producer surplus. The rows in table 6 represent the premerger predictions of the model (first row) and the four scenarios we consider after the merger.[38] Column 1 presents the 95% confidence interval for the average fare (share weighted across markets). Column 2 presents the total consumer welfare across all markets in millions of dollars, column 3 is the total profit for AA and US (summed over all markets) in millions of dollars.[39]

Under the best-case scenario, the confidence interval for average prices is slightly greater than the baseline, although the two intervals overlap, [$229.32, $239.50] versus [$232.06, $242.38]. Consumer welfare would increase from [$8,969 million, $10,063 million] to [$9,509 million, $10,733 million], as would the profit of the new merged firm compared with the sum of the premerger AA and US profit, [$2,403 million, $2,711 million] to [$3,303 million, $3,796 million]. This welfare increase is likely unreasonably large but highlights the importance of merger efficiency assumptions, as our other assumptions on merger efficiency imply lower consumer welfare.

In the subcases where only the best observable characteristics are inherited by the merged firm, consumer welfare may fall, as does the merged firm's profit. In the case where the new firm inherits the average premerger characteristics, consumer welfare may fall by even more. These results

[38] Notice that, in contrast to the standard approach that takes market structure as exogenous, the premerger simulations will not necessarily match the observed data on a market-by-market basis.

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

foreshadow an observation that we will again make later: unobservable characteristics of the firms play a crucial role in determining the welfare effects of a merger.

In table 7, we report changes in predicted entry probabilities after the merger for all four cases. Specifically, we display 95% confidence intervals 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 [0.391, 0.395] to [0.808, 0.812] of markets. The increase in entry is not surprising given that AA inherits all of US Airways'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.

Under the other three scenarios, AA sees a more modest but still substantial increase in the number of markets served, and the other airlines realize very slight increases in aggregate entry probabilities. In the remaining discussion in this section, we go deeper into the mechanisms that explain these aggregate changes by considering changes in particular types of markets.

We begin our analysis by looking at two sets of markets that are polar opposites in terms of postmerger effects: markets that were not served by any airline before the merger and markets that were served by American and US Airways 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 merger, which is clearly a strong reason for the antitrust authorities to allow for a merger to proceed. We also want to examine premerger duopolies, which are markets that are most likely to see high price increases and large welfare losses postmerger.

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 postmerger (columns) conditional on observing a market structure premerger (rows).[40] The 2×2 table consists of the two premerger market structures, with no firm in the market and with a duopoly of US and AA. The postmerger market structures are those markets with no firm in the market and with an AA/US monopoly.[41]

Table 8 shows that under the best-case scenario, the probability that the merged firm AA/US will enter a market as a monopolist that was not previously being served is 47.9%–48.3%, which is a large and positive effect

[40] Although our model is static, we use the term "transition" in order to convey predicted changes premerger to postmerger.

[41] The complete transition table would be of dimension 64×32 for each premerger market structure, which we do not present for practical purposes. Instead, we take slices of these tables.

TABLE 7
Entry Probabilities, Postmerger

	AA	DL	LCC	UA	US	WN
Premerger	[.391, .395]	[.742, .745]	[.185, .189]	[.514, .518]	[.485, .490]	[.459, .464]
Postmerger:						
Best case	[.808, .812]	[.739, .742]	[.183, .187]	[.511, .515]		[.457, .462]
Mean unobservables	[.637, .642]	[.742, .745]	[.185, .189]	[.514, .518]		[.459, .464]
New unobservables	[.585, .590]	[.742, .745]	[.186, .189]	[.515, .519]		[.459, .464]
Average case	[.538, .544]	[.744, .747]	[.187, .191]	[.517, .521]		[.461, .466]

Note.—Presented are entry probabilities across all markets in the sample described in the text. Confidence intervals are constructed using the subsampling routine described in the text.

