Identification and Estimation of a Discrete Game of Complete Information

Patrick Bajari, Han Hong, and Stephen Ryan*
University of Minnesota and NBER
Stanford University
MIT and NBER

May 21, 2007

Abstract

We discuss the identification and estimation of discrete games of complete information. Following Bresnahan and Reiss (1990, 1991), a discrete game is a generalization of a standard discrete choice model where utility depends on the actions of other players. Using recent algorithms to compute all of the Nash equilibria to a game, we propose simulation-based estimators for static, discrete games. We demonstrate that the model is identified under weak functional form assumptions using exclusion restrictions and an identification at infinity approach. Monte Carlo evidence demonstrates that the estimator can perform well in moderately-sized samples. As an application, we study entry decisions by construction contractors to bid on highway projects in California. We find that equilibria in this game are more likely to be played if they are in mixed strategies and if they maximize joint profits.

*We would like to thank seminar participants and Victor Aguirregabiria, Lanier Benkard, Federico Ciliberto, Jeremy Fox, Stephanie Houghton, John Krainer, Andy McLennan, Andrew Sweeting and Chris Timmins for helpful discussions. This research was funded in part by National Science Foundation grants to Bajari (SES-0339328) and Hong (SES-0452143). We would like to thank Denis Nekipelov for research assistance. Special thanks are due to Ted Turocy for extending the Gambit software package to use homotopy solution methods at our request.

1 Introduction

In this paper, we study the identification and estimation of static, discrete games of complete information. These are the canonical normal form games of basic microeconomic theory, with a history dating back to the seminal work of Nash (1951). Econometrically, a discrete game is a generalization of a standard discrete choice model, such as the conditional logit or multinomial probit, that allows an agent's utility to depend on the actions of all other agents. The utilities of all agents are common knowledge, and we assume that observed outcomes are generated by a Nash equilibrium, where agents play strategies that are mutual best responses. Discrete game models been applied to diverse topics such as labor force participation (Bjorn and Vuong (1984), Kooreman and Soetevent (2006), entry (Bresnahan and Reiss (1990, 1991), Berry (1992), and Jia (2006)), product differentiation (Seim (2001), Mazzeo (2002)), technology choice (Ackerberg and Gowrisankaran (2006) and Ryan and Tucker (2006), Manuszak and Cohen (2004)), advertising (Sweeting (2006)), long term care and family bargaining (Stern and Heideman(1999), Stern and Engers (2002)), analyst stock recommendations (Bajari, Hong, Nekipelov and Krainer (2004)) and production with discrete units (Davis (2005)).

A generic feature of normal form games is that, for a given set of payoffs, there are often multiple Nash equilibria to the game. Therefore, the model does not satisfy the standard coherency condition of a one-to-one mapping between the model primitives and outcomes, which is problematic for identification and estimation. The literature has taken three approaches to dealing with multiple Nash equilibria. The first approach is to introduce an equilibrium selection mechanism which specifies which equilibrium is picked as part of the econometric model. Examples include random equilibrium selection in Bjorn and Vuong (1984) and the selection of an extremal equilibrium, as in Jia (2006). The second approach is to restrict attention to a particular class of games, such as entry games, and search for an estimator which allows for identification of payoff parameters irrespective of the presence of multiple equilibria. For example, Bresnahan and Reiss (1990, 1991) and Berry (1992) study models in which the number of firms is unique even though there may be multiple Nash equilibria. They propose estimators in which the number of firms, rather than the entry decisions of individual agents, is treated as the dependent variable. A third method, proposed by Tamer (2002), uses bounds estimation to estimate an entry model. The bounds are derived from the necessary conditions for pure strategy Nash equilibria which imply that the entry decision of one agent must be a best response to the entry decisions of other agents. Bounds estimation has also been used by Ciliberto and Tamer (2003), Pakes, Porter, Ho and Ishii (2005) and Andrews, Berry, Jia (2005). Berry and Tamer, (2006) and Reiss (2006) provide recent surveys of the econometric analysis on discrete games.

In this paper, we study identification and estimation of discrete complete information games, explicitly allowing for both multiple and mixed strategy equilibria. We propose a simulation-based estimator for these games. The model primitives include player utilities and an equilibrium selection mechanism which determines the probability that a particular equilibrium to the game is played. Using these primitives, we define a Method of Simulated Moments (MSM) estimator. We exploit recent algorithms that can compute all of the equilibria for general discrete games (see McKelvey and McLennan (1996)). Finding the entire set of Nash equilibria is computationally expensive in all except the most simple games. For moderately sized games, for example five players each of whom has two potential actions, we have found that it may take up to 20 minutes of CPU time on a 3.0 GHz single processor workstation to compute all the Nash equilibria. Therefore, we construct a smooth simulator for our model using an approach related to work on importance sampling by Ackerberg (2004) and Keane and Wolpin (1997, 2001). As we shall demonstrate in Section 3, this algorithm significantly reduces the computational burden of estimation and can be easily implemented as a parallel process. In a Monte Carlo study, we demonstrate that it is possible to construct estimates and standard errors for our model with less than a day of CPU time on a standard processor. We provide Monte Carlo evidence that our estimator works well even with moderately size samples. Finally, we apply our framework to study entry in an asymmetric first-price auction model. Using a unique data set, we study the strategic decision of contractors to bid on highway repair contracts in California and we estimate the probability of alternative equilibrium to the entry game.

