Market Structure and Multiple Equilibria in Airline Markets*

Federico Ciliberto†
North Carolina State University

Elie Tamer‡
Princeton University

INCOMPLETE WORKING PAPER
THIS VERSION: Nov 2003

Abstract

We provide a framework for inference in discrete games that involve multiple decision makers and use it to study airline market structure in the US. We make inferences on a "class of models" rather that looking for point identifying assumptions that pin down a unique model. We extend the empirical literature on entry and market structure started in Bresnahan and Reiss (1990). We allow for a more flexible model of entry, heterogeneity and player identities without making assumptions on equilibrium selection. Our estimation strategy is directed at a class of models that obey the fundamental assumption that if a firm enters a market it expects nonnegative profits. This fundamental condition provides a set of inequality restrictions on regressions that we exploit to learn about the profits of various firms. We then examine airline market structure focusing on the strategic behavior between a set of airlines. This allows us to study the effect of specific airlines on each others entry decision.

*A part of this paper was circulated under the title "Empirical Strategies for Estimating Discrete Games with Multiple Equilibria." We thank T. Bresnahan, A. Cohen, P. Haile, C. Manski, M. Mazzeo, J. Panzar, A. Paula, R. Porter, W. Thurman, and seminar participants at many institutions for comments. We also thank Shinichi Sakata for help with a version of his genetic algorithm that we use and Tom Whalen at the Department of Justice for useful insights on airlines' entry decisions. Berna Karali provided excellent research assistance.

†Department of Agricultural and Resource Economics, North Carolina State University, `Federico_Ciliberto@ncsu.edu`, ph: 919-515-8946.

‡Department of Economics, Princeton University, `tamer@princeton.edu`, ph: 609-258-4032.

Support from the National Science Foundation and from the Chow Econometrics Research Program at Princeton is gratefully acknowledged.

1 Introduction

We provide a framework for inference in discrete games that involve multiple decision makers and use it to study airline market structure in the US. Generally, multiple pure strategy equilibria are common in discrete games with a sufficiently rich environment. Making inferences in the presence of multiple equilibria in complete information games is complicated since the underlying econometric model is incomplete, i.e., for a given parameter and a given value for the exogenous variables, the model does not predict a unique outcome. Strategies to deal with multiplicity have involved a variety of assumptions to point identify the parameters of interest. A common approach to these strategies is to base inference on a *unique* feature that is common to all equilibria. In 2×2 entry games for example, Bresnahan and Reiss (1991) show that the model predicts a unique number of players in the market which allows one to use standard methods (like maximum likelihood) to do inference based on the number of firms. This approach has a few shortcomings. First, it is not clear how one would look for these unique features in general discrete games involving many players and/or large strategy space. Second and more important, strong assumptions, such as not allowing for player specific heterogeneity or externalities, are needed to guarantee uniqueness across all equilibria. In this paper we take a different approach.

We make inferences directly on a "class of models" rather than looking for (point identifying) assumptions that pin down a unique model. Taking a class of models approach to modelling in game theoretic settings, one "abandon(s) the aim of identifying some unique equilibrium outcome. Instead, we admit some class of candidate models (each of which may have one or more equilibria) and ask whether anything can be said about the *set* of outcomes that can be supported as an equilibrium of any candidate model" (For more on this, see Sutton (1991)). The inferential strategy that we propose in this paper implements this approach by estimating the identified features of the class of models under consideration. Moreover, our framework is practical and can be used in a variety of empirical games.

We apply our method to entry models and estimate profit functions of firms in the airline industry. Following Bresnahan and Reiss (1990) and Berry (1992), we consider a two-stage model of entry. In the first period firms decide whether or not to enter into the market. In the second period the firms, producing a homogeneous good, compete in quantities. We study the entry model under the following critical but plausible assumption that we impose throughout.

Assumption 1 *All entrants in a market make nonnegative profits from entry, while firms that do not enter expect negative profits from entry.*

This assumption, which represents a necessary and weak condition on behavior, is satisfied in a class of entry models and is similar to the "viability condition" of Sutton. The question that we ultimately answer is what can we learn about the profit function of different firms under this fundamental (weak but credible) assumption. A heuristic answer which the paper elaborates on, is that given a cross section of markets or firms, the identified feature of the model is the *set* of profit functions that obey our assumption and are consistent with the empirical evidence. This approach to inference has been used in a class of English auctions studied in Haile and Tamer (2003). Andrews and Berry (2003) provide another approach to inference in entry models with multiple equilibria.

More generally, the paper presents a practical estimator that can be used to conduct robust inference in parametric discrete games involving multiple decision makers. This estimator is based on minimizing the distance between the empirical choice probabilities (which can be consistently estimated from a cross section of markets or firms for example) and the predicted ones defined in terms of inequality restrictions following assumption 1 and provides a set that covers the identified feature (a set of parameters or a point) with a prespecified probability. Tamer (2003) studied identification in a 2×2 game where no assumptions were made about equilibrium selection in different markets. However, he showed that to guarantee point identification, strong support conditions (a lot of variation in the regressors) were needed and these conditions are not easy to obtain in general more practical games involving many players and/or large strategy space. In this paper, we abandon this search for point identifying assumptions and provide empirical strategies to estimate the identified features of the game.

The airline industry is a natural choice to study entry in the presence of multiple equilibria. Similar markets differ in the number and identities of carriers that serve them. For example Detroit-Pittsburgh was served by 3 airlines in 1998, while Boston-Pittsburgh was only served by 1 airline.[1] We identify the effect of technological (such as distance) and demographic characteristics from the effect of airlines' strategic behavior on market structure. Strategic behavior is sometimes airline specific and our empirical strategy allows for a potential entrants' profits to depend on the *identity* of other players in the market not just the number of competitors. This is crucial for example when dealing with airlines that follow a unique business model, such as Southwest.

Multiple equilibria, that sometimes involve a *different* number of firms, are likely to occur in the entry game for a variety of reasons. First, allowing for simple heterogeneity in the effect each airline has on the other can result in multiple equilibria that differ in the number of firms. For instance, major airlines contribute to the construction and maintenance of the airports more than small airlines do, creating free-riding opportunities for low cost carriers[2]. On the other hand, entrants must often sublease airport facilities from the larger airlines, usually at higher costs than if they leased them directly from the airport (U.S. General

[1] These data are from the T-100 Domestic Segment Dataset, which contains domestic non-stop segment data for the United States (See the appendix for more information on the data).

[2] This is why airports and airlines sign majority in interest agreements, which we discuss in section 2 (U.S. General Accounting Office (1989)). Major carriers are American, Continental, Delta, Northwest, TWA, USAir, and United.

Accounting Office (1989)). Thus major carriers can strategically accommodate new entrants, favoring the entry of smaller carriers that have a complementary network (where small carriers bring in passengers that use markets served by the major airlines) or make it harder for direct competitors to enter. Second, demand for airline travel is highly nonlinear as business travellers have different demand schedules than leisure travellers[3]. Third, Hendricks, Piccione, and Tan (1997) show that there can be multiple equilibria in a game where hub-and-spoke networks are explicitly modelled and small regional carriers have lower costs than the major airlines.

In this paper we do not sort out which of these possible reasons give rise to multiple equilibria. To do this, one would need to estimate simultaneously the demand and supply of airline services, and the entry decision as in Reiss and Spiller (1989).[4] Here, we contribute to the literature started by Bresnahan and Reiss (1990) and Berry (1994) and focus on estimating airline profit functions. However, rather than estimating the probability of observing a given number of firms, we provide upper and lower probabilities on observing a particular market structure that involves certain firms. Our approach allows us to include, for the first time in the literature, *firm specific indicator variables* in the profit functions of other airlines.

We address important questions concerning the airline industry: Is the dispersion in the number of firms across airline markets the result of the carriers' strategic behavior? Do major carriers behave more aggressively toward some low-cost carrier than others? These questions have wide policy relevance as the recent report by the Transportation Board writes, "where entry is not artificially impeded, competing services will ensure that the fares charged to passengers are, in the long run, reflective of the full cost of efficiently providing the type of service desired."[5]

In section 2, we review the literature on entry and market structure in the airline industry. This sheds light on heterogeneity between different airlines. We then provide a summary of our empirical model of entry that we use and relate it to models of entry in the literature. Our econometric framework for estimating discrete games with multiple equilibria is provided in section 4 below. We then provide a detailed description of our data in section 5. Section 6 provides our empirical results and section 7 concludes.

2 Entry and Market Structure in the Airline Industry

With the notable exceptions of Berry (1992) and Reiss and Spiller (1989), most of the previous literature has studied the effect of airport and route concentration on airline fares rather than on market structure. Borenstein (1988), Evans and Kessides (1993), and Hurdle, Johnson, Joskow, Werden, and Williams (1989) find that airport concentration is associated

[3]Even a simple homogeneous good oligopoly entry game can have multiple equilibria. See Tirole (1988).
[4]This challenging work was done for small markets and under more restrictive conditions.
[5]Transportation Research Board (1999), page 1-7.

with higher fares for consumers who fly to or from the airport.[6] However, omitted cost and demand variables or selection effects could explain the observed correlations between market concentration and fares, and these same unobservable factors could also affect entry (Reiss and Spiller (1989)).[7] To address this concern, Reiss and Spiller (1989) model simultaneously market structure and pricing behavior in the airline industry. They estimate the probability that a firm offers direct service conditional on the number of firms offering indirect service. In their econometric analysis, they impose the restrictive condition that "the number of indirect firms (does not) vary with the number of direct firms," and "thus ensures that the likelihood function is well defined, even for ranges of the unobservable that would otherwise make the model's prediction non-unique."[8] Reiss and Spiller find that entry introduces a selection bias in equations explaining fares or quantities and that there is considerable variation in competitive conduct both within and across routes. Their results suggest that unobservable firm heterogeneity in different markets may be important in determining the effect of market power on airline fares.

Berry (1992) studies the effect of market presence on entry in the airline industry. Berry focuses on estimating the probability of observing a given number of firms, conditional on exogenous market characteristics. The number of firms is endogenous in Berry's analysis, and unrestricted. Moreover, Berry models observable (and unobservable) firm heterogeneity, which is a critical achievement in the analysis of market structure.[9]. Berry finds that airport presence has an important role in determining airline profitability, providing support to the studies that show a strong positive relationship between airport presence and airline fares. Berry also finds that profits decline rapidly in the number of entering firms. The main methodological innovation in Berry (1992) (and in Bresnahan and Reiss (1991)) is to avoid the problem of multiple equilibria by providing assumptions on the model that guarantee uniqueness of the predicted *number* of firms in the market rather than trying to predict which particular firms are more likely to serve a market. By focusing on the number of firms rather than on their identities, Berry (1992) can then study market structure in models where the number of firms is endogenous. Mazzeo (1997) used a similar approach to study markets where firms produce differentiated goods. Berry, Mazzeo, and Bresnahan and Reiss ensure the uniqueness of equilibrium by assuming that the second stage Cournot game has a unique

[6]Bruekner (2002) estimates the effect of market concentration on delays at American airports using the analogous approach that Borenstein and the rest of the literature has used to study the effect of market concentration on prices.

[7]Borenstein (1992) instruments market concentration with the probability that potential entrants do actually enter into the market. This identification assumption is valid only if the number and identity of potential entrants does not depend on market fares, which is unlikely (Borenstein [1992]).

[8]Reiss and Spiller (1989), page 188.

[9]Bresnahan and Reiss (1990) model firm heterogeneity as a function of observable market specific variables and firm specific parameters. The firm specific parameters are estimated using an assumption on the order of entry of firms. See Bresnahan and Reiss (1990), page 549. The difference with Berry (1992) is that Berry introduces firm heterogeneity as an observable firm specific variable, and that Berry's method can be extended to games with many players in a straightforward fashion

equilibrium, and by assuming that the firms' entry costs are independent of which firms are in the market. Our approach avoids both assumptions.

The methods that we use are complementary to those proposed by Bresnahan and Reiss (1991), Berry (1992), and Mazzeo (1997). Rather than estimating the probability of observing a given number of firms, we will provide upper and lower probabilities on observing a particular market structure that involves certain firms.[10]. Furthermore, we include firm specific indicator variables in the profit function of each other airline. An early attempt to investigate market structure considering the identity of firms was made by Morrison and Winston (1990). They ran a probit specification where the dependent variable was whether or not American Airlines serves a market. In their analysis, independent variables included the number of enplanements of American and of its competitors at the endpoints of the market, which Morrison and Winston assumed to be exogenous. However, as we know from Reiss and Spiller (1989) and Berry (1992), entry selection most likely biases Morrison and Winston (1990)'s estimates. We will show that this is actually the case in our empirical analysis.