TABLE 8
Market Structures in AA and US Monopoly and Duopoly Markets

	Postmerger Entry		Postmerger %Δ Price AA Monopoly
	No Firms	AA Monopoly	
Best-case scenario:			
No firms	[.517, .521]	[.479, .483]	
AA/US duopoly	[.000, .000]	[.958, .966]	[+13.6, +14.7]
Mean unobservables:			
No firms	[.821, .823]	[.177, .179]	
AA/US duopoly	[.000, .000]	[.947, .959]	[+4.9, +5.2]
New unobservables:			
No firms	[.588, .591]	[.409, .412]	
AA/US duopoly	[.322, .327]	[.559, .569]	[+16.5, +17.6]
Average-case scenario:			
No firms	[.906, .908]	[.092, .094]	
AA/US duopoly	[.000, .000]	[.904, .916]	[+4.8, +5.2]

Note.—Results from a counterfactual merger between AA and US. "Postmerger entry" is the likelihood of observing a column market structure given a premerger row market structure. "Postmerger %Δ price" is the percentage change in price for AA after the merger. Confidence intervals are constructed using the subsampling routine described in the text. The large price changes in the table are likely driven by extreme outliers in the simulation draws because they occur when the underlying probability of the event is very small.

of the merger that would be ignored by the standard economic analysis with exogenous market structure. We also find that there is a probability of 95.8%–96.6% that a market with a AA/US duopoly would be served by the merged firm as a monopoly after the merger. In those two-to-one cases, the merged firm would charge a higher price (13.6%–14.7%).

Results under the best-case scenario are different from the new unobservables scenario in terms of the transition probabilities but similar in terms of prices. Thus, the prices are computed on fewer markets under the second scenario, which is consistent with the firms self-selecting into markets. Notice that the unobservables under the best-case scenario are necessarily good ones because firms decided to enter into those markets premerger with those unobservables. We interpret this finding as supporting our self-selection hypothesis.

The predictions from the other two scenarios are remarkably different from the best-case scenario, 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 9.2%–9.4%, much lower than the best case. We also find it very likely that AA/US duopolies would turn into monopolies, and prices would increase by 4.8%–5.2% for those markets.

Comparing the four scenarios, we conclude that the unobservable characteristics play a crucial role in determining the effect of the merger on higher prices. This observation allows us to make an important point.

In any merger simulation that uses empirical industrial organization techniques (e.g., our method or more traditional methods like BLP), synergies from a merger could come through variables modeled by the researcher or variables unobserved by the researcher. This is an important distinction because it may be more viable for practitioners to successfully defend or prosecute a merger based on observable and measurable variables that can be clearly associated with the mechanisms of synergy. It would help to have direct information on the synergies claimed by the parties and how they are merger specific. We could use that knowledge to develop a fifth scenario to compare with the other four that would allow us to check on the credibility of the claimed synergies (under the maintained assumption that the model we are considering is correctly specified, of course).

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 AA/US duopoly premerger.

Under all scenarios, we find very little evidence that other competitors would enter. The most likely carriers to replace US are DL and UA, the two other major airlines. In those cases, we would expect prices to change by between −3.3% and 1.3% (average case, DL) or between 0.2% and 6.9% (average case, UA). There is up to a 3.6% chance that one of these legacy carriers replaces US for the average-case scenario. Overall, the

TABLE 9
Entry in Former AA and US Duopoly Markets

	Duopoly AA/ US and DL	Duopoly AA/ US and LCC	Duopoly AA/ US and UA	Duopoly AA/ US and WN
Best-case scenario:				
Probability of market structure	[.014, .018]	[.003, .005]	[.010, .013]	[.005, .008]
%Δ price AA	[+8.0, +24.8]	[+13.1, +39.9]	[+25.2, +38.2]	[+27.9, +43.5]
Mean unobservables:				
Probability of market structure	[.015, .023]	[.004, .007]	[.009, .013]	[.006, .010]
%Δ price AA	[−1.0, +2.8]	[+4.6, +17.5]	[−5.1, +4.1]	[+−.6, +11.7]
New unobservables:				
Probability of market structure	[.013, .020]	[.006, .010]	[.008, .012]	[.008, .011]
%Δ price AA	[+52.4, +73.4]	[+61.6, +90.1]	[+35.9, +90.4]	[+39.4, +78.1]
Average-case scenario:				
Probability of market structure	[.028, .036]	[.008, .011]	[.026, .030]	[.015, .018]
%Δ price AA	[−3.3, +1.3]	[+3.2, +13.4]	[−.2, +5.0]	[+.2, +6.9]