Our approach makes several contributions to the literature on estimating static discrete games. First, our approach can be applied any normal form game of complete information. Several of the previous approaches in the literature restrict attention to specific classes of games, such as entry games or games with strategic complementarities. Also, to the best of our knowledge, our estimator is the only approach which can accommodate mixed strategy equilibria. In Section 2, we demonstrate that unless strong restrictions are made on the underlying payoffs or on the support of the error terms, every discrete game with complete information generates equilibrium sets that contain no pure strategy equilibria with positive probability. Moreover, some research argues that mixed strategy equilibria are likely in some settings, such as zero-sum games. For example, in their study of penalty kicks, Chiapporri, Groseclose, and Levitt (2002) find evidence in favor of mixed strategies. Levin and Smith (2001) conduct an experimental study of entry in auctions and find evidence in favor of the mixed strategy entry equilibrium compared to the pure strategy entry equilibrium. In experimental studies, El-Gamal and Grether (1995) and Shachat and Walker (2004) both found that mixed strategy equilibria can be consistent with an unobserved mixture of Bayesian learning by players.

Second, we explicitly model and estimate the equilibrium selection mechanism. McKelvey and McLennan (1997, 1997) have established that normal form games generically have large numbers of Nash equilibria that increase at an exponential rate as the number of players and/or actions grows. Estimating the selection mechanism allows the researcher to simulate the model, which is central to performing counterfactuals. This contrasts with the earlier literature on discrete games, which proposed estimators which do not specify which equilibrium is selected, making it impossible to simulate the model.

Understanding how equilibria are selected in actual plays of a game is also a topic of independent interest. There is a large and influential literature on refinements of the Nash equilibrium solution concept, such as trembling hand perfection or stability. However, there may be a large number of Nash equilibria which satisfy even the strongest refinements. Currently, there is no generally accepted method in economic theory for selecting between alternative equilibria to a normal form game. As a result, in some applications, the usefulness of game theory may be limited because the economist is forced to either make simplifying assumptions which guarantee a unique outcome or propose an *ad hoc* rule for selecting between multiple equilibria. We contribute to the literature by taking an empirical approach to the problem of equilibrium selection. We believe that an empirical approach may be useful given the lack of theory for selecting between alternative equilibria in many applications.

Our third contribution is to propose sufficient conditions for the semiparametric identification of both the structural parameters underlying the payoff functions and the parameters of the equilibrium selection mechanism. We propose two separate sets of conditions. The first identification strategy is based on an identification at infinity argument. Here we suppose that the structural utility parameters can be defined as a linear index, and that the covariates have a sufficiently rich support. We demonstrate that it is possible to identify the structural parameters of our model by examining choice behavior for sufficiently large values of the covariates. The second strategy is based on finding an appropriate exclusion restriction. For example, if there are covariates that shift the utility of one player, but can be excluded from the utility of another player, then we demonstrate that both payoffs and the equilibrium selection mechanism are locally identified. In an entry game, for example, we would search for a covariate that shifts the profitability of one firm for entering a particular market that can be excluded from the profits of other firms. An example of this could be distance from headquarters, as in empirical studies of entry by discount retailers (see Jia (2006) and Holmes (2006)). This excluded variable allows us to generate variation in all of the choice probabilities, while only changing the payoff for one player. We demonstrate that this type of variation is sufficient to locally identify the structural parameters of our model.

Our identification strategy is closely related to approaches found in treatment effect and sample selection models. The probability that a particular equilibrium is played is analogous to the selection equation, and the equation that determines utility corresponds to the treatment equation. In sample selection models, it is well known that identification under weak functional form assumptions often requires an exclusion restriction or identification at infinity (see Heckman (1990)). These models have simpler structure than our model of a discrete game since they only consider the actions of a single agent acting in isolation. It follows that equally, or even more, stringent assumptions will be required in our more complicated models. Both exclusion restrictions and index restrictions have been commonly used to identify econometric models of discrete games. Bresnahan and Reiss (1991) and Tamer (2002) use these restrictions to identify latent utility parameters in two by two games. These restrictions are necessary because, as shown in Bresnahan and Reiss (1991), without any restrictions all outcomes are observationally equivalent in games other than two by two games. To the best of our knowledge, we are the first to use these restrictions to identify both payoffs and the equilibrium selection mechanism in general normal form games.