3 Empirical Models of entry

Following Reiss and Spiller (1989), Bresnahan and Reiss (1990), Berry (1992), Toivanen and Waterson (2000), and Mazzeo (1997), we consider a two-stage model of entry and assume that a pure strategy equilibrium exists. We assume that a reduced form of the profit for firm i in market m can be written as:

$$\ln \pi^*_{im} \simeq \alpha_{i0} + \alpha_{iX} \ln X_m + \sum_{j \neq i} \delta_{ji} d_{jm} + \epsilon_{im}.$$

where X_m are exogenous determinants of demand. ϵ_{im} is the part of firms' profits that we cannot measure but that firms observe and act on. A market is a route, or spoke, between two airports. We are specially interested in the estimation of δ_{ji}, the effect that the presence of one of the other airlines has on the probability of observing firm i in the same market. For example, the parameters δ_{ji} could measure a particular aggressive behavior of one airline (e.g. American) against another airline (e.g. Southwest), or could measure cost externalities among airlines at airports. δ_{ji} captures the effect that the entry by airline j has on i's unobserved profits. In this way we capture the possibility that the unobserved part of the firms' profits might change with the number and identity of entrants. We show that δ_{ji} can be positive, exactly as Bresnahan and Reiss (1990) found in the study of their "*Model* 3". Bresnahan and Reiss found that if they did not constrain their parameters so that the duopoly profits were smaller than the monopoly profits, then they would find that the event

[10] Assuming an order of entry it is possible to estimate the probability of observing a particular market structure that involves certain firms. We do not assume any order of entry. See Berry [1992], page 916

of higher duopoly profits than monopoly profits would take on positive probability for some range of market sizes. With our new estimation methodology we do not need to impose that duopoly profits are smaller than the monopoly profits. Our method is flexible enough to allow us to estimate the duopoly and monopoly profits for each firm only using Assumption 1.

Firm i enters market m if:

$$\alpha_{i0} + \alpha_{iX} \ln X_m + \sum_{j \neq i} \delta_{ji} d_{jm} + \epsilon_{im} \geq 0. \tag{1}$$

This leads into the following statistical model:

$$P\left(d_{im} = 1\right) = P\left(\pi_{im} \geq 0\right) = P\left(\alpha_{i0} + \alpha_{iX} \ln X_m + \sum_{j \neq i} \delta_{ji} d_{jm} + \epsilon_{im} \geq 0\right).$$

We now show that the econometric models used by Bresnahan and Reiss (1990) and Berry (1992) are special cases of the specification above.

3.1 Bresnahan and Reiss (1992): Unobservable heterogeneity in a simultaneous move entry game

Bresnahan and Reiss assume that players' profits are the same except for a mean zero unobserved heterogeneity variable that we denote here by ϵ_{im}. Therefore firm i enters into the market if:

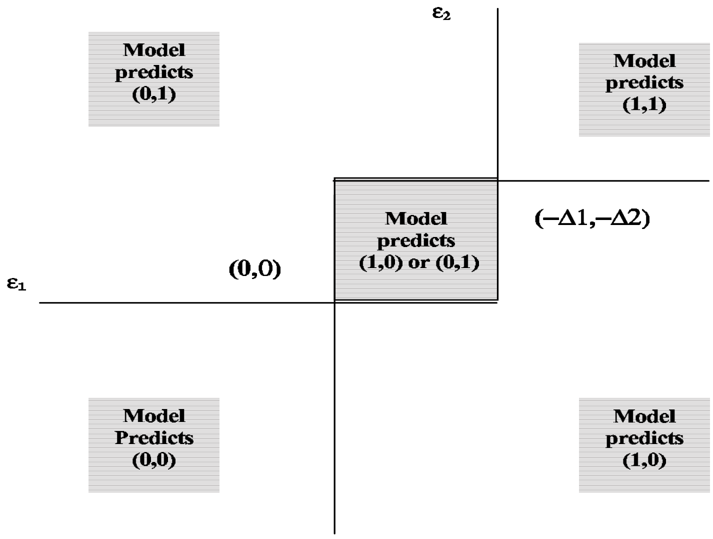
$$\alpha_0 + \alpha_X \ln X_m + \delta N_m + \epsilon_{im} \geq 0,$$

where no firm specific variables enter the observed profit function ($\beta = 0$ and $\delta = \delta_{ji}$ for all i and j). Bresnahan and Reiss show that even this simple model leads to multiple equilibria. To see why, consider a simple 2x2 entry game, where the payoff for player i ($i = 1, 2$) is

$$y_i = 1[\Delta_i y_{3-i} + \epsilon_i \geq 0] \tag{2}$$

Δ_i represents the change in player $i's$ profits from having player $3 - i$ enter the market. We omit regressors for simplicity. BR assume that duopoly profits are lower than monopoly profits, $\Delta_i < 0, i = 1, 2$. Here, multiple equilibria in the identity, but not number, of firms arise when $0 \leq \epsilon_i \leq -\Delta_i$ for $i = 1, 2$. This can be seen in Figure 1. The shaded center region of the figure contains payoff pairs where either firm could enter as a monopolist in the simultaneous-move entry game. This implies that the model predicts only upper and lower probabilities on the outcomes $(0, 1)$ and $(1, 0)$. To deal with this, BR transform the model into one that predicts the number of entrants in the market. However as our new empirical model and results show, different equilibria can exist with different number of players in richer games. Heuristically, in 3-player games where one is a large firm and the other two are small

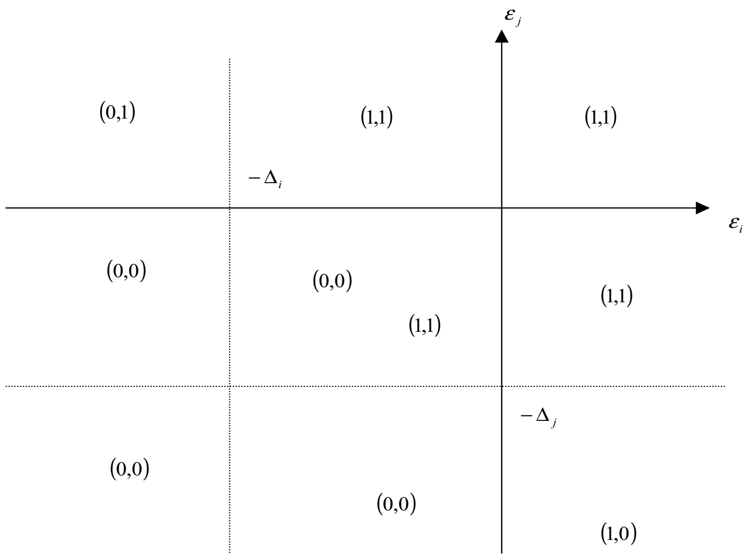
Figure 1: Prediction of the Bivariate Model for $\Delta_1 < 0$ and $\Delta_2 < 0$



firms, there can be multiple equilibria where one equilibrium includes the large firm as a monopolist, while in the other the smaller two firms enter as duopolists (more on this in the empirical section). This is common in games where the effect of one firm ("its Δ") is allowed to be different than the effect (the Δ's) of other firms. This type of heterogeneity is not allowed in the Bresnahan and Reiss (1990) and the Berry (1992) settings. Another source of multiple equilibria are the externalities that can exist among entrants because airlines share airport costs and because carriers can strategically accommodate new entrants. The relevance of externalities can be represented in the simple $2x2$ discrete game, illustrated in figure 2. In this case where $\Delta_i > 0$ for $i = 1, 2$, we have that for $-\Delta_i \leq \epsilon_i \leq 0$ both players enter or no player enters. Here, a player benefits from having the other player entering the market. We can again use BR's approach and estimate the probability of the outcome $(1, 0)$, of the outcome $(0, 1)$, and of the outcome "either $(1, 1)$ or $(0, 0)$."

The presence of heterogeneity and the externalities among firms can explain the reason why we see a large variation in the number of airlines serving otherwise similar markets (See Table 3). We show that this variation is not explained solely by technological and/or demographic differences across markets. The class of models that we consider under assumption 1 above contains models that allow for heterogeneity and externalities to arise.

Figure 2: Prediction of the Bivariate Model for $\Delta_1 > 0$ and $\Delta_2 > 0$



3.2 Berry (1994): Observable heterogeneity in sequential-move entry game.

Berry proposes a method to add observable firm heterogeneity to Bresnahan and Reiss's framework while allowing one to examine games with many players. [11]

In particular, firm i enters into the market if:

$$\alpha_0 + \alpha_X \ln X_m + \delta N_m + \beta \ln Z_{im} + \epsilon_{im} \geq 0,$$

The crucial difficulty with this approach when using standard econometric techniques is to estimate β. Berry shows that with an order of entry assumption, one can estimate β using (simulated) method of moments. In particular Berry either adds a dynamic element to the analysis using information from previous periods to identify incumbents, or he assumes that the order of entry follows the rank of profitability of the firms. While intuitive from an economic point of view, the main drawback of Berry's approach is that it is not necessary that the firms that enter first are the ones that make the highest profits (e.g. face the lowest fixed costs). In addition, the model above does not allow the fixed costs to depend on other players in the market nor does it allow for inference on the identity of the firms without an assumption on the order of entry.

[11] As we discuss in footnote 9, Bresnahan and Reiss also model firm heterogeneity, but in their context it is a function of observable market characteristics and firm specific parameter. We introduce observable heterogeneity in a companion paper Ciliberto and Tamer (2003).

As Sutton (1998) suggests, one does not know the specific form of entry process and we are better off beginning with a general class of multistage games. In particular, one can think of a multistage game where each firm decides whether to enter at some date $t \leq T$, its date of arrival in the market. By the final date T, one can summarize the outcome of the entry process by describing the characteristics of the entrants and modelling the form of competition. The firms have reached a long-run equilibrium that one uses to infer their functions.

Other models of entry are Mazzeo (2002), who allows for firm types, and Seim (2002), who changes the informational structure of the game by transforming it into one of incomplete information.

4 Estimating Discrete Response Models with Multiple Equilibria

Simultaneity in binary response models is an important topic in econometrics which dates back at least to the work of Heckman (1978). It is also examined in chapter 5 of Maddala (1983). There, a class of dummy endogenous models was studied and *coherency* conditions were imposed to ensure that the likelihood is well defined (see Tamer (2003) for more detail). In the case where the model is a representation of a discrete game with two decision makers, imposing the coherency conditions eliminates simultaneity, an essential feature the model is trying to capture. Tamer (2003) studied an incomplete simultaneous discrete response model that is written as a set of inequality restrictions on regressions. His inference method, which requires point identification of the parameters using support conditions in a bivariate game, does not extend in a simple way to larger games with many players and/or large strategy space. In this paper, we take a slightly different approach. We do not look for point identification conditions, but focus on the identified *features* of the set of inequality restrictions on regressions that are derived from the necessary conditions that define the set of models we consider. This allows us to study general games with many players and a large strategy space without making any equilibrium selection assumptions or other assumptions limiting the type of heterogeneity allowed. Most importantly, this strategy is practical and can be implemented. It uses a modified minimum distance estimator that heuristically minimizes the distance between the vector of (conditional) empirical choice probabilities (the empirical evidence or the data) and the *region* of feasible choice probabilities predicted by the model. The method relies on one being able to *derive analytically* the inequality restrictions. We highlight our method first in a simple 2×2 binary game.