Note.—Results from a counterfactual merger between AA and US. "%Δ price AA" refers to percentage of change postmerger. Confidence intervals are constructed using the subsampling routine described in the text.

large price effects should be taken with a good dose of caution because they are computed out of few observations.

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 premerger.

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 consider only those markets that were monopolies before the merger.

In column 1, we display the likelihood that AA replaces the monopolist after the merger, and in column 2, 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 1.4% and 1.6% for the best-case scenario. It is much more likely that AA enters to form a duopoly, between 49.7% and 50.3%, in which case, the DL prices would fall by roughly 2%. 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 1%–2%.

TABLE 10
Postmerger Entry of AA in Former Monopolies

	AA Replacement Entry Probability	AA Entry	
		Entry Probability	Price Change (%)
Best-case scenario:			
DL	[.014, .016]	[.497, .503]	[−2.4, −2.2]
LCC	[.091, .106]	[.399, .415]	[−4.5, −4.2]
UA	[.039, .046]	[.490, .499]	[−3.4, −3.2]
WN	[.025, .029]	[.432, .438]	[−3.3, −3.1]
Mean unobservables:			
DL	[.005, .005]	[.201, .204]	[−2.1, −1.9]
LCC	[.027, .034]	[.163, .171]	[−4.2, −3.9]
UA	[.011, .013]	[.192, .196]	[−3.1, −2.9]
WN	[.008, .010]	[.176, .182]	[−2.9, −2.7]
New unobservables:			
DL	[.010, .011]	[.434, .440]	[−1.9, −1.8]
LCC	[.064, .076]	[.376, .390]	[−4.0, −3.8]
UA	[.025, .028]	[.441, .449]	[−2.8, −2.7]
WN	[.014, .018]	[.375, .382]	[−2.8, −2.7]
Average-case scenario:			
DL	[.001, .001]	[.081, .083]	[−1.3, −1.2]
LCC	[.009, .014]	[.082, .089]	[−3.4, −3.1]
UA	[.003, .004]	[.077, .082]	[−2.3, −2.2]
WN	[.002, .003]	[.085, .087]	[−2.1, −1.9]

Note.—Results from a counterfactual merger between AA and US. Postmerger entry is the likelihood of observing a column market structure given a premerger row market structure. Price change refers to the percentage change in price for the incumbent monopolist after AA joins as a duopolist postmerger. Confidence intervals are constructed using the subsampling routine described in the text.

Under the 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 greater competition in entry from the other major carriers but reduced competition from low-cost 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,

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

	Probability of Exit	AA Price Change (%)
Best-case scenario:		
DL	[.009, .010]	[+5.1, +11.2]
LCC	[.048, .077]	[−7.4, +1.6]
UA	[.014, .020]	[+6.4, +11.7]
WN	[.015, .020]	[+2.5, +11.2]
Mean unobservables:		
DL	[.005, .008]	[−11.5, −4.3]
LCC	[.032, .048]	[−16.6, −3.1]
UA	[.011, .015]	[−12.2, −5.8]
WN	[.008, .012]	[−9.6, +2.4]
New unobservables:		
DL	[.006, .007]	[−12.2, −2.9]
LCC	[.023, .031]	[−43.0, −30.9]
UA	[.012, .014]	[−17.8, −10.9]
WN	[.005, .008]	[−25.8, −13.2]
Average-case scenario:		
DL	[.002, .003]	[−12.3, −1.4]
LCC	[.013, .023]	[−33.6, −8.5]
UA	[.003, .006]	[−19.0, −7.9]
WN	[.002, .006]	[−32.4, −16.0]

Note.—Results from a counterfactual merger between AA and US. Postmerger entry is the likelihood of observing a column market structure given a premerger row market structure. Row market structures are a duopoly between AA and the listed airline. Price change refers to the percentage change in price for AA after the merger and subsequent rival exit. Confidence intervals are constructed using the subsampling routine described in the text.

therefore lowering price and making it difficult for the rival to earn enough variable profit to cover fixed costs.[42] Second, AA might become more attractive to consumers after the merger and steal business from rivals. For ease of exposition, we consider only markets where AA and other incumbents were in the market, and we do not report the results for the other merging firm, US Airways.