Finally, we consider an application of our estimator to the study of entry in auctions. Entry in auctions has been considered in earlier research, but researchers have not formally treated the possibility of multiple equilibria to the auction game (see Bajari and Hortacsu (2003) and Athey, Levin, and Seira (2006)). We construct a data set of bidder entry into procurement auctions for highway paving projects in California. This application fits our modeling assumptions well. First, contractors' entry decisions can reasonably be modeled as a simultaneous move game. Contractors are prohibited by antitrust law from communicating before submitting their bids, enforced by the threat of both civil and criminal penalties. Second, an observation in our data set is the decision to bid for a single, precisely specified construction project with a fixed duration. In our data, we find that backlog and other dynamic factors are fairly minor in accounting for bidding behavior. Thus, we argue that our entry decision can be reasonably modeled as static, isolated instances of the entry game. In other applications, entry decisions will involve competing in a market for an indeterminate period of time which suggest allowing for a dynamic model may be important. The focus of our application is the estimation of an equilibrium selection mechanism. We allow the probability that a particular equilibrium is observed to depend on if the equilibrium is in pure strategies, maximizes joint profits, has the highest Nash product among pure strategies, and if it is dominated. To the best of our knowledge, this is the first empirical test of alternative criteria for selecting between Nash equilibrium in a normal form game.

2 The Model

The model is a simultaneous move game of complete information, commonly referred to as a normal form game. There are $i = 1, \dots, N$ players, each with a finite set of actions A_i. Define $A = \times_i A_i$ and let $a = (a_1, \dots, a_N)$ denote a generic element of A. Player i's von Neumann-Morgenstern (vNM) utility is a map $u_i : A \to R$, where R is the real line. Let π_i denote a mixed strategy over A_i. A Nash equilibrium is a vector $\pi = (\pi_1, \dots, \pi_N)$ such that each agent's mixed strategy is a best response.

Following Bresnahan and Reiss (1990, 1991), assume that the vNM utility of player i can be written as:

$$u_i(a, x, \theta_1, \epsilon_i) = f_i(x, a; \theta_1) + \epsilon_i(a). \tag{1}$$

We will sometimes abuse notation and write $u_i(a)$ instead of $u_i(a, x, \theta_1, \epsilon_i)$. In Equation 1, i's vNM utility from action a, $u_i(a)$, is the sum of two terms. The first term is a function $f_i(x, a; \theta_1)$, which depends on a, the vector of actions taken by all of the players, covariates x, and parameters θ_1. The second term is $\epsilon_i(a)$, a random preference shock. The term $\epsilon_i(a)$ reflects information about utility that is common knowledge to the players, but not observed by the econometrician. In games where there are a small number of players who know each other well, they will observe important information about each other that is not observed to the econometrician. Note that the preference shocks depend on the entire vector of actions a, not just the actions taken by player i. In much of the literature, stochastic shocks are only a function of player i's own actions, which is less general than the present model. The $\epsilon_i(a)$ are assumed to be i.i.d. with a density $g_i(\epsilon_i(a)|\theta_2)$ and joint distribution $g(\epsilon(a)|\theta_2) = \prod_i \prod_{a \in A} g_i(\epsilon_i(a)|\theta_2)$. We could easily modify our estimator to allow the $\epsilon_i(a)$ to only depend on the actions of i or to drop the independence assumption, for example by including random effects to account for unobserved heterogeneity.[1] We discuss the independence assumption in more detail in our section on identification.

Let $u_i = (u_i(a))_{a \in A}$ denote the vector of utilities for player i, and let $u = (u_1, \dots, u_N)$. Given that there may be more than one equilibrium for a particular u, let $\mathcal{E}(u)$ denote the set of Nash equilibria. We now introduce a mechanism for how a particular equilibrium is selected in the data. We let $\lambda(\pi; \mathcal{E}(u), \beta)$ denote the probability that equilibrium $\pi \in \mathcal{E}(u)$ is selected, where β is a vector of parameters. In order for λ to generate a well-defined

[1] If $\epsilon_i(a)$ has full support, the set of games u that can be drawn has full support. Therefore, the support of the likelihood function is A for all parameter values and covariates. If $\epsilon_i(a)$ only depends on i's actions, a_i, the likelihood may not have full support, A. This may lead to a severe specification problem since the model could predict that some events may have zero probability.