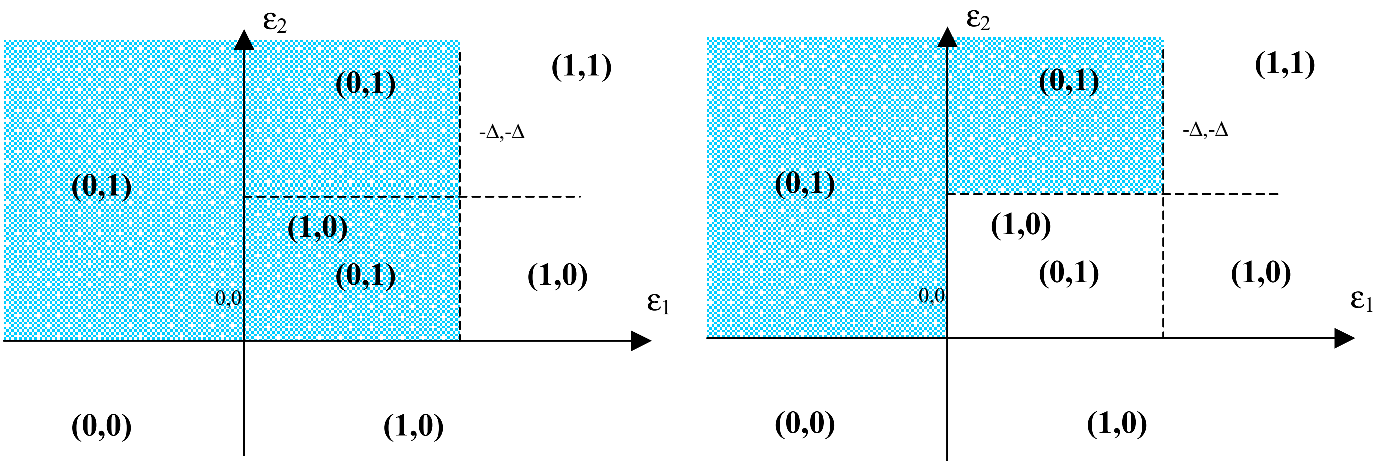
4.1 A Bivariate 2×2 game example:

Consider the 2×2 game studied in Section 3.1:

$$y_i = 1[\Delta_i y_{3-i} + \epsilon_i \geq 0] \tag{3}$$

where we omit the observable regressors for simplicity. In the case where $\Delta_1 < 0$ and $\Delta_2 < 0$ the model (see figure 1) does not predict a unique pure strategy equilibrium when (ϵ_1, ϵ_2) falls in the middle square. Here, either firm 1 is a monopolist or firm 2 is a monopolist. As we

Figure 3: Upper and Lower probability Bounds on the $\Pr(0,1)$:



The shaded area in the graph on the right hand side represents the region for (ϵ_1, ϵ_2) that would predict the outcome $(0,1)$ uniquely. The shaded region in the graph on the left hand side represents the region where $(0,1)$ would be predicted if we *always* select (0,1) to be the equilibrium in the region of multiplicity. The probability of the epsilons falling in the respective regions provide an upper and a lower bound on the probability of observing (1,0).

can see from figure 3, the model predicts *upper and lower* probabilities (as opposed to exact) on the probability of the (1,0) and (0,1) outcomes. Hence, one can then calculate explicitly the bounds on the choice probability. In particular, let the probability of the middle square in figure 1 be $H(\boldsymbol{\theta}) = H(\Delta_1, \Delta_2, \Omega)$ where Ω is the variance covariance matrix of (ϵ_1, ϵ_2). Let:

$$H_1(\theta) = \Pr(\epsilon_1 \leq 0; \epsilon_2 \leq -\Delta_2) + \Pr(\epsilon_1 \leq \Delta_1; \epsilon_2 \geq -\Delta_2) \tag{4}$$

$$H(\theta) = \Pr(0 \leq \epsilon_1 \leq -\Delta_1; 0 \leq \epsilon_2 \leq -\Delta_2) \tag{5}$$

The model predicts the following set of equality and *inequality* restrictions on the conditional regressions:

$$\begin{aligned} \Pr((1,1)) &= \Pr(\epsilon_1 \geq -\Delta_1, \epsilon_2 \geq -\Delta_2) \\ \Pr((0,0)) &= Pr(\epsilon_1 \leq 0; \epsilon_2 \leq 0) \\ H_1(\theta) &\leq \Pr(1,0) \leq H_1(\theta) + H(\theta) \end{aligned} \tag{6}$$

The objective of the new method is to minimize the distance between the observed empirical choice probabilities and the predicted choice probabilities *regions* defined in 6. [12]

4.2 Inference in Discrete Games: the General Case

We base our inferential approach in this paper directly on the inequality restrictions that the model provides (for example those are inequalities (6) in the 2×2 game above). The identified feature of the model is the set of parameter values that satisfy these inequality restrictions. Without loss of generality, we consider a k-player binary game where the strategy of player i is $y_i = 1$ or 0 depending on whether the i's utility crosses a threshold. For player i, this can be written as

$$y_i \quad = \quad \Pi_i(\mathbf{x_i}, \mathbf{y_{-i}}, \theta, \epsilon_i) \geq 0 \tag{7}$$

where $\mathbf{y_{-i}}$ is the $k-1$ binary zero-one vector of other players' strategies, x_i is a vector in R^d of regressors, and ϵ_i is the part of player i's utility that is unobserved to the analyst. Here, for simplicity, we do not have market specific subscripts. We assume in this paper that the observed part of utility $\Pi()$ is known up to the finite dimensional parameter θ and that the unobserved part of the profit function (the ϵ's) is independent of x. The following assumptions summarizes the model.

Assumption 2 *We observe a random iid sample of observations* $(\mathbf{y_i}, \mathbf{x_i})$ *for* $i = 1, \ldots, N$. *Let the model defined in (7) hold. Moreover, let* $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_k)$ *be a mean zero random variable, independent of* $\mathbf{x} \in \mathrm{R}^d$, *and has a known (up to a finite dimensional parameter* Ω*) distribution* F_Ω *that is absolutely continuous on* R^k. *The parameter space* Θ *is a compact subset of* R^l.

Given the decision rule in (7) and assumption 2, the model provides inequality restrictions on regressions:

$$\mathbf{H}_1(x, \theta) \leq \Pr(\mathbf{y}|x) \leq \mathbf{H}_2(x, \theta) \tag{8}$$

where $\Pr(\mathbf{y}|x)$ is a 2^k vector of of choice probabilities that can be consistently estimated using the data. The $\mathbf{H}$'s are functions of θ and the distribution function F_Ω. For example, these functions were derived analytically in the 2×2 game above. In general games, it is not possible to derive these functions since obtaining these functions entails *solving* for the equilibria of the game, a task that can be very complicated. We provide below a simulation procedure that can be used to obtain an estimate of these functions for a given x and a given parameter value.

[12] Identification in this 2×2 game is studied in Tamer (2002) where identification in the case where $\Delta_1 \times \Delta_2 < 0$ is also studied. This is a case where the game admits, for some values of ϵ's, no pure strategy equilibria (only equilibria in mixed strategies).

4.2.1 Identification

The approach we take to identification in this paper is as follows. The class of economic models we study provides the inequality restrictions in (8) above. Heuristically, the identified *set* is the set of parameter values that obey these restrictions for all x almost everywhere and represents the set of economic models that is consistent with the empirical evidence. Next we formally define the identified set.

Definition 1 *Let Θ_I be such that*

$$\Theta_I = \{\theta \in \Theta \quad s.t.\ inequalities\ (8)\ are\ satisfied\ at\ \theta \quad \forall \mathbf{x} \quad a.s.\} \tag{9}$$

We say that Θ_I is the identified set of interest.

The definition above can be generalized naturally to games with many players and/or large strategy space. In those games, the set θ_I is not a singleton and it is hard to characterize it since the inequality restrictions are nonlinear. In some games, the richer the support of $\mathbf{x}$ the "smaller" the set Θ_I is. For example, in a bivariate game with binary strategies, Tamer (2002) showed that the set Θ_I shrinks to a singleton (point identification) with enough variation in the regressors

4.2.2 A Minimum Distance Sharp Estimator of Θ_I

In this section, we provide a modified minimum distance estimator that uses the inequality restrictions in (8) to estimate the set Θ_I. This estimator only requires that one is able to obtain (or solve for) explicitly the functions $\mathbf{H}_1$ and $\mathbf{H}_2$.

A cross- section of markets allow us to identify the conditional choice probabilities. Heuristically, for a given parameter value, the estimator is based on minimizing the distance between this vector of choice probabilities and the set of predicted probabilities. On the set Θ_I, this distance is minimized. The estimator is a *sharp* two step minimum distance estimator: in the first step we estimate non-parametrically the conditional choice probabilities, and in the second stage we obtain an estimate of the set Θ_I. This sharp set contains all the possible parameter values that are consistent with the set of economic models that obey our fundamental assumption.

First, we define the probability set $G(\mathbf{t};\mathbf{x})$ as the set of all *feasible* predicted probability distributionsd for a given value of $\mathbf{t} \in \Theta$ and $\mathbf{x}$.

Definition 2 *Let $\boldsymbol{t} \in \Theta$ and $\mathbf{x} \in \mathrm{R}^l$. Let the set of predicted choice probabilities be defined as*

$$G(\mathbf{t};\boldsymbol{x}) = \left\{ \mathbf{g} = (g_1, \ldots, g_{2^k}) \in [0,1]^{2^k};\ \sum_{i=1}^{2^k} g_i = 1 \Big| \ \mathbf{H_1}(\mathbf{x},\mathbf{t}) \le \mathbf{g} \le \mathbf{H_2}(\mathbf{x},\mathbf{t}) \right\}$$

This set contains all the feasible choice probabilities that are predicted by the model for a given choice of the parameter vector $\mathbf{t}$. For all $\mathbf{x}$ and $\mathbf{t}$ this set is nonempty. Moreover, in games with no multiple equilibria, this set shrinks to a singleton (i.e. the case where $G = \{\mathbf{g} = \mathbf{H_1} = \mathbf{H_2}\}$). The estimator of the identified set Θ_I is based on the following proposition.

Proposition 3 *Let $\boldsymbol{\theta} \in \Theta$. Define the following function*

$$Q(\boldsymbol{\theta}) = \int d\left[\Pr(\mathbf{y}|\mathbf{x}), G(\boldsymbol{\theta}; \mathbf{x})\right] dF_{\mathbf{x}} \tag{10}$$

where $d[\mathbf{p}, G]$ is an appropriate distance function between the vector p and the set G such that $d[\mathbf{p}, G] = 0$ if and only if $\mathbf{p} \in G$, otherwise $d[\mathbf{p}, G] > 0$. We have that for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $Q(\mathbf{b}) \geq 0$. Moreover, $Q(\mathbf{b}) = 0$ if and only if $\boldsymbol{\theta} \in \Theta_I$.

One can for example use the following distance function d in the above

$$d\left(\Pr(\mathbf{y}|\mathbf{x}), G(\mathbf{t}; \mathbf{x})\right) = \min_{\mathbf{g} \in G(\mathbf{t};\mathbf{x})} \|\Pr(\mathbf{y}|\mathbf{x}) - \mathbf{g}\|$$

First, for a given $\mathbf{t}$, the inequalities in (6) above provide a feasible set of choice probabilities that are predicted by the model. Usually, if this set is a singleton, one then minimizes the distance between the predicted probability and the observed one (which is $\Pr(\mathbf{y}|\mathbf{x})$ and hence "observed" here is used to mean "can be consistently estimated."). This is the minimum distance method. In the case where for a given $\boldsymbol{\theta}$ the model provides a set of predicted probabilities $G(\boldsymbol{\theta}; \mathbf{x})$, the estimator minimizes the distance between the observed choice probability and this set. This distance is minimized on the set Θ_I where for all x, $d(\Pr(\mathbf{y}|\mathbf{x}), \mathbf{G}(\mathbf{x}, \boldsymbol{\theta}) = 0)$ for all $\boldsymbol{\theta} \in \Theta_I$.
To estimate the set Θ_I, we minimize a feasible sample analog of $Q(\mathbf{t})$. To do that, we first replace $\Pr(\mathbf{y}|\mathbf{x})$ by a consistent nonparametric estimator $P_N(\mathbf{x})$. Then, $\widehat{\Theta}_I$ the estimate of the set Θ_I is defined as

$$\widehat{\Theta}_I = \left\{ t \in \Theta \mid Q_n(t) \leq \min_{t \in \Theta} Q_n(t) + \epsilon_N \right\} \tag{11}$$

where ϵ_N is a nonnegative real number that goes to zero as $N \to \infty$, and

$$\begin{aligned} Q_n(t) &= \frac{1}{N} \sum_{i=1}^{N} d_n[P_n(\mathbf{x}), G(\mathbf{t}, \mathbf{x}_i)] \\ &= \frac{1}{N} \sum_{i=1}^{N} \min_{\mathbf{g} \in G(\mathbf{t};\mathbf{x}_i)} \|P_N(\mathbf{x}_i) - \mathbf{g}\| \end{aligned}$$

Notice that $\widehat{\Theta}_I$ is the set of parameter values that are ϵ away from minimizing the objective function. This guarantees that the distance between the set V_n and V goes to zero as the sample size increases. To guarantee this, we need to drive ϵ to zero at a slow enough rate (the exact rate is in Theorem 1) . The results are stated in the next theorem.

Theorem 4 (Consistency) *Let assumptions 1 and 2 above hold. Let $N \to \infty$. Suppose that $P_N(x) \to_{a.s.} P(x)$ uniformly in x. Moreover, let*

$$\sup_t \|Q_n(t) - Q(t)\| = o_p(\epsilon_n) = o_p(n^{-\alpha}) \tag{12}$$

where $\alpha > 0$. Then

$$\rho(\widehat{\Theta}_I, \Theta_I) \equiv \sup_{t \in \widehat{\Theta}_I} \inf_{t' \in \Theta_I} |t - t'| \to_{a.s.} 0$$

$$\rho(\Theta_I, \widehat{\Theta}_I) \equiv \sup_{t \in \Theta_I} \inf_{t' \in \widehat{\Theta}_I} |t - t'| \to_{a.s.} 0$$

Under standard uniform convergence conditions and consistency of the first step nonparametric estimation of the conditional choice probabilities, the theorem asserts that the Hausdorff distance between the set Θ_I and the estimated set $\widehat{\Theta}_I$ converges in probability to zero. The set $\widehat{\Theta}_I$ is constructed by taking all the parameter values that are within ϵ_n from minimizing the sample objective function Q_n where as we see above $\epsilon_n = O(n^{-\alpha})$ and hence $\widehat{\Theta}_I$ is a particular level set of the objective function. This theorem is similar to theorem 5 in Manski and Tamer (2002).