The first row of column 1 in table 11 shows that, for the best-case scenario, there is a probability of 0.9%–1.0% that DL will leave the duopoly market with AA after the merger. In such cases, AA's price will be 5.1%–11.2% higher. Overall, the greatest likelihood of exit, by far and across all scenarios, is for the LCC airline.

B. *The Economics of Mergers at a Concentrated Airport: Reagan National Airport*

The DOJ reached a settlement with American and US Airways to drop its antitrust challenge if the two 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.[43]

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.[44]

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 an AA/DL market with probability [0.959, 1.000] for the best-case scenario and [0.922, 0.954] for the average-case

[42] AA could experience either 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.

[43] 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. Two recent papers have looked specifically at slot divestitures: Ali (2020) and Park (2020).

[44] 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.

TABLE 12
POSTMERGER ENTRY AND PRICING AT DCA

	AA/DL	AA/UA	AA/DL/LCC	AA/DL/UA	AA/DL/WN
Best-case scenario:					
AA, US, DL markets:					
Market structure transitions	[.959, 1.000]	[.000, .000]	[.000, .023]	[.000, .005]	[.000, .025]
%Δ shares weighted price	[+6.5, +7.4]	NA	[−18.1, +34.2]	[+2.8, +2.8]	[−23.2, +3.8]
AA, US, DL, UA markets:					
Market structure transitions	[.001, .003]	[.000, .000]	[.000, .000]	[.987, .999]	[.000, .000]
%Δ shares weighted price	[+18.5, +20.4]	NA	NA	[+4.9, +5.4]	NA
Average-case scenario:					
AA, US, DL markets:					
Market structure transitions	[.922, .954]	[.000, .000]	[.020, .038]	[.012, .024]	[.009, .025]
%Δ shares weighted price	[+3.1, +4.1]	NA	[−8.2, +4.3]	[−4.1, +5.9]	[−26.2, +2.8]
AA, US, DL, UA markets:					
Market structure transitions	[.000, .000]	[.000, .000]	[.000, .000]	[.954, .986]	[.000, .000]
%Δ shares weighted price	NA	NA	NA	[+3.4, +3.6]	NA

NOTE.—Counterfactual predictions for markets with DCA as one endpoint. Comparisons are based on premerger market structures of AA/DL/UA and AA/US/DL/UA. Postmerger market structures are listed in the columns. Confidence intervals are constructed using the subsampling routine described in the text. NA = not applicable.

scenario, for example. We find that this would result in a rise in prices in both scenarios, but with a higher price rise in the best-case scenario.

In none of the premerger markets where AA and US were both present were LCC or WN 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 DOJ's 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. This is accompanied by an increase in prices of about 3%–5%, depending on the scenario. Similar results are found for the average-case scenario.

Overall, our results suggest that the decisions made by the DOJ to facilitate access to airport facilities by new entrants were justified and should help control the postmerger increase in prices and promote low-cost carrier coverage at DCA.

VII. 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) that allowing for the selection of firms into markets based on unobservables can lead to different estimates of price elasticities and markups than those we find when we assume that market structure is exogenous to the pricing decision and (ii) that 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 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 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 up-front 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; Berry 1994) that are taken as exogenous. More specifically, the classical model takes as exogenous the following: 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 has been done by Li et al. (2021). 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." *J. Econometrics* 168 (1): 156–73.

Ali, Ratib. 2020. "Effect of Airport Slots in Competition and Antitrust Policy: Evidence from a Recent Merger." Working paper, Boston Coll.