distribution it must be the case that, for all u and β:

$$\sum_{\pi \in \mathcal{E}(u)} \lambda(\pi; \mathcal{E}(u), \beta) = 1.$$

Economic theory or the specifics of a particular application might suggest factors which favor some types of equilibria. For instance, economic theory suggests that an equilibrium is more plausible if it satisfies a refinement such as trembling hand perfection. Berry (1992) and Ciliberto and Tamer (2003) suggest an equilibrium could be more likely if it is in pure strategies, or if it maximizes the joint profits of firms in the industry. Given u and $\mathcal{E}(u)$ we could create dummy variables for whether a given equilibrium, $\pi \in \mathcal{E}(u)$, satisfies any of these criteria. Let $y(\pi, u)$ denote a vector of variables that we generate in this fashion. For instance, to construct an equilibrium selection mechanism based on the three factors listed above, define the following quantities:

$$y_1(\pi, u) = \begin{cases} 1 & \text{if } \pi \text{ is trembling-hand perfect,} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

$$y_2(\pi, u) = \begin{cases} 1 & \text{if } \pi \text{ is a pure strategy equilibrium,} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

$$y_3(\pi, u) = \begin{cases} 1 & \text{if } \left(\sum_i \sum_a \pi(a) u_i(a)\right) - \hat{u} = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where

$$\pi(a) = \prod_i \pi_i(a_i),$$

and

$$\widehat{u} = \max_{\pi' \in \mathcal{E}(u)} \left\{ \sum_i \sum_a \pi'(a) u_i(a) \right\}.$$

A parsimonious, parametric model of λ is then:

$$\lambda(\pi; \mathcal{E}(u), \beta) = \frac{\exp(\beta \cdot y(\pi, u))}{\sum_{\pi' \in \mathcal{E}(u)} \exp(\beta \cdot y(\pi', u))}. \tag{5}$$

Note that in Equation 5 the sum is taken over the distinct elements of the equilibrium set $\pi' \in \mathcal{E}(u)$. For each π', we calculate the vector $y(\pi, u) = (y_1(\pi, u), y_2(\pi, u), y_3(\pi, u))$ as above. Then we evaluate the standard logit formula where β weights the probability that a particular type of equilibrium is selected. The example above is meant to be a simple illustration of what a selection mechanism might look like in practice. It is easy to generalize λ to allow for a less restrictive functional form and a richer set of variables, $y(\pi, u)$.

Computing the set $\mathcal{E}(u)$, all of the equilibria to a normal form game, is a well studied problem. McKelvey and McLennan (1996) survey the available algorithms in detail. The free, publicly available software package, Gambit, has routines that can be used to compute the set $\mathcal{E}(u)$ using these methods.[2] Finding all of the equilibria to a game is not a polynomial time computable problem. However, the available algorithms are fairly efficient at computing $\mathcal{E}(u)$ for games of moderate size. Readers interested in the details of the algorithms are referred to McKelvey and McLennan (1996). In the next sections, we shall take the ability to compute $\mathcal{E}(u)$ as given.

2.1 Discussion

2.1.1 Mixed Strategies

Allowing for mixed strategies in our framework is necessary because there is a strictly positive probability of a mixed strategy equilibrium if the error term has large enough support. As a result, our estimator would be ill-defined if we restricted attention to pure strategy equilibria. Consider the well-known game of matching pennies, illustrated in the figure below:

Matching Pennies

	H	T
H	(1,-1)	(-1,1)
T	(-1,1)	(1,-1)

In matching pennies, each player simultaneously chooses heads (H) or tails (T). If the choice of strategies match, then player one receives utility of one and player two receives a utility of negative one. If the strategies differ, the payoffs are reversed. The only equilibrium to this game is in mixed strategies with each player placing probability 1/2 on H and 1/2 on T. Consider games that have payoffs in a neighborhood of matching pennies by perturbing the payoffs as follows:

Perturbed Game

	H	T
H	(1+$\varepsilon_1(H,H)$,-1+$\varepsilon_2(H,H)$)	(-1+$\varepsilon_1(H,T)$,1+$\varepsilon_2(H,T)$)
T	(-1+$\varepsilon_1(T,H)$,1+$\varepsilon_2(T,H)$)	(1+$\varepsilon_1(T,T)$,-1+$\varepsilon_2(T,T)$)

[2]Gambit can be downloaded on the web from http://econweb.tamu.edu/gambit/.

For sufficiently small, but still non-zero, values of ϵ it can easily be verified that there is no pure strategy equilibrium to this game. For example, (H,H) cannot be a pure strategy equilibrium since player 2 would have an incentive to deviate and play T. Thus, there is an open set of payoffs for which the game that only has an equilibrium in mixed strategies. As a result, in our model the probability that payoffs are in this open set is strictly positive and the probability of a mixed strategy equilibrium is also strictly positive. It is straightforward to show that this result can be generalized to games with more players and more strategies. If we only allowed for pure strategies, the model would have no equilibrium with probability greater than zero and would not be well defined.