4.3 Simulating H_1 and H_2

The game has 2^k potential outcomes $\mathbf{y}_l = (y_{1l}, ..., y_{kl})$ where $y_{jl} \in \{0, 1\}$, for $l = 1, ..., 2^k$, and $j = 1, ..., k$. The observed data in this model identifies the 2^k dimensional choice probability vector $\Pr(\mathbf{y}|\mathbf{x}) = [\Pr(\mathbf{y_1}|\mathbf{x}), ..., \Pr(\mathbf{y_{2^k}}|\mathbf{x})]$. The model provides the following restrictions

$$\mathbf{H}_1(x, \theta) = \begin{bmatrix} H_1^1(\mathbf{x}, \boldsymbol{\theta}) \\ \vdots \\ H_1^{2^k}(\mathbf{x}, \boldsymbol{\theta}) \end{bmatrix} \leq \begin{bmatrix} \Pr(\mathbf{y}_1|\mathbf{x}) \\ \vdots \\ \Pr(\mathbf{y}_{2^k}|\mathbf{x}) \end{bmatrix} \leq \begin{bmatrix} H_2^1(\mathbf{x}, \boldsymbol{\theta}) \\ \vdots \\ H_2^{2^k}(\mathbf{x}, \boldsymbol{\theta}) \end{bmatrix} = \mathbf{H}_2(x, \theta) \tag{13}$$

where the inequalities are to be interpreted element by element. The presence of simultaneity in the model makes the derivation of the H_1 and H_2 complicated. One basically needs to "solve" for the equilibrium of the game for every $\boldsymbol{\theta}$ and $\mathbf{x}$. This is a difficult task in general especially if the game involves many players and/or a large strategy space. Moreover, these two functions depend on whether we allow for mixed strategy equilibria to occur. For simplicity, we restrict ourselves in this paper to pure strategy equilibria and leave the topic of mixed strategies to future papers. Essentially, the problem is that with many players it is hard to obtain the equivalent of the bounds we obtained in (6) above for the 2×2 game.

To deal with this, we propose a procedure to simulate H_1 and H_2. Simulation procedures in discrete choice models are well known in econometrics. The simulation estimators are usually used to approximate choice probability in multivariate probit models. These probabilities are expectations of indicator functions over the joint multivariate normal distribution of the unobserved random variables. The simulation procedure to obtain H_1 and H_2 is summarized below.

Procedure to obtain predicted probability bounds for a given t and x

Set $\widehat{\mathbf{H}}_1(\mathbf{x},\mathbf{t}) = \widehat{\mathbf{H}}_2(\mathbf{x},\mathbf{t}) = \mathbf{0}$.

- **Step 1:**
  Simulate a random draw from the joint distribution of $(\epsilon_1^r, \ldots, \epsilon_k^r)$ with covariance matrix specified in $\mathbf{t}$.

- **Step 2:**
  Using the profits functions in 7 above, calculate

  $$\mathbf{\Pi}(\mathbf{y_l}, \mathbf{x}, \mathbf{t}, \boldsymbol{\epsilon}^{\mathbf{r}}) = [\Pi_1(\mathbf{y_{-1}}, \mathbf{x}, \mathbf{t}, \epsilon_1^r), ..., \Pi_k(\mathbf{y_{-k}}, \mathbf{x}, \mathbf{t}, \epsilon_k^r)]$$

  for all $l = 1, ..., 2^k$.

- **Step 3:**
  This step finds the equilibria of the game:

  1. For all $l \in \{1, ..., 2^k\}$ such that $\mathbf{\Pi}(\mathbf{y_l}, \mathbf{x}, \mathbf{t}, \boldsymbol{\epsilon}) \geq \mathbf{0}$, set $\widehat{H}_2^l = \widehat{H}_2^l + 1$.
  2. If there is an $l \in \{1, ..., 2^k\}$ such that $\mathbf{U}(\mathbf{y}_l, \mathbf{x}, \mathbf{t}) \geq 0$ *uniquely*, i.e., there is no $l' \neq l$ such that $\mathbf{U}(\mathbf{y}_{l'}, \mathbf{x}, \mathbf{t}, \epsilon) \geq 0$, then $\widehat{H}_1^l = \widehat{H}_1^l + 1$.
  3. If neither (1) or (2) above holds then for this value of the parameter, the game does not have a pure strategy equilibrium.

Repeat steps 1-3 above R times to obtain the simulation estimators

$$\frac{1}{R}\widehat{\mathbf{H}}_2(\mathbf{x}, t) \text{ and } \frac{1}{R}\widehat{\mathbf{H}}_1(\mathbf{x}, \mathbf{t})$$

Comments on the simulation procedure: First fix x and t. In step 1 above, we simulate a draw from a multivariate random normal. In step 2, we obtain the "profits" for every player i as a function of other players' strategies. If $\mathbf{\Pi}(\mathbf{y}_l, \mathbf{x}, b) \geq \mathbf{0}$ for some $l \in \{1, ..., 2^k\}$, then $\mathbf{y_l}$ is an equilibrium of that game. If this equilibrium is unique, then we add 1 to the lower bound probability for outcome y_l and add 1 for the upper bound probability. If the equilibrium is not unique, then we add a 1 only to the upper bound of each of the multiple equilibria's upper bound probabilities. To illustrate, consider the game in figure 1 above. If (ϵ_1, ϵ_2) lies in the upper right hand side corner, then $(1, 1)$ is the unique equilibrium of the game and $\widehat{H}_1^{(1,1)} = \widehat{H}_1^{(1,1)} + 1$ and $\widehat{H}_2^{(1,1)} = \widehat{H}_2^{(1,1)} + 1$. On the other hand, if (ϵ_1, ϵ_2) lies in the middle square, then $\widehat{H}_2^{(1,0)} = \widehat{H}_2^{(1,0)} + 1$ and $\widehat{H}_2^{(0,1)} = \widehat{H}_2^{(0,1)} + 1$. This way we see that since $(1,1)$ is a unique equilibrium for this game, then $\widehat{H}_2^{(1,1)} = \widehat{H}_1^{(1,1)}$ and the left hand side probability bound is equal to the right hand side. We repeat the above procedure R times to obtain a simulation based estimator of the upper and lower bounds on the outcome of interest. For example, the upper bound on the outcome probability $\Pr(1,...,1|\mathbf{x})$ is

$$\widehat{\mathbf{H}}_2^{2^k}(\mathbf{x},\theta) = \frac{1}{R}\sum_{j=1}^{R} 1\left[\Pi_1(x_1,\theta;\mathbf{y}_{-1}^{2^k},\epsilon_1^j) \geq 0,\ldots,\Pi_{2^k}(x_{2^k},\theta;\mathbf{y}_{-2^k}^{2^k},\epsilon_{2^k}^j) \geq 0\right]$$

where $\mathbf{1}[*]$ is equal to one if the logical condition $*$ is true. One can use methods developed by McFadden (1989) and Pakes and Pollard (1989) to show that $\widehat{\mathbf{H}}_{@}(\mathbf{x},\theta)$ converges almost surely uniformly in θ and x to $\mathbf{H}_2(\mathbf{x},\theta)$ as the number of simulations increases.

4.4 Practical Estimator and Confidence Regions (Preliminary):

In this section, we provide a slight variant of the estimator (10) above that is more easily implementable especially in the case where we have a rich set of regressors x. Consider the following modified objective function

$$Q'(t) = \int \left[\| \left(P(x) - H_1(x,t)\right)_- \| + \left\|\left(P(x) - H_2(x,t)\right)_+\right\|\right] dF_x$$

where $(A)_- = [a_1 1[a_1 \leq 0],\ldots,a_{2^k}1[a_{2^k} \leq 0]]$ and similarly for $(A)_+$ for a 2^k vector A and where $\|.\|$ is the Euclidian norm. It is easy to see that $Q'(t) \geq 0$ for all $t \in \Theta$ and that $Q'(t) = 0$ if and only if $t \in V$, the identified set in definition 1 above. Then, $\widehat{\Theta}'_I$ the estimate of the set Θ_I based on an empirical analog of Q' is defined as

$$\widehat{\Theta}'_I = \left\{t \in \Theta \mid Q'_n(t) \leq \min_{t\in\Theta} Q'_n(t) + \epsilon_N\right\} \tag{14}$$

where ϵ_N is a nonnegative real number that goes to zero as $N \to \infty$, and

$$Q'_n(t) = \frac{1}{N}\sum_{i=1}^{N}\left[\left\|\left(P_N(x_i) - \widehat{H}_1(x_i,t)\right)_-\right\| + \left\|\left(P_N(x_i) - \widehat{H}_2(x_i,t)\right)_+\right\|\right] \tag{15}$$

and

$$\left\|\left(P_N(x_i) - \widehat{H}_1(x_i,t)\right)_-\right\| = \left(P_N^1(x_i) - \widehat{H}_1^1(x_i,t)\right)_-^2 + \ldots + \left(P_N^{2^k}(x_i) - \widehat{H}_1^{2^k}(x_i,t)\right)_-^2$$

We summarize our set consistency result in the next theorem.

Theorem 5 (Consistency) *Let assumptions 1 and 2 above hold. Let* $N \to \infty$. *Suppose that* $P_N(x) \to_{a.s.} P(x)$ *uniformly in* x. *We have that for* $\epsilon_n > 0$ *and* $\epsilon_n \to 0$ *such that*

$$\frac{\sup_t \|Q'_n(t) - Q'(t)\|}{\epsilon_n} \to_p 0 \tag{16}$$

Then

$$\rho(\widehat{\Theta}'_I, \Theta_I) \equiv \sup_{t \in \widehat{\Theta}'_I} \inf_{t' \in \Theta_I} |t - t'| \to_{a.s.} 0$$
$$\rho(\Theta_I, \widehat{\Theta}'_I) \equiv \sup_{t \in \Theta_I} \inf_{t' \in \widehat{\Theta}'_I} |t - t'| \to_{a.s.} 0$$

Here, we will assume that the regressors x have a discrete support, i.e., $x \in \{x_1, \ldots, x_K\}$. This is meant to facilitate obtaining confidence regions for the identified set. In this case,

$$P_N(x = x_k) = \frac{1}{N} \sum_i 1[x_i = x_k]$$

The above method does not deliver a confidence region for the set Θ_I. We describe a method to construct such a region in the appendix. This method is based on recent results on constructing confidence regions for set identified models obtained in Chernozhukov, Hong, and Tamer (2002). Basically, we look for a set $\widehat{\Theta}_I$ that covers the identified set Θ_I with a prespecified probability α as sample size increases:

$$\lim_{n \to \infty} \Pr(\Theta_I \subseteq \widehat{\Theta}_I) = \alpha \tag{17}$$

The confidence regions are appropriately constructed level sets similar to those in (14) where an appropriate 'ϵ_n" is provided that guarantees coverage. This subsampling based procedure provides a cutoff level where the corresponding level set obeys condition (17) above. This subsampling procedure is described in the appendix.

5 The Data

We use the T-100 Domestic Segment Dataset, which contains domestic non-stop segment data by aircraft type and service class for passengers transported, freight and mail transported, available capacity, scheduled departures, departures performed, and aircraft hours. The dataset is publicly available at TransStats, the Intermodal Transportation Database prepared by the Bureau of Transportation Statistics. [13]

The definition of the market is an important issue. Berry (1992) defines a market as the market for air passenger travel between two cities, irrespective of intermediate transfer

[13] We choose the month of May of 1998 because data on gate leases at airports are available from TRB [1999] for that year. We use these data in Ciliberto and Tamer (2003).

points. This approach, however, ignores the fact that some segments of the consumer demand can have strong preferences between airports in the same city. Following Borenstein [1989], we assume that flights to different airports are in separate markets. In one of our specifications our econometric procedure allows for spatial correlation among markets between the same two cities, for example ORDCLE and MDWCLE. This will result in larger parameter bounds. We restrict the analysis to forty major airports, which are listed in the Appendix. Furthermore, our dataset permits to define the market as as the non-stop market (the route, or spoke) between two airports. This definition of the market takes into account the fact that the critical element of the entry decision is whether to commit assets like planes and lease facilities at airports. This decision concerns the non-stop service between two airports.