Berry, Steven, Michael Carnall, and Pablo T. Spiller. 2006. "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity." In *Advances in Airline Economics 1: Competition Policy and Antitrust*, edited by Darin Lee, 183–214. Bingley, United Kingdom: Emerald Group.

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

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.

Berry, Steven T. 1990. "Airport Presence as Product Differentiation." *A.E.R.* 80 (2): 394–99.

———. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica* 60 (4): 889–917.

———. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND J. Econ.* 25 (2): 242–62.

———. 1999. "Free Entry and Social Inefficiency in Radio Broadcasting." *RAND J. Econ.* 30 (3): 397–420.

Borenstein, Severin. 1989. "Hubs and High Fares: Dominance and Market Power in the US Airline Industry." *RAND J. Econ.* 20 (3): 344–65.

Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." *J. Indus. Econ.* 35 (4): 457–82.

———. 1989. "Empirical Studies of Industries with Market Power." *Handbook Indus. Org.* 2:1011–57.

Bresnahan, Timothy F., and Peter C. Reiss. 1990. "Entry in Monopoly Market." *Rev. Econ. Studies* 57 (4): 531–53.

———. 1991. "Entry and Competition in Concentrated Markets." *J.P.E.* 99 (5): 977–1009.

Brueckner, Jan K., and Pablo T. Spiller. 1994. "Economies of Traffic Density in the Deregulated Airline Industry." *J. Law and Econ.* 37 (2): 379–415.

Chernozhukov, Victor, Han Hong, and Elie Tamer. 2007. "Estimation and Confidence Regions for Parameter Sets in Econometric Models." *Econometrica* 75 (5): 1243–84.

Ciliberto, Federico, Amalia R. Miller, Helena Skyt Nielsen, and Marianne Simonsen. 2016. "Playing the Fertility Game at Work: An Equilibrium Model of Peer Effects." *Internat. Econ. Rev.* 57 (3): 827–56.

Ciliberto, Federico, and Elie Tamer. 2009. "Market Structure and Multiple Equilibria in Airline Markets." *Econometrica* 77 (6): 1791–828.

Ciliberto, Federico, and Jonathan W. Williams. 2014. "Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry." *RAND J. Econ.* 45 (4): 764–91

Cohen, Andrew M., and Michael J. Mazzeo. 2007. "Market Structure and Competition among Retail Depository Institutions." *Rev. Econ. and Statis.* 89 (1): 60–74.

Draganska, Michaela, Michael Mazzeo, and Katja Seim. 2009. "Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions." *Quantitative Marketing and Econ.* 7 (2): 105–46.

Eizenberg, Alon. 2014. "Upstream Innovation and Product Variety in the US Home PC Market." *Rev. Econ. Studies* 81 (3): 1003–45.

Ellickson, Paul B., and Sanjog Misra. 2012. "Enriching Interactions: Incorporating Outcome Data into Static Discrete Games." *Quantitative Marketing and Econ.* 10 (1): 1–26.

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

Fan, Ying. 2013. “Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market.” *A.E.R.* 103 (5): 1598–628.

Fan, Ying, and Chenyu Yang. 2020. “Competition, Product Proliferation and Welfare: A Study of the US Smartphone Market.” *American Econ. J. Microeconomics* 12 (2): 99–134.

Gandhi, Amit, and Jean-François Houde. 2019. “Measuring Substitution Patterns in Differentiated Products Industries.” Working paper, Univ. Wisconsin–Madison and Wharton School.

Gentry, Matthew, and Tong Li. 2014. “Identification in Auctions with Selective Entry.” *Econometrica* 82 (1): 315–44.

Goolsbee, Austan, and Chad Syverson. 2008. “How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines.” *Q.J.E.* 123 (4): 1611–33.

Gronau, Reuben. 1974. “Wage Comparisons—A Selectivity Bias.” *J.P.E.* 82 (6): 1119–43.

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*, vol. 5, no. 4, edited by Sanford V. Berg, 475–92. Cambridge, MA: NBER.

———. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica* 47:153–161.

Ho, Katherine. 2009. “Insurer-Provider Networks in the Medical Care Market.” *A.E.R.* 99 (1): 393–430.

Jeziorksi, Przemyslaw. 2014a. “Effects of Mergers in Two-Sided Markets: The US Radio Industry.” *American Econ. J. Microeconomics* 6 (4): 35–73.

———. 2014b. “Estimation of Cost Synergies from Mergers: The US Radio Industry.” *RAND J. Econ.* 45 (4): 816–46.

Lee, Robin. 2013. “Vertical Integration and Exclusivity in Platform and Two-Sided Markets.” *A.E.R.* 103 (7): 2960–3000.

Li, Sophia, Joe Mazur, James Roberts, Yongjoon Park, Andrew Sweeting, and Jun Zhang. 2021. “Repositioning and Market Power after Airline Mergers.” Working paper, Univ. Maryland.

Li, Tong, and Bingyu Zhang. 2015. “Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects.” *American Econ. J. 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.” *Rev. Econ. Studies* 76 (4): 1397–429.

Mazzeo, Michael J. 2002. “Product Choice and Oligopoly Market Structure.” *RAND J. Econ.* 33 (2): 221–42.

Mazzeo, Michael J., Katja Seim, and Mauricio Varela. 2018. “The Welfare Consequences of Mergers with Endogenous Product Choice.” *J. Indus. Econ.* 66 (4): 980–1016.

Mizuno, Toshihide. 2003. “On the Existence of a Unique Price Equilibrium for Models of Product Differentiation.” *Internat. J. Indus. Org.* 21 (6): 761–93.

Nevo, Aviv. 2000. “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry.” *RAND J. Econ.* 31 (3): 395–421.

———. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica* 69 (2): 307–42.

Nocke, Volker, and Nicolas Schutz. 2018. “Multiproduct-Firm Oligopoly: An Aggregative Games Approach.” *Econometrica* 86 (2): 523–57.

Olley, G. Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64 (6): 1263–97.

Pakes, Ariel, J. Porter, Joy Ishii, and Kate Ho. 2015. “Moment Inequalities and Their Application.” *Econometrica* 83:315–34.

Panzar, John C. 1979. “Equilibrium and Welfare in Unregulated Airline Markets.” *A.E.R.* 69 (2): 92–95.

Park, Yongjoon. 2020. “Structural Remedies in Network Industries: An Assessment of Slot Divestitures in the American Airlines/US Airways Merger.” Working paper, Univ. Maryland.

Petrin, Amil, and Boyoung Seo. 2017. “Identification and Estimation of Discrete Choice Demand Models When Observed and Unobserved Product Characteristics Are Correlated.” Working paper, Univ. Minnesota.

Reiss, Peter C., and Pablo T. Spiller. 1989. “Competition and Entry in Small Airline Markets.” *J. Law and Econ.* 32 (2): S179–S202.

Roberts, James, and Andrew Sweeting. 2013. “When Should Sellers Use Auctions?” *A.E.R.* 103 (5): 1830–61.

Rosse, James N. 1970. “Estimating Cost Function Parameters without Using Cost Data: Illustrated Methodology.” *Econometrica* 38 (2): 256–75.

Schmalensee, Richard. 1989. “Inter-Industry Studies of Structure and Performance.” In *Handbook of Industrial Organization*, vol. 2, edited by Richard Schmalensee and Robert Willig, 951–1009. Amsterdam: North-Holland.

Seim, Katja. 2006. “An Empirical Model of Firm Entry with Endogenous Product-Type Choices.” *RAND J. Econ.* 37 (3): 619–40.

Sovinsky Goeree, Michelle. 2008. “Limited Information and Advertising in the US Personal Computer Industry.” *Econometrica* 76 (5): 1017–74.

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): 1763–803.

Tamer, Elie. 2003. “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria.” *Rev. Econ. Studies* 70 (1): 147–65.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* Cambridge: Cambridge Univ. Press.