Previous research on complete information games generally limits attention to entry games (see Bresnahan and Reiss (1990, 1991), Berry (1992) and Tamer (2002)). These papers carefully restrict payoffs to guarantee the existence of a pure strategy equilibrium. Thus, the estimators proposed in these papers, which restrict attention to mixed strategies, do not need to accommodate mixed strategies. However, since we are interested in a more general specification of payoffs, we must allow for mixed strategies. To the best of our knowledge, our framework is the only available method which can accommodate mixed strategies.

2.1.2 Equilibrium Selection

A unique aspect of our framework is that we include the equilibrium selection mechanism, λ, in our econometric model. The inclusion of λ is useful for two reasons. First, there are frequently multiple Nash equilibria to a normal form game. Including λ specifies the probability of each equilibrium and therefore allows us to simulate the model. This is necessary for both the construction of our estimator in the next section and for counterfactual analysis. Second, equilibrium selection is an extremely important question in game theory and there is very little empirical work in this area. Using our modeling framework, we are able to empirically investigate equilibrium selection, which is important given that economic theory may provide little guidance about which equilibrium to select.

For example, consider the pure coordination game below, where player one chooses {T,B}, top or bottom, and player 2 chooses {L,R}, left or right.

Coordination Game

	L	R
T	(1,1)	(0,0)
B	(0,0)	(1,1)

This game has three equilibria (T,L), (B,R) and a mixed strategy equilibrium where each player plays each strategy with probability 1/2. Economic theory provides little guidance about which equilibrium is most likely in this game. For example, it does not seem possible to use theory to predict whether the (T,L) or (B,R) equilibrium is more plausible. Both equilibrium generate the same payoffs and only differ in the names assigned to the strategies. The inability of economic theory to select a unique equilibrium is not specific to this example. Many games generate multiple equilibria that satisfy the best known refinements in the theoretical literature.

Our approach allows for an empirical approach to equilibrium selection in this example. Suppose that the payoff matrix is known and that the economist has access to data on repeated plays of this game. With a sufficiently large number of observations, the economist will be able to precisely estimate the probability of observing the strategy pairs (T,L), (T,R), (B,L) and (B,R). In this example, knowing λ requires the economist to specify the probability that each of the three equilibria is played. Since the economist has knowledge of four probabilities, three of which are linearly independent, it follows that λ is identified. Therefore, while economic theory cannot be used to determine equilibrium selection, our simple example suggests that an empirical approach to this problem may sometimes be possible.

In our identification section, we investigate conditions under which both the equilibrium selection mechanism and the payoff matrix can be simultaneously identified. We demonstrate, similar to Bresnahan and Reiss (1991) that in general our problem is underidentified. However, we also describe two sets of sufficient conditions for identification that may be useful in some applications. Our identification results extend those of Bresnahan and Reiss (1991) and Tamer (2002) to general normal form games with flexible equilibrium selection mechanisms.

2.1.3 Comparison with Incomplete Information Games

An alternative approach used in the applied literature is to assume that the error terms only depend only player i's own actions and are private information. Incomplete information games are attractive for empirical work since it is often possible to estimate these models using a simple two-step procedure.[3] However, discrete games with incomplete information have a very different equilibrium structure than games with complete information. For example, in a coordination game Bajari, Hong, Krainer and Nekipelov (2006) show that the number of equilibria decreases as the number of players in the game increase. In fact, the equilibrium is typically unique when there are more than four players. In a complete information game, by comparison, the average number of Nash equilibrium will increase as

[3] See Pesendorfer and Schmidt-Dengler (2003), Aguirregabiria and Mira (2004), Brock and Durlauf (2001), Sweeting (2006) and Bajari, Hong, Krainer and Nekipelov (2006).

players are added to the game (see McKelvey and McLennan (1996)). Thus, the assumption of incomplete information appears to refine the equilibrium set.

While games of incomplete information have been used in several applied papers, they have not been studied in detail in economic theory (with the notable exception of Brock and Durlauf (2001) and McKelvy and Palfrey (1995)). Theorists have not characterized the structure of the equilibrium set and there is no literature on equilibrium refinement and selection for this class of models. On the other hand, for complete information games there is an extensive literature which characterizes the equilibrium set and which discusses equilibrium selection and refinements. Therefore, we believe it worthwhile to study games of complete information since the model is more closely linked to existing economic theory. However, there is currently no method for determining whether the data generating process corresponds to a complete information game or incomplete information game. We believe that evaluating the merits of games with complete and incomplete information is an important topic for future research.[4]