In our data set there are 23 carriers. Table 1 presents the airlines' characteristics in terms of their number of markets served (a measure of the size of the network). One important

Table 1: **Carrier Characteristics**

Carrier Code	Name Carrier (Network Size)	Number Markets
HA	Hawaiian	1
YX	Midwest Express	2
XJ	Mesaba	4
FF	Tower	4
KP	Kiwi	4
NK	Spirit	4
QQ	Reno	6
EV	Atlantic Southeast	7
9N	Trans States	8
NJ	Vanguard	9
TZ	ATA	9
F9	Frontier	10
FL	AirTran	14
AS	Alaska	16
TW	TWA	48
HP	America West	56
WN	Southwest	69
CO	Continental	91
NW	Northwest	92
US	USAir	106
AA	American	112
UA	United	125
DL	Delta	151

issue is how to treat regional airlines that operate through code-sharing with national airlines. We assume that the decision to serve a spoke is made by the regional carrier, which then signs code-share agreements with the national airlines. As long as the regional airline is independently owned, we treat it separately from the national airline. [14]

There are 766 markets in the dataset, of which 200 are not served by any airline. The maximum number of airlines serving a non-stop route is 8.[15] We have classified markets by the log of the product of the populations of the connected cities. Population determines the potential demand for air travel between two cities (Berry (1992)). The relevant issue in presenting the descriptive statistics is whether market size alone determines market structure (Bresnahan and Reiss (1990)).

The first row of table 2 shows the average number of carriers in each market by potential demand size, which is very similar across city pairs of different population size. Looking at the standard deviation in the first rows one notices that there is some variation in the number of firms across markets and this variation is only slightly decreasing with market size.

Table 2: **Market Characteristics**

	Served				
	All Markets	Large Markets	Medium-Large Markets	Medium Small Markets	Small Markets
Variable	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean(s.d.)	Mean (s.d.)
# of Carriers	1.24 (1.03)	1.38 (1.16)	1.26 (1.08)	1.17 (0.94)	1.15 (0.89)
Log Distance	6.89 (0.72)	6.87 (0.87)	6.95 (0.68)	6.94 (0.61)	6.79 (0.69)
Log Population	30.34 (1.08)	31.77 (0.53)	30.68 (0.21)	29.94 (0.21)	29.00 (0.43)
N	766	191	190	193	192

The secopnd row presents the log distance in nonstop miles, which does not differ across markets. In fact, the average of the log of the product of the populations of the connected cities does not vary much across the four quartiles. In short, demographic and technological characteristics do not seem to vary much among markets.

To further investigate the variation in the number of firms across markets Table 3 provides the distribution of the number of firms by potential demand size. Table 3 shows that the variation in the number of firms across markets cannot be explained by the size of the potential demand alone.

[14]For example, Atlantic Southeast Airlines joined the program Delta Connection on May 1, 1984. In 1986, Delta bought some shares of ASA. Only in 1999, however, Delta acquired ASA. Trans State Airlines provides service today for AA and USAir. In the past TSA provided service also for NW, DL, Alaska, TWA. We code TSA and ASA as independent firms. Since this decision affects less than 2 percent of the markets, the results are unchanged if we code the three regional carriers as part of the national airlines. See Appendix for a discussion on small regional and independent airlines.

[15]The route is: LASLAX. The airlines are: American, Delta, Hawaiian, America West, Reno Air, United, Southwest.

Table 3: **Distribution of the Number of Carriers by Market Size**

Number of Firms	Large	Medium Large	Medium Small	Small	Total
0	49	50	55	46	200
1	60	66	67	88	281
2	54	57	56	43	210
3	20	13	14	14	61
4	4	3	1	1	9
5	4	0	0	0	4
8	0	1	0	0	1
Total	191	190	193	192	766

6 Empirical Results

The main objective of this paper is to provide a method to estimate a system of simultaneous discrete choice equations where, for a given set of parameters, the model does not predict a unique distribution for the outcomes. It seems natural to compare the results that we find using the method developed in Section 4 with the results that we would find when the firms' decisions are assumed to be strategically independent. This is analogous to the case when we compare the results from the estimation of a system of simultaneous equations to the results that we would find if we were to run each equation separately. The next section provides simple probit results that ignore endogeneity.

6.1 Simple Probits

In order to show how striking the difference in the results can be when we allow for strategic dependence, we allow the probability of a firm being in the market to depend on the identity of the other firms in the market. In particular, we (separately) estimate the following regressions:

$$y_{im} = 1\left[\alpha_0 + \alpha_X \ln X_m + \sum_{j\in\tilde{I}_m\setminus\{i\}} \delta_j y_{jm} + \epsilon_{im} \geq 0\right].$$

where $\tilde{I} = \{\text{AA,DL,WN,OL,OS}\}$. This includes American, Delta, and Southwest. It also includes an index OL which is equal to 1 if there is another large carrier in the market after counting American, Delta, and Southwest. For example, $OL = 1$ if United is in the market. To reduce the computational burden we use OL and not a categorical variable for each firm. This should not be a big problem since only 11.10% of the markets have more than one of the other large firms. We also use an index OS to indicate whether there is at least one small low-cost carrier in the market. In this case also, only about 0.78% of the markets have more than one small low cost carrier. Table 4 presents the results of the estimation.[16]

Table 4: **Simple Probit Results**

LHS / RHS	AA	DL	WN	Another Large	At Least One Small
AA	x	0.64 (0.15)	-0.85 (0.33)	0.43 (0.14)	0.01 (0.18)
DL	0.62 (0.14)	x	-0.14 (0.18)	-0.47 (0.12)	0.43 (0.14)
WN	-0.91 (0.35)	-0.05 (0.19)	x	-0.35 (0.17)	0.02 (0.20)
Another Large	0.36 (0.13)	-0.41 (0.11)	-0.35 (0.15)	x	-0.01 (0.13)
At Least One Small	-0.02 (0.18)	0.47 (0.15)	0.03 (0.20)	-0.04 (0.15)	x
LogPop	0.34 (0.06)	-0.15 (0.05)	-0.16 (0.07)	0.02 (0.04)	0.18 (0.06)
Log Distance	-0.06 (0.06)	-0.17 (0.54)	-0.57 (0.08)	-0.31 (0.05)	-0.40 (0.06)
Constant	-2.13 (0.25)	-0.27 (0.20)	0.50 (0.28)	0.96 (0.17)	-0.86 (0.24)
Nobs	766	766	766	766	766
Log-Likelihood	-280.67	-351.02	-191.81	-491.06	-247.34

Every column provides the equation by equation probit estimates. So the second column provides estimates from a probit regression of AA on the rest of the variables in the first column (except AA).

Each column presents the results of a separate probit regression for each carrier. For example, the first column presents the results of a probit regression where we study the probability of observing AA in a market conditional on observing the other airlines, and on the other control variables. The second column estimates the probability of observingDL in a market. We find that at the mean values, having DL in a market increase the probability of observing AA by 14.8 percent.[17] The presence of one small firm (OS) decreases the probability of observing AA in the market by less than 1 percent.

An eyeball study of the table Probit suggests that there is perfect symmetry in the estimation results, in the sense that if WN decreases the probability of observing AA in a market, then AA decreases the probability of observing WN in the market. The reason for this finding is that the probit estimation only reports simple correlation patters between the airline entry, but does not account for the fact that the entry decisions are strategically related.

These results are puzzling. We would expect that the presence of a large carrier decreases the probability of observing another large carrier in the market. Next, we investigate whether the unintuitive sign of the coefficients is explained by the fact that we do not control for the strategic dependance among carriers.

[16]We find almost identical results if we consider a dataset of the largest 50 metropolitan statistical areas.

[17]A table of the marginal effects of the variables in the probit estimation is available on request from the authors. We present the coefficient estimates for sake of comparison with the results in the next section.

6.2 Main Results

In this section, we estimate systems of discrete response models involving many decision makers. We start with a specification where we allow for example Southwest to have a different effect on American than on Delta.

6.2.1 Specification with heterogeneous impacts

The first model we estimate is:

$$\text{(Gen. Spec.)} \begin{cases} y_{AA,m} = 1\left[\alpha_{AA} + \alpha_{AA}\ln X_m + \sum_{j\in\tilde{I}_m/\{AA\}} \delta_j^{AA} y_{jm} + \epsilon_{AA,m} \geq 0\right] \\ y_{DL,m} = 1\left[\alpha_{DL} + \alpha_{DL}\ln X_m + \sum_{j\in\tilde{I}_m/\{DL\}} \delta_j^{DL} y_{jm} + \epsilon_{DL,m} \geq 0\right] \\ y_{WN,m} = 1\left[\alpha_{WN} + \alpha_{WN}\ln X_m + \sum_{j\in\tilde{I}_m/\{WN\}} \delta_j^{WN} y_{jm} + \epsilon_{WN,m} \geq 0\right] \\ y_{OL,m} = 1\left[\alpha_{OL} + \alpha_{OL}\ln X_m + \sum_{j\in\tilde{I}_m/\{OL\}} \delta_j y_{jm}^{OL} + \epsilon_{OL,m} \geq 0\right] \\ y_{OS,m} = 1\left[\alpha_{OS} + \alpha_{OS}\ln X_m + \sum_{j\in\tilde{I}_m/\{OS\}} \delta_j^{OS} y_{jm} + \epsilon_{OS,m} \geq 0\right]. \end{cases}$$

In this model the effect of firms on each other are allowed to be different. For example, we want to let American have a different effect on Delta than it has on a low cost carrier. Notice that the dummies introduce a measure of heterogeneity, as they capture how each firm affects the entry decision of the other five firms. Table 5 resents our results from the estimator above.[18]

We can compare this table to the probit results of table 4. The second column in Table 5, denoted AA, reports the set estimates for the first equation in the general specification (See the footnote to Table 5 for more on the estimates). For example, δ_{AA}, which captures the effect of Delta on American being in the market, is in $[1.817, 13.636]$. The other bounds should be interpreted similarly.

The row "Correct Predictions" of Table 5 reports the percentage of outcomes that were observed in the data and that the we would predict them as possible equilibria using the parameter vector θ where the distance function $Q_n'(.)$ is minimized. This means that 28.12 percent of the observed outcomes were one of the equilibria predicted by the estimated model in each market. The last row reports the value of the distance function at the parameter values where it is minimized.

First, Table 5 shows that firms benefit from the presence of other firms. We estimate several of the δ's to be positive. For example, the presence of Delta in a market does increase, on average, the probability of observing American as well.

Second, Table 5 shows that there is still considerable symmetry in the table and that the signs of the coefficients are generally similar. There are some cases where simmetry does not

[18]The optimization was done using simulated annealing and its adaptive version. This is helpful since genetic algorithms, although slow, scan the surface of the function and thus allows us to obtain the level sets needed to construct our set estimates.

Table 5: **The General Specification**

	AA	DL	WN	Another Large	At Least One Small
AA	x	[1.817,13.636]	[-19.999,-0.182]	[1.203,6.518]	[-0.866,0.360]
DL	[1.559,2.310]	x	[0.987,2.653]	[-7.531,-5.755]	[2.687,3.430]
WN	[-19.972,-0.885]	[12.247,14.056]	x	[5.117,6.873]	[-19.894,-0.653]
Another Large	[1.662,1.930]	[-14.046,-13.198]	[1.538,1.945]	x	[2.018,2.514]
At Least One Small	[-0.672,0.082]	[13.651,14.573]	[-19.932,-1.653]	[5.898,7.744]	x
Log Pop	[0.178,0.268]	[-0.297,-0.119]	[-0.262,-0.080]	[0.052,0.121]	[-0.056,0.125]
Log Distance	[-0.123,0.022]	[-0.209,-0.111]	[-0.586,-0.269]	[-0.288,-0.231]	[-0.389,-0.181]
Constant	[-3.460,-3.179]	[-0.305,0.018]	[-2.174,-1.699]	[0.552,0.727]	[-3.461,-2.974]
Nobs	766				
Correct predictions	0.2812				
Function Value	58.363				

This table provides estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

hold: the effect of Delta on the entry decision of Another Large is much weaker than the reverse. Also the effect of Another Large's presence on the entry decision of Southwest is much weaker than the effect of Southwest entry on Another Large's decision to enter. These results suggest that firms' behavior is not symmetric.

Third, Table 5 shows that the probit estimates are biased. For example, in Table 4 we find that the coefficient of the effect of Delta on American is equal to 0.62 but Table 5 shows that the coefficient is included in $[1.559, 2.310]$. Also, we find that the coefficient of Another Large in the regression for Southwest (Column 4, Table 4) is -0.35 while in Table 6 we get $[1.538, 1.945]$, confirming that simple probit regressions inconsistently estimate the parameters of the model.

Finally, Table 5 shows that American and Delta display different strategic behaviours. Entry of Southwest has a positive effect on Delta's decision to enter but a negative effect on American's decision. On the contrary, the presence of Another Large carrier decreases the probability of observing Delta, but increases the probability of American being in the market. Small carriers do not affect American's decision to enter, and viceversa. But entry by small carriers positively affects the decision of Delta to enter into the market. We conclude that American and Delta display different strategic behaviors.[19]

To examine the existence of multiple equilibria, we simulate results where for every market (the only market specific variables in this specification are logdistance and logpop) in the following way. We draw 1000 times from the joint distribution of the errors and calculate the number of multiple pure strategy equilibria observed. Doing this, we get a lot of multiple

[19]We further investigate the airlines' different strategic behaviors in Ciliberto and Tamer (2003).

equilibria under this specification. In fact in 67.85% of the markets, there are multiple equilibria and in 57.07% of these markets there are multiple equilibria in the number of firms. This critical result suggests that either airlines behave differently with each other, or that there exist externalities among airlines which depend on their identities.

The estimated coefficient bounds provide information on the effect of airlines on each other, but from an economic standpoint we are mostly interested in estimating the probability of each outcome.

To estimate the probability that an equilibrium occurs, we proceed as follows. We take the parameter vector θ where the distance function $Q_n'(.)$ is minimized. We then compute the lower and upper bound of the probabilities on the 32 possible equilibria for each of 1000 random draws from the joint distribution of the errors. We finally take the average of these 1000 lower and upper bounds as estimates of the actual lower and upper bound of each possible equilibrium. Table 6 presents the results in terms of probabilities of observing different equilibria for *all* markets.[20]

Table 6 provides information on the probability of industry configurations. For example, the estimated probability of observing American, Delta, Southwest, another large carrier, and a small low-cost carrier (this equivalent to the first row $(1,1,1,1,1)$ in the table) is null.

We can use Table 6 to understand which market configurations are equilibria and which are not. We can thus infer whether a combination of firms, such as American and Southwest, can be in equilibrium. This has critical implications both from a positive standpoint, as we can predict whether a firm can successfully enter a market, and from a normative standpoint, as we can assess whether American behaves differently towards low cost carriers than towards Southwest Airline.

Only 12 configurations can be equilibria. Among the equilibria, there are some that have low probability of occurring. For example, it is very unlikely to observe Delta, Southwest and a Another Large in the same market. Southwest is only in markets where it is a duopolist with another large airline. Small airlines are sometimes in markets with a large airline, but most likely they are in markets that are highly competitive. American is the only large airline that is never a monopolist. Finally, there is a large probability that markets are not served by any airline.

The last column reports the empirical probabilities of market structures. For example, the market structure with no firms serving a market occurs 26.11 percent of the times in the dataset. We estimate the probability of observing a market with no firms to be included

[20] Using probit regressions to estimate the model (we know this is inconsistent), we might get parameter estimates that are close to the bounds obtained using our estimator. For prediction purposes though, it is not clear, without using our intuition, how one would predict market structure since the model (probit) is incomplete. One can use the (inconsistent) probit estimates to derive upper and lower probabilities on the outcomes of interest. We compared these "probit predictions" to our own in the general specification (Gen. Spec.) above and found that the probit predictions, not reported here, were very different from those presented in Table 6.

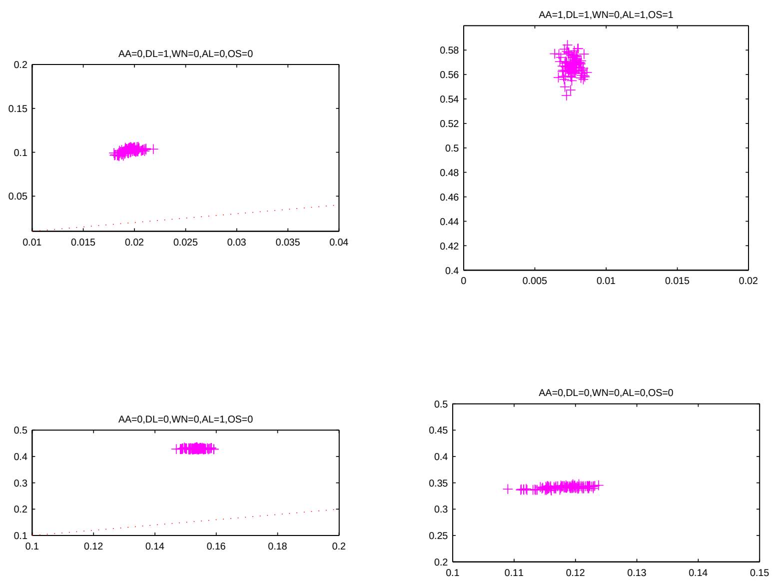
Table 6: **Predicted Outcome Probabilities for General Specification**

Equilibrium											
AA	DL	WN	OL	OS	Lower Bound IID	Upper Bound IID	LB Perfect Corr	UB Perfect Corr	LB Corr	UB Corr	Empirical Probability
1	1	1	1	1	0	0	0	0	0.0000	0.0363	0.0013
1	1	1	1	0	0	0	0	0	0	0.1508	0
1	1	1	0	1	0	0	0	0	0.0000	0.0000	0
1	1	1	0	0	0	0	0	0	0	0.0205	0
1	1	0	1	1	0.0076	0.5679	0	0.3328	0	0.3191	0.0091
1	1	0	1	0	0	0	0	0	0	0	0.0170
1	1	0	0	1	0	0	0	0	0.0000	0.2918	0.0013
1	1	0	0	0	0.0001	0.1275	0	0.2116	0.0000	0.3132	0.0196
1	0	1	1	1	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	0	0	0.0064	0
1	0	1	0	1	0	0	0	0	0	0	0.0013
1	0	1	0	0	0	0	0	0	0	0.0013	0
1	0	0	1	1	0	0	0	0	0	0.0366	0.0026
1	0	0	1	0	0.0003	0.1352	0	0.1506	0	0.1863	0.0718
1	0	0	0	1	0	0	0	0	0	0	0.0078
1	0	0	0	0	0	0	0	0	0	0.0523	0.0144
0	1	1	1	1	0	0	0	0.0069	0.0000	0.1277	0.0013
0	1	1	1	0	0.0002	0.0204	0	0.0279	0.0000	0.2600	0.0065
0	1	1	0	1	0	0	0	0	0	0	0
0	1	1	0	0	0.0020	0.0824	0	0.0710	0	0.1270	0.0117
0	1	0	1	1	0.0033	0.0511	0	0.6672	0	0.1283	0.0196
0	1	0	1	0	0	0	0	0	0	0	0.0379
0	1	0	0	1	0.0028	0.0481	0	0	0	0.3661	0.0091
0	1	0	0	0	0.0190	0.1005	0	0.3641	0	0.2019	0.0627
0	0	1	1	1	0	0	0	0	0	0	0.0039
0	0	1	1	0	0.0067	0.0654	0	0.1157	0	0.1283	0.0339
0	0	1	0	1	0	0	0	0	0	0	0.0078
0	0	1	0	0	0	0	0	0	0	0	0.0222
0	0	0	1	1	0.0080	0.0428	0	0	0	0.1646	0.0366
0	0	0	1	0	0.1525	0.4284	0	0.4544	0.0292	0.5013	0.3211
0	0	0	0	1	0	0	0	0	0	0	0.0183
0	0	0	0	0	0.1190	0.3432	0	0.3512	0.0088	0.4272	0.2611

in $[0.1190, 0.3432]$. Hence, the observed probability is included in the estimated probability bounds. The same consideration is valid for almost all of the equilibria that have non-null probability.

Table 6 provides the results at the parameter value in the estimated region at which the sample objective function is minimized. Next, we check whether the results in that table are robust to using other parameters in the estimated region. Figure 4 shows the probability bounds for four market outcomes when we compute the probability bounds at 100 randomly chosen parameter values included in the coefficient bounds presented in Table 5. The bounds are shown to be very stable for various choices of coefficient vectors.

Figure 4: Bounds on Equilibria for Four Equilibria in the Baseline Specification



These figures provide bound estimates on the probability of the (0,1,0,0,0), (1,1,0,1,1), (0,0,0,1,0), and (0,0,0,0,0)outcomes. Every cross is a bound: the x-axis is the *lower* bound and the y-axis is the *upper* bound. For every parameter in the set estimate (the level set of the objective function), we plot a cross which represents a *bound* on the probability of observing the outcome. So, every graph is a map between the set estimates of the parameters and upper and lower bounds on the probabilities of the four outcomes.

6.3 Perfectly Correlated Unobservable Profits

So far the unobservable part of the profits has been assumed to be independent and identically distributed across firms and markets. This is equivalent to assume, for example, that the unobservable fixed costs of American between two airports are independent of the fixed costs of Delta between the same two airports. This is clearly very restrictive since both American and Delta use the same two airports and their fixed costs are likely to be correlated. As a first step we consider the case when the unobservable airline profits are perfectly correlated across firms within each market. This is equivalent to assume that the unobservable profits of American and Delta, for example, are identical in each market. This specification is specially useful to compare our methodology with Bresnahan and Reiss (1990) "$Model$ 3". Bresnahan and Reiss's "$Model$ 3" is the specification where the unobservable profits are perfectly correlated among firms in a market but the error faced by a monopolist is different than the error faced by a duopolist. In this part of the paper, we also assume that the unobservable profits are perfectly correlated and that the unobservable profits might change with the identity and number of firms. In particular, we continue to include δ's in our regressions. δ's exactly capture the unobservable part of the profits that change with the identity and number of firms.

Table 7 presents the results of this estimation:

Table 7: **Perfectly Correlated Unobservables**

	AA	DL	WN	Another Large	At Least One Small
AA	x	[-1.335,12.341]	[-19.947,-1.094]	[-1.185,5.477]	[-10.232,2.150]
DL	[1.681,3.011]	x	[1.075,2.709]	[-9.499,-3.574]	[3.171,16.593]
WN	[-19.994,-2.611]	[10.411,12.746]	x	[1.370,7.228]	[-19.872,-3.946]
Another Large	[2.068,2.250]	[-14.854,-12.657]	[1.437,1.877]	x	[2.360,12.951]
At Least One Small	[-4.146,0.297]	[13.874,19.982]	[-19.976,-0.237]	[5.951,19.974]	x
Log Pop	[0.193,0.235]	[-0.220,0.119]	[-0.164,0.046]	[-0.050,0.040]	[-1.481,0.980]
Log Distance	[-0.176,-0.088]	[-0.513,-0.109]	[-0.532,-0.334]	[-0.304,-0.254]	[-3.281,0.180]
Constant	[-3.549,-3.381]	[0.575,1.556]	[-1.812,-1.445]	[1.040,1.131]	[-14.430,-3.645]
Nobs	766				
Correct Predictions	0.3227				
Function Value	63.525				

This table provides estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

First, the value of the function at the minimum is somewhat higher in this specification than in the general specification. However, this specification gives a higher percentage of "Correct Predictions" than the general specification discussed in Table 5. We will return on the fit of the specifications to the data later, when we discuss the case of (not perfectly) correlated unobservables.

Second, the coefficient bounds overlap with those that we estimated under the assumption of iid unobservables, and in general the bounds from the two specifications are very close to each other. The results concerning the effect of market size are relevant at this point. With the exception of the positive effect of market size on American's decision to enter, the effect on the decisions of the other firms is ambiguous: the coefficient bounds include zero as a possible value for the market size parameter. For example, we find that the effect of market size on the presence of Southwest is now included in $[-0.164, 0.046]$, while in Table 5 it was included in $[-0.262, -0.080]$. The other exception to the equality of bounds across Table 5 and Table 7 concerns the effect of Delta on American, which now is not statistically different from zero.

We find, as we did in Table 5, that some δ's are positive. In their investigation of this specification without firm specific indicators, Bresnahan and Reiss (1990) found that the event of higher duopoly profits than monopoly profits could take on positive probability for some range of market sizes, and had to constrain the parameters of the model to avoid such occurrence. We do find that monopoly profits can be lower than oligopoly profits, but our methodology does not require any restriction on the parameters to be estimated.

We also compute the bounds for the probability that equilibria occur, and we show them in Table 6 ("LB Perfect Corr" and "UB Perfect Corr"). We notice that the equilibria of the game that have nonnegative probability are those that we also found in the case of iid errors. The probability bounds are larger.

6.4 Correlated Unobservable Profits

We now consider the case when the unobservable profits are correlated among firms in a market. This consists of estimating a variance-covariance matrix of the errors. This is equivalent to allow the unobservable profits of American and Delta to be correlated in the airline markets: for example, in markets where American faces high fuel costs, Delta also faces high fuel costs. Another possibility is that there are unobservable characteristics of a market that we are unable to observe and that affect American and Southwest differently, so that when American enters, Southwest does not, and vice versa. Table 8 presents the results of this estimation:

The value of the distance function is much lower in this specification than in the specifications with perfectly correlated or iid unobservables. Moreover, the bounds are generally tighter and the percentage of correct predictions is higher in this specification than in the previous two. Hence, we conclude that this is the best specification in terms of fitting the data.