2.2 Examples of Discrete Games

The model that we propose is quite general and could be applied to many discrete games considered in the literature. We discuss three examples: entry, technology adoption with network effects, and peer effects. The first example is static entry into a market (see Bresnahan and Reiss (1990, 1991), Berry (1992), Tamer (2002), Ciliberto and Tamer (2003), and Manuszak and Cohen (2004)). In applications of entry games, the economist observes a cross section of markets and the players correspond to a set of potential entrants. The potential entrants simultaneously choose whether to enter: $a_i = 1$ denotes a decision by i to enter the market and $a_i = 0$ not to enter the market. In empirical work, the function f_i typically takes the form:

$$f_i = \begin{cases} \theta_1 \cdot x + \delta \sum_{j \neq i} 1\{a_j = 1\} & \text{if } a_i = 1, \\ 0 & \text{if } a_i = 0. \end{cases} \tag{6}$$

In Equation 6 the mean utility from not entering is set equal to zero.[5] The covariates x are variables which influence the profitability of entering a market, such as the number of consumers in the market, average income, and market-specific cost variables. The term δ measures the influence entry by other firms on firm i's profits. If profits decrease from having another firm enter the market, then $\delta < 0$. The error $\epsilon_i(a)$ capture shocks to the profitability of entry that are commonly observed by all firms in the market, but which

[4] In addition, it may be easier to compute the set of all equilibria in games of complete information. See Bajari, Hong, Nekipelov and Krainer (2004) for a discussion.

[5] We formally discuss this normalization in our section on identification.

are unobserved to the econometrician. In applied work, it might be desirable to include a market-specific random effect in the specification of the error term $\epsilon_i(a)$ in order to account for factors which shift profits that are commonly observed by the firms, but are unobserved by the econometrician. As discussed above, this extension to our estimator is completely straightforward.

A second example is technology adoption in the presence of network effects, as in Ackerberg and Gowrisankaran (2006) who model the decision by banks in spatially separated markets to adopt the Automated Clearing House (ACH) payment system. The players in the game are the existing banks in some market. Let $a_i = 1$ denote a decision to adopt ACH and $a_i = 0$ denote non-adoption. A priori, network effects are likely since the customers of bank i are able to transfer funds to customers of bank j if both i and j adopt ACH. An empirical model of network effects could take the form:

$$f_i = \begin{cases} \theta_1 \cdot x + \delta \sum_{j \neq i} 1\{a_j = 1\} \cdot c_j \cdot c_i & \text{if } a_i = 1, \\ 0 & \text{if } a_i = 0. \end{cases} \tag{7}$$

In Equation 7, x_i denotes some factors which influence the costs and benefits to adoption by firm i, such as the number of customers of bank i and their characteristics (e.g. large corporate or government agencies commonly use ACH to make automatic payroll deposits). The term c_i is the current number of clients of bank i. If $\delta > 0$, the term $\delta \sum_{j \neq i} 1\{a_j = 1\} \cdot c_j \cdot c_i$ captures the network effect. The marginal benefit of i's adoption of ACH is a function of the number of adopting customers at i's bank interacted with the number of adopters at other banks. The term $\epsilon_i(a)$ captures benefits to adoption observed by the banks but unobserved by the economist. Once again, market specific random effects could be used to account for unobserved heterogeneity in the benefits from adopting ACH.

A third example is peer effects, as in Manski (1993) and Brock and Durlauf (2001, 2003). A peer effect connotes a situation in which there is a benefit from conforming to the average or norm behavior. For example, consider the decision by a high school senior to take calculus. The players in the game are all of the students who could potentially take the class. Let $a_i = 1$ if student i decides to take calculus and $a_i = 0$ otherwise. The utility of student i is:

$$f_i = \begin{cases} \theta_1 \cdot x_i + \delta \sum_{j \neq i} 1\{a_j = 1\} \cdot s_j & \text{if } a_i = 1, \\ 0 & \text{if } a_i = 0. \end{cases} \tag{8}$$

In Equation 8, the covariates x_i could include terms that shift a student's incentives to take calculus, such as the educational status of her parents. The term s_i denotes the score of student i on a standardized achievement test and is commonly used to proxy for ability. If $\delta > 0$, the term $\delta \sum_{j \neq i} 1\{a_j = 1\} \cdot s_j$ captures a positive peer effect, i.e. the utility to student i from taking calculus in an increasing function of the number of other students who take calculus, interacted with the test scores of student i's peers.

The modeling framework we propose could be applied beyond these three examples. In principal, the framework above could be used to model any discrete choice where 1.) the payoffs of agents are interdependent, 2.) decisions are made simultaneously, and 3.) there is complete information. If the number of players or actions is very large, our estimator may not be computationally feasible due to the computational cost of solving for the entire equilibrium set. However, in the next section we describe an estimator which reduces the computational burden of estimation through the use of a parallel algorithm.