Table 8 shows that now American's entry has a negative effect ($[-3.222, -0.167]$) on the decision of small firms to enter into the market. In Table 5 we found that the entry decision of American and One Small airlines were not statistically significant. An explanation for the difference in the results is provided by a close look at Table 9, where we observe that the correlation between the unobservables of American and One Small carrier is included in $[0.299, 1]$. The positive correlation implies that there are unobservable factors that encourage

Table 8: **Correlated Unobservables**

	AA	DL	WN	Another Large	At Least One Small
AA	x	[0.058,4.635]	[-6.055,-0.136]	[2.021,3.610]	[-3.222,-0.167]
DL	[2.155,5.505]	x	[3.589,7.679]	[-9.823,-3.925]	[2.934, 7.426]
WN	[-7.001,-1.085]	[11.483,19.427]	x	[2.044,9.995]	[-8.142,-0.469]
Another Large	[1.423,2.385]	[-19.059,-12.607]	[3.249,5.521]	x	[1.998,4.078]
At Least One Small	[-3.827,-1.019]	[12.736,19.953]	[-10.456,-1.276]	[2.400,9.340]	x
Log Pop	[0.135,0.454]	[-0.426,0.464]	[-0.761,0.282]	[-0.108, 0.037]	[-0.431,0.427]
Log Distance	[-0.310,-0.007]	[-0.548,0.060]	[-1.038,0.094]	[-0.274,-0.179]	[-0.760,0.118]
Constant	[-3.899,-2.927]	[-1.900,0.599]	[-5.852,-2.953]	[0.439,0.817]	[-5.509,-2.093]
Nobs	766				
Correct Predictions	0.3516				
Function Value	37.983				

This table contains estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

both American and One Small entry. Interestingly, the correlation between American and Southwest unobservables is also positive. On the contrary the unobservables of One Small and Southwest are negatively correlated with the unobservables of Delta and Another Large. This suggests that American follows a different entry strategy than Delta and other large carriers.

As usual, we compute the probability bounds for this specification and we present them in Table 6 ("LB Corr" and "UB Corr"). More outcomes can occur with positive probability now. For example, the outcome $(1,1,1,1,1)$ which in the data is observed 0.0013 percent of the time, is not estimated to occur with a probability included in $[0.0000, 0.0363]$. On the other hand, outcomes that never occur in the data have now a nonnull probability of realization. Clearly, we are unable to conclude whether we do not observe them because this is just a sample, or because the parameters are not perfectly estimated.

6.5 Spatial Correlation

One of the main concerns when studying market structure and measuring market power in the airline market is that many large cities have more than one airport. For example, it is possible to fly from San Francisco to Washington on nine different routes. Following Borenstein [1989], we have assumed that flights to different airports are in separate markets. This makes sense because some segments of the demand, especially the business travelers, should put a strong preference for airports that are closer to the city[21] (DCA is few minutes

[21] More precisely here, cities refer to MSA's (Metropolitan Statistical Area).

Table 9: **The Variance-Covariance Matrix**

	Correlation
σ_{AADL}	[-0.998,-0.320]
σ_{AAWN}	[-0.012,0.998]
σ_{AAAL}	[-0.998,-0.615]
σ_{AAOS}	[0.299,1]
σ_{DLWN}	[-0.997,0.344]
σ_{DLAL}	[0.289,0.998]
σ_{DLOS}	[-0.998,-0.059]
σ_{WNAL}	[-0.978,-0.176]
σ_{WNOS}	[-0.077,1]
σ_{ALOS}	[-0.997,-0.360]

This table provides estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

from downtown Washington, while BWI is in Baltimore). On the other hand, other segments of the demand are probably willing to drive many miles to save several hundred dollars. One way to address this concern is by allowing the firms' unobservables to be correlated across markets between the same two cities. For example, American decides to enter in the Miami-Cleveland market (CLEMIA) because, given its strategic interaction with the other firms in the market and give its CLEMIA specific unobservables, American makes a nonnegative profit. Among the unobservables there could be the international network to the Caribbean and to South America that America controls from Miami. American also serves Fort Lauderdale (FLL), but American only provides very few international flights from FLL. AA's CLEMIA specific unobservables should then be correlated with the AA's CLEFLL specific unobservables. Table 10 shows the coefficient bounds when we estimate the general specification allowing for spatial correlation. In practice we proceed as follows. Whenever a market is included in the subsample that we draw to construct the parameter bounds, we also include any other market between the same two cities. For example, if the market CLEMIA is included in the subsample to compute the parameter bounds,' then we also include the market CLEFLL in the subsample.

The results of Table 10 should be compared with those of Table 8. The unit of observation is no longer an airport pair market, but a city pair market. As expected, the bounds are larger, since the number of independent observations that we use is smaller. As a result, few effects that were statistically different from zero in Table 8, are now not different from it. For example, we estimate the effect of American's entry on Delta's decision to enter to be now in $[-0.111, 4.635]$, while before we found it to be in $[0.058, 4.635]$. Otherwise, the results are very similar in Table 10, and Table 8[22].

Table 10: **Spatial Correlation**

	AA	DL	WN	Another Large	At Least One Small
AA	x	[-0.111,4.635]	[-6.055,0.146]	[2.021,3.610]	[-3.222,-0.136]
DL	[2.148,5.531]	x	[3.589,7.950]	[-10.121,-3.572]	[2.508,7.615]
WN	[-7.001,-1.085]	[6.374,19.427]	x	[1.689,10.123]	[-8.142,-0.469]
Another Large	[1.401,2.385]	[-19.059,-12.607]	[3.249,5.521]	x	[1.998,4.078]
At Least One Small	[-3.827,-0.606]	[12.488,19.953]	[-10.456,-0.977]	[2.400,9.366]	x
Log Pop	[0.135,0.455]	[-0.426,0.464]	[-0.761,0.282]	[-0.129,0.062]	[-0.431,0.447]
Log Distance	[-0.310,0.014]	[-0.603,0.060]	[-1.038,0.094]	[-0.293,-0.165]	[-0.760,0.189]
Constant	[-3.899,-2.927]	[-1.900,0.812]	[-5.852,-2.953]	[0.352,0.902]	[-5.727,-2.093]
Nobs	766				
Correct Predictions	0.3516				
Function Value	37.983				

This table contains estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

7 Estimates with Equilibrium Selection Rules

Our analysis does not assume any equilibrium selection rule. Bresnahan and Reiss (1990) and Berry (1992) assume that firms enter in a particular sequential order to introduce observable and unobservable heterogeneity in their estimation. For example, Bresnahan and Reiss assume that General Motors always enters into a market before Ford. Berry assumes that firms enter into markets in their order of profitability. In this final section, we consider two different selection rule, and compare the results with our previous findings.

The first selection rule is based on Berry (1992): whenever we have multiple outcomes that can be equilibria of the entry game, we choose that equilibrium where there is the firm that makes the highest profit. The idea is that the most profitable firms enter first.

The second selection rule is intended to provide a simple welfare comparison to the probability estimates in Table 5: whenever we have multiple outcomes that can be equilibria of the entry game, we choose that equilibrium where the sum of the profits of the entrants is the largest.

The analysis that follows assumes that the unobservables are iid among firms.

[22] Similar conclusions hold for the variance-covariance matrix estimates, which we do not present here for sake of brevity.

7.1 Selection Rule 1: The Firm Making Highest Profit Always Moves First

Probably the most intuitive equilibrium selection rule is the one that has the firms entering according to their order of profitability. The first firm to enter is the firm that makes the highest profit. For example, consider a market where there are two possible outcomes that can be an equilibrium. The first outcome is Delta as a monopolist, and the second outcome are American and Southwest as duopolists. Suppose that Delta makes the highest profit among the three in the two possible outcomes. Then the selection rule would tell that the chosen equilibrium is the one with Delta as a monopolist.

Table 11 presents the results of this estimation:

Table 11: **firm with highest profit**

	AA	DL	WN	Another Large	At Least One Small
AA	x	0.906	1.039	0.601	0.300
DL	-0.033	x	-0.004	-0.031	0.003
WN	-0.664	-0.043	x	0.883	-6.730
Another Large	-0.326	-0.874	-0.522	x	-0.525
At Least One Small	0.015	0.533	-0.759	0.148	x
Log Pop	0.095	-0.172	-0.324	-0.011	0.011
Log Distance	-0.169	-0.261	-0.016	-0.224	-0.435
Constant	-0.360	0.317	-0.695	0.591	0.164
Nobs	766				
Correct Predictions	0.2105				
Function Value	233.608				

This table provides estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

We immediately observe that the value of the distance function is much larger now than in Table 5. We also observe that the percentage of correct predictions is down to 21.05 percent from 28.12 percent. The best way to verify whether the selection rule describes the true firms' behavior is comparing the point identified parameters in Table 11 with the parameter bounds in Table 5.

Ideally, if it is true that the most profitable firm always moves first, then the point identified parameters in Table 11 should fall within the bounds in Table 5. This is not the case. Take, for example, the parameter estimate of the effect of Delta on American: In Table 11 the parameter is estimated to be equal to -0.033, while in Table 5 we found the parameter to be included in $[1.559, 2.310]$. This is true for almost all the estimated parameters. Thus, the selection rule that picks the firm with the highest profit does not appear to be supported by the data.

7.2 Selection Rule 2: The Equilibrium with the Highest Aggregate Profit is Selected

Consider now the selection rule that among the outcomes that can be an equilibrium in one market, we choose the one where the sum of the profits of the firms is the highest. This selection rule provides an interesting welfare implication, since it is the one that maximizes the firms' wealth.

Table 12 presents the results of this estimation:

Table 12: **market with largest sum of profits**

	AA	DL	WN	Another Large	At Least One Small
AA	x	6.290	-0.049	5.064	-0.666
DL	1.236	x	1.453	-5.665	2.351
WN	-2.046	12.922	x	5.284	-0.158
Another Large	1.482	-13.304	1.655	x	1.975
At Least One Small	-0.244	14.536	-1.657	9.507	x
Log Pop	0.228	-0.143	-0.127	0.034	0.035
Log Distance	-0.022	-0.170	-0.437	-0.289	-0.614
Constant	-2.995	-0.152	-1.713	0.712	-4.409
Nobs	766				
Correct Predictions	0.2219				
Function Value	259.344				

This table provides estimates of the 8-dimensional cube that contains the set estimates. These set estimates are appropriately constructed level sets of the sample objective function that cover the sharp identified set (which might not be convex) with 95% (See Chernozhukov, Hong, and Tamer (2002) for more details). The estimates in this table are the max and the min taken for every parameter.

We again observe that the value of the distance function is much larger now than in Table 5, and as it was the case for the Selection Rule 1, we also observe that the percentage of correct predictions is down to 22.19 percent from 28.12 percent. As before, the best way to verify whether the selection rule describes the true firms' behavior is comparing the point identified parameters in Table 12 with the parameter bounds in Table 5. The results are now extremely interesting. Most of the parameter estimates are very close to the bounds or are within the bounds presented in Table 5. This finding is interesting for two reasons.

First, it appears that a theory of equilibrium selection where firms maximize the total firm wealth cannot be rejected as a plausible selection rule. To explain how firms end up selecting this maximum aggregated wealth equilibrium is, however, beyond the scope of this paper. Second, and more generally, it seems clear that Selection Rule 2 can be used as a critical benchmark in the application of this methodology to the airline industry. Policy analysis based on the above rule will provide sharper predictions than ones done using the more general model with no specific selection mechanism.

8 Conclusions

In this paper, we have provided a framework for inference in discrete games that involve multiple decision makers and use it to study airline market structure in the US. We have made inferences on a "class of models" rather that looking for point identifying assumptions that pin down a unique model. We have estimated the profit functions of the players (airlines) under the assumption that their unobservables are iid; perfectly correlated; or correlated according to a variance-covariance matrix that we have estimated. We have also allowed for spatial correlation in the sample.

We find that there are multiple equilibria in the number and identity of firms in most markets. We also find that firms follow different entry strategies. In particular, American is less likely to be in market where Southwest or a small low cost carrier is also present. American's entry is positively related to the entry of Delta and other large carriers. On the contrary, Delta enters in markets where Southwest is present, and where one small low cost carrier is also present. Interestingly, Southwest and other low cost carriers avoid to be present in the same markets, suggesting that there is normally only space for one low cost carrier in each market.

We have also considered whether the evidence is consistent with two equilibrium selection rules. The first selection rule assumes that the firm making highest profit always moves first. The evidence rejects this selection rule. The second selection rule assumes that the equilibrium with the highest aggregate profit is selected. The evidence does not reject this selection rule, and we conclude that this selection rule can be used as a useful benchmark for policy analysis. More generally, our framework allows us to test among different equilibrium selection rules. This provides for example a set of equilibrium selection rules that is consistent with the model and the observed data. These set of *consistent or allowable* models can then be used to sharpen policy analysis.

References

U.S. General Accounting Office (1989): "Barriers to competition in the Airline Industry," .

Andrews, D., and S. Berry (2003): "On Placing Bounds on Parameters of Entry Games in the Presence of Multiple Equilibria," Working Paper.

Berry, S. (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 Journal of Economics*, 25(2), 242–262.

Borenstein, S. (1988): "On the Efficiency of Competitive Markets for Operating Licenses," *Quarterly Journal of Economics*, pp. 357–385.

——— (1992): "The Evolution of US Airline Competition," *Journal of Economic Perspectives*, 6(2), 45–73.

Bresnahan, T., and P. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

Bresnahan, T., and P. Reiss (1991): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57—81.

Bruekner, J. (2002): "Airport Congestion when Carriers have Market Power," *American Economic Review*, 92(5), 1357–1375.

Chernozhukov, V., H. Hong, and E. Tamer (2002): "Inference in Incomplete Econometric Models," Department of Economics, Princeton University.

Ciliberto, F., and E. Tamer (2003): "Barriers to Entry in Airline Markets," Working Paper.

Evans, W., and I. Kessides (1993): "Localized Market Power in the U.S. Airline Industry," *Review of Economics and Statistics*, 75(1), 66–75.

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 111, 1–51.

Heckman, J. (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46, 931–960.

Hendricks, K., M. Piccione, and G. Tan (1997): "Entry and exit in hub-spoke networks," *Rand Journal of Economics*, 28(2), 291–303.

Hurdle, G., R. Johnson, A. Joskow, G. Werden, and M. Williams (1989): "Concentration, Potential Entry, and Performance in the Airline Industry," *Journal of Industrial Economics*, 38(2), 119–39.

Maddala, G. S. (1983): *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press.

Manski, C., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–547.

Mazzeo, M. (1997): "Product Differentiation and Market Structure: An Empirical Model of the Hotel Industry," mimeo, Northwestern University.

Morrison, S., and C. Winston (1990): "The Dynamics of Airline Pricing and Competition," *American Economic Review*, 80(2), 389–393.

Politis, D., J. Romano, and M. Wolf (1999): *Subsampling*, Springer Series in Statistics. Springer:New York.

Reiss, P., and P. Spiller (1989): "Competition and Entry in Small Airline Markets," *Journal of Law and Economics*, 32, S179–S202.

Seim, K. (2002): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," Working Paper, Stanford Business School.

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

——— (1998): *Technology and Market Structure*. MIT Press.

Tamer, E. T. (2003): "Incomplete Bivariate Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70, 147–167.

Tirole, J. (1988): *The Theory of Industrial Organization*. MIT Press.

Toivanen, O., and M. Waterson (2000): "Empirical Research on Discrete Choice Game Theory Models of Entry: An Illustration," *European Economic Review*, 44(4-6), 985–92.

Transportation Research Board (1999): "Entry and Competition in the U.S. Airline Industry: Issues and Oppurtunities," Discussion Paper Special Report 255, National Research Council.

9 Appendix [Preliminary]

10 Proof of theorem 5:

Here assume that x has discrete support (this is not essential not essential for the proof of the present theorem. It will be needed later when constructing confidence regions.). In particular, let x take K values $x_1, \ldots, x_K$. Hence, the sample objective function we consider is

$$\begin{aligned}
Q_n(t) &= \frac{1}{N}\sum_{i=1}^{N} (P_N(x_i) - \hat{H}_1(x_i,t))^2 1[P_N(x_i) \le \hat{H}_1(x_i,t)] + (P_N(x_i) - \hat{H}_2(x_i,t))^2 1[P_N(x_i) \ge \hat{H}_2(x_i,t)] \\
&= \frac{1}{N}\sum_{i=1}^{N} (P_N(x_i) - \hat{H}_1(x_i,t))^2_- + (P_N(x_i) - \hat{H}_2(x_i,t))^2_+ \\
&= \sum_{j=1}^{K} \frac{n_j}{n}(P_N(x_j) - \hat{H}_1(x_j,t))^2_- + (P_N(x_j) - \hat{H}_2(x_j,t))^2_+
\end{aligned}$$

where $n_j = \sum_i 1[x_i = x_i]$, $u^2_- = u^2 1[u \le 0]$ and similarly for u^2_+. Also, we are considering scalar vectors H_1 and H_2 for simplicity. This consistency theorem is similar to the theorem 5 in Manski and Tamer (2002). Hence what we need essentially, is to show that the objective function $Q_N(t)$ converges to its expectation *uniformly* in t. For this it is sufficient to show that

$$\operatorname*{Sup}_t \left\| \frac{n_j}{n}(P_N(x_j) - \hat{H}_2(x_j,t))^2_+ - p_j(P(x_j) - H_2(x_j,t))^2_+ \right\| = o_p(1) \tag{18}$$

where $\frac{n_j}{n} \to p_j$ as $n \to \infty$ (the case for H_1 is similar. We have

$$\begin{aligned}
\left\| \frac{n_j}{n}(P_N(x_j) - \hat{H}_2(x_j,t))^2_+ - p_j(P(x_j) - H_2(x_j,t))^2_+ \right\| &\le \left\| \frac{n_j}{n}\left((P_N(x_j) - \hat{H}_2(x_j,t))^2_+ - (P(x_j) - \hat{H}_2(x_j,t))^2_+\right) \right\| \\
&\quad + \left\| \frac{n_j}{n}(P(x_j) - \hat{H}_2(x_j,t))^2_+ - p_j(P(x_j) - H_2(x_j,t))^2_+ \right\| \\
&= (1) + (2)
\end{aligned} \tag{19}$$

First, consider (1)

$$\begin{aligned}
(1) \le \frac{n_j}{n} &\left\| (P_N(x_j) - \hat{H}_2(x_j,t))^2 - (P(x_j) - \hat{H}_2(x_j,t))^2 \right\| 1[(P_N(x_j) \ge \hat{H}_2(x_j,t); P(x_j) \ge \hat{H}_2(x_j,t)] \\
&+ (P_N(x_j) - \hat{H}_2(x_j,t))^2 1[P_N(x_j) \ge \hat{H}_2(x_j,t) \ge P(x_j)] \\
&+ (P(x_j) - \hat{H}_2(x_j,t))^2 1[P(x_j) \ge \hat{H}_2(x_j,t) \ge P_N(x_j)] \\
&= (1a) + (1b) + (1c)
\end{aligned}$$

we have

$$(1a) = \frac{n_j}{n}\left\|(P_N(x_j) - P(x_j) - 2\hat{H}_2(x_j,t))((P_N(x_j) - P(x_j))^2\right\| 1[(P_N(x_j) \geq \hat{H}_2(x_j,t); P(x_j) \geq \hat{H}_2(x_j,t)]$$
$$\leq \frac{n_j}{n} M((P_N(x_j) - P(x_j))^2 = o_p(1)$$

where $M \leq \infty$ since $(P_N(x_j) - P(x_j) - 2\hat{H}_2(x_j,t))$ is bounded by definition. Moreover,

$$(1c) \leq ((P_N(x_j) - P(x_j))^2 = o_p(1)$$

since (1c) is constrained to be on the set $1[P_N(x_j) \geq \hat{H}_2(x_j,t) \geq P(x_j)]$ and similarly for (1b) which is constrained on the set $1[P_N(x_j) \geq \hat{H}_2(x_j,t) \geq P(x_j)]$. Consider (2'),

$$(2') = \frac{1}{n}\sum_n (P(x_i) - \hat{H}_2(x_i,t))_+^2 - E(P(x_i) - \hat{H}_2(x_i,t))_+^2 \tag{20}$$

Here, one can use a uniform law of large number similar to the one used in Pakes and Pollard (1989) to show that the above converges to zero uniformly in t. Moreover,

$$E(P(x_i) - \hat{H}_2(x_i,t))_+^2 - E(P(x_i) - H_2(x_i,t))_+^2 = o_p(1)$$

uniformly in t as sample size increases. This combined with (20) shows that (2) $= o_p(1)$ uniformly in t.

11 Constructing confidence regions:

We describe a practical method, based on Chernozhukov, Hong, and Tamer (2002) (CHT), that delivers a confidence region for the identified set. This confidence region is an appropriately constructed level set of the objective function similar to confidence regions based on inverting a likelihood. First assume that we have a sample objective function, such as $Q_n(t)$ in (15) above that obeys the conditions of Theorem 2.5, i.e., that uniformly converges to an objective function $Q(.)$ that is minimized on a set Θ_I (this set can be a singleton). Our confidence regions are level sets of the form

$$\widehat{\Theta}_I = \left\{\theta \in \Theta : n\left(Q_n(\theta) - \min_t Q_n(t)\right) \leq c\right\}$$

The results in CHT are derived for the case the the estimated set converges to the identified set at the parametric rate. To guarantee this, we assume that our data has finite points of support. First, define the coverage index $\rho(c)$ as follows:

$$\rho(c) = \quad c - \sup_{\theta \in \Theta_I} n\left(Q_n(\theta) - q_n\right)$$

where $q_n = \min_t Q_n(t)$. The coverage index ρ is convenient to use because events such as whether $\Theta_I \subseteq C_n(c)$ or not is determined by the sign of this index. For example, for $\theta \in \Theta_I$ such that $\theta \notin C_n(c)$, we have $n(Q_n(\theta) - q_n) > c$ which implies that $\rho(c) < 0$. Next, we state the main result from CHT where basically the appropriate confidence region is the one based on a level set derived at a particular quantile of the statistic

$$\mathcal{C}_n = \sup_{\theta \in \Theta_I} n\left(Q_n(\theta) - q_n\right)$$

where

$$\lim_{n \to \infty} P\left\{\Theta_I \subseteq C_n(\hat{c}_\alpha) \equiv \widehat{\Theta}_I\right\} = \alpha$$

and $\hat{c}_\alpha \to c_\alpha$ and c_α is the α-quantile of $\mathcal{C}$ where $\mathcal{C}_n \to^d \mathcal{C}$. We use a subsampling procedure that provides an estimate of c_α. This subsampling procedure is as follows. First, we construct all subsets B_n of size $b << n$. We take b to be equal to $n/4$.[23] We then compute

$$\widehat{\mathcal{C}}_{i,b,n,c_0} = \sup_{\theta \in C_n(c_0)} b(Q_b(\theta) - Q_b)$$

for each i-th subset, $i \leq B_n$ where we take c_0 to be an appropriate chisquared critical value. We then compute the α-quantile of the numbers $\widehat{\mathcal{C}}_{i,b,n,c_0}$ which will get us the appropriate ϵ at which we obtain our level set which has the appropriate coverage properties (asymptotically).

12 Data Construction

Data are from the Web site *http://itdb.bts.gov/*, managed by the Bureau of Transportation Statistics.

We limit the analysis to the continental USA, because routes in Alaska and Hawaii are heavily subsidized by the federal government (e.g. for postal transportation). We consider 40 airports. These are ATL (Atlanta), AUS (Austin), BOS (Boston), BWI (Baltimore), CLE (Cleveland), CLT (Charlotte), CVG (Cincinnati), DAL (Dallas Lovie Field), DCA (Washington Reagan), DEN (Denver), DFW (Dallas Fort Worth), DTW (Detroit), EWR (Newark), FLL (Fort Lauderdale), HOU (Houston Hobby), IAD (Washington Dulles), IAH (Houston International), JFK (New York Kennedy), LAS (Las Vegas), LAX (Los Angeles), LGA (New York's La Guardia), MCI (Kansas City), MCO (Orlando), MDW (Chicago Midway), MEM (Memphis), MIA (Miami), MSP (Minneapolis), OAK (Oakland), ORD (Chicago O'Hare), PDX (Portland), PHL (Philadelphia), PHX (Phoenix), PIT (Pittsburgh), SAN (San Diego), SEA (Seattle), SFO (San Francisco), SJC (San Jose in California), SLC (Salt Lake City), STL (Saint Louis), TPA (Tampa).

[23]There is not general theory of picking a subsample size. See Politis, Romano, and Wolf (1999) for more on this. However, trying different $b's$ in this paper led to similar results.

We exclude markets between airports in the same Metropolitan Statistical Area, such as between Fort Lauderdale and Miami, or between Chicago Midway and Chicago O'Hare. Even when people are transported between airports in the same city, these are likely coding errors.

We exclude from the dataset markets that had less than four departures scheduled or performed at the origin airport and at the destination airport. For example, AUSPHL had only 1 departure scheduled in the month of May 1998 from PHL to AUS. We code the market AUSPHL as being not served on a non-stop basis by any airline.

The following table provides summary comments on the regional and small independent carriers.

Table 13: **Regional and Independent Carriers**

Name	Comments
Trans State	Operates today as an independent carrier with AA and US. In 1998 it operated also with other airlines, e.g. TWA.
Atlantic Southeast Airlines	Independent until 1999 when bought by Delta. Previously serving as DL Connection.
Mesaba	NW tried to buy it in 2001 but gave up. Still today is independent.
Frontier	
Tower Air	Ceased in 2000
Airtran	
Kiwi Intl.	
Vanguard	
Spirit	
Reno	Independent until 1999 then bought by American
ATA	
Midwest	