3 Simulation

Next, we propose a computationally efficient Method of Simulated Moments (MSM) estimator for θ and β, the parameters governing agent payoffs and the equilibrium selection mechanism, respectively. Let $P(a|x, \theta, \beta)$ denote the probability that a vector of strategies, $a = (a_1, \ldots, a_N)$, is observed conditional on x, θ, and β. MSM estimation requires an accurate and computationally efficiently method for simulating $P(a|x, \theta, \beta)$, which can be written as:

$$P(a|x,\theta,\beta) = \int \left\{ \sum_{\pi \in \mathcal{E}(u(x,\theta,\epsilon))} \lambda\left(\pi; \mathcal{E}\left(u(x,\theta_1,\epsilon)\right), \beta\right) \left(\prod_{i=1}^{N} \pi_i(a_i) \right) \right\} g(\epsilon|\theta_2) d\epsilon. \tag{9}$$

In principal, this integral could be simulated using a straightforward Monte Carlo procedure. First, pseudo random values of the random preference shocks $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$ are drawn from the distribution $g(\epsilon|\theta_2)$. Second, for each pseudo random error draw $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$, utilities are computed using Equation 1; we denote the utilities as $u(x, \theta, \epsilon)$ to emphasize their dependence on the parameters, covariates and preference shocks. Third, the equilibrium set $\mathcal{E}(u(x, \theta, \epsilon))$ is computed. And finally, the probability an event is observed is computed by summing over the equilibria $\pi \in \mathcal{E}$ and computing 1.) $\lambda(\pi)$, the probability that the equilibrium π is selected, and 2.) $\prod_{i=1}^{N} \pi_i(a_i)$, the probability that a is observed given π. By averaging over a large number of draws of ϵ, the economist could precisely simulate $P(a|x, \theta, \beta)$.

Unfortunately, this straightforward approach is not practical for applied work in all but the simplest games. The reason is that $P(a|x, \theta, \beta)$ must be simulated a large number of times for many different parameter values, and the equilibrium set $\mathcal{E}(u(x, \theta, \epsilon))$ must be recomputed every time the parameter values are changed. This can be computationally expensive; for example, we have found that it may take up to 20 minutes to compute $\mathcal{E}(u(x,\theta,\epsilon))$ for 5 player games with two strategies, for *each* draw of ϵ from $g(\cdot)$. As a result, the computational costs of using straightforward Monte Carlo integration for simulating Equation 9 may be prohibitive for applied work.

In order to lower the computational burden of simulating $P(a|x,\theta,\beta)$, we borrow from Keane and Wolpin (1997), Keane and Wolpin (2001) and Ackerberg (2004). First, we change the variable of integration in Equation 9 from ϵ to u. Let $h(u|\theta,x)$ denote the density u, conditional on θ and x. In many models, this density is trivial to compute. For instance, suppose that the preference shocks $\epsilon_i(a)$ are i.i.d. normal with density $\phi(\cdot|\mu,\sigma)$, with mean $\mu = 0$ and standard deviation σ. Then, the density $h(u|\theta,x)$ is:

$$h(u|\theta,x) = \prod_i \prod_{a\in A} \phi(u_i(a) - f_i(x,a;\theta_1)|0,\sigma)$$

which can easily be computed very quickly using standard programming packages, such as Fortran, C or Matlab. If we change the variable of integration from ϵ to u, then Equation 9 becomes:

$$P(a|x,\theta,\beta) = \int \left\{ \sum_{\pi\in\mathcal{E}(u)} \lambda\left(\pi;\mathcal{E}(u),\beta\right) \left(\prod_{i=1}^{N} \pi_i(a_i)\right) \right\} h(u|\theta,x)du. \tag{10}$$

In our simulator, we will use importance sampling; therefore, we rewrite Equation 10 as:

$$P(a|x,\theta,\beta) = \int \left\{ \sum_{\pi\in\mathcal{E}(u)} \lambda\left(\pi;\mathcal{E}(u),\beta\right) \left(\prod_{i=1}^{N} \pi_i(a_i)\right) \right\} \frac{h(u|\theta,x)}{q(u|x)} q(u|x)du,$$

where $q(u|x)$ is the importance density. For a given value of x, we draw a pseudo random sequence $u^{(s)} = (u_1^{(s)},\ldots,u_N^{(s)})$, $s = 1,\ldots,S$ of random utilities from the importance density $q(u|x)$. We then compute the equilibrium sets $\mathcal{E}(u^{(s)})$, a step which can be performed in parallel across several CPU's. Since computing the equilibria is the most burdensome step of the computation, the time required to estimate the model is roughly proportional to the inverse of the number of processors that the economist can exploit.

We can then simulate $P(a|x,\theta,\beta)$ as follows:

$$\widehat{P}(a|x,\theta,\beta) = \frac{1}{S}\sum_{s=1}^{S} \left\{ \sum_{\pi\in\mathcal{E}(u^{(s)})} \lambda(\pi;\mathcal{E}(u^{(s)}),\beta) \left(\prod_{i=1}^{N} \pi_i(a_i)\right) \right\} \frac{h(u^{(s)}|\theta,x)}{q(u^{(s)}|x)} \tag{11}$$

This simulator has three practical advantages for applied work. First, $\widehat{P}(a|x,\theta,\beta)$ will be a unbiased estimator of $P(a|x,\theta,\beta)$. Second, given the simulation draws $u^{(s)}$, the parameters θ and β do not enter into the expression for the equilibrium set $\mathcal{E}(u^{(s)})$. Therefore, when we change the parameter values, it is not necessary to recompute the equilibrium set $\mathcal{E}(u^{(s)})$. Exploiting this property will drastically reduce the time required to compute our estimator. Third, this simulator is a smooth function of the underlying parameters. As a result, the minimization of our MSM objective function will be numerically well behaved.[6]

The theory of importance sampling proves that $\widehat{P}(a|x,\theta,\beta)$ is a smooth and unbiased simulator for any choice of the importance density $q(u^{(s)}|x)$ that has sufficiently large support. However, as a practical matter, it is important that the ratio $\frac{h(u^{(s)}|\theta,x)}{q(u^{(s)}|x)}$ does not become too large. In order to ensure this, we need to make sure that the tails of the importance density $q(u^{(s)}|x)$ are not too thin in a neighborhood of the parameter that minimizes our MSM estimator. In our applied work, we have often constructed the importance density $q(u|x)$ by first estimating a version of the model in which the error terms $\epsilon = (\epsilon_1,\ldots,\epsilon_N)$ are private information instead of common knowledge. We then use the method proposed in Bajari, Hong, Nekipelov and Krainer (2004) to estimate the parameters of the private information version of the model. This is an extremely simple estimation problem and can be quickly programmed using a standard statistical package such as STATA. The importance density $q(u|x)$ is then set equal to the distribution of utilities conditional on x in the private information version of the game.[7]

3.1 The Estimator

The econometrician observes a sequence (a_t, x_t) of actions and covariates, $t = 1,\ldots,T$. Equation 11 can be used to form a maximum simulated likelihood estimator (MSL) for these observations. As is well known, MSL is biased for any fixed number of simulations. In order to obtain $\sqrt{T}$ consistent estimates, one needs increase the number of draws S so that

[6]We note that while we follow Keane and Wolpin (1997), (2001) and Ackerberg (2004) in constructing the importance sampler, its use in normal form game estimation is new. In addition, there is also a subtle difference between our use of the importance sampling and the use by previous authors. The dynamic discrete choice model in these earlier papers requires a complete random coefficient specification to allow the importance sampler to reduce the computation burden. Similarly, in a model of incomplete information games, the separation of equilibrium computation and numerical optimization in the estimation procedure also requires random coefficient specification. The complete information normal form game has the interesting feature that it does not require a random coefficient specification in order for the importance sampler to save on the computation burden of the estimator.

[7]This estimator can be performed in two stages. In the first stage, the economist flexibly estimates the choice probabilities $P(a|x)$ using standard methods. In the second stage, the economist assumes that these estimated choice probabilities represent the agent's equilibrium expectations. These choice probabilities are then substituted into the utility function. Bajari, Hong, Nekipelov and Krainer (2004) present a simple, static example of how to use these estimators.

$\frac{S}{\sqrt{T}} \to \infty$.[8]

Alternatively, one can estimate the parameters using MSM. An advantage of MSM is it generates an unbiased and consistent estimator for a fixed value of S. To form the MSM estimator, enumerate the elements of A from $k = \{1, \ldots, \#A\}$. Note that, because the probabilities of all of the elements of $a \in A$ must sum to one, one of these probabilities will be linearly dependent on the others, so there are effectively $\#A-1$ conditional moments. Let $w_k(x)$ be a vector of weight functions, with dimension larger than the number of parameters, for each k and let $1(a_t = k)$ denote the indicator function that the tth vector of actions is equal to k. The function $P(k|x, \theta, \beta)$ denotes the probability that the observed vector of actions is k given x and the parameters θ and β. This is probability is defined in Equation 9. At the true parameters of the data-generating process the predicted probability of each action equals its empirical probability for each action k:

$$E\left[1(a_t = k) - P(k|x, \theta, \beta)\right] w_k(x) = 0.$$

Using the sample counterpart of the above expectation, we form a vector of $\#A-1$ moments, where the k-th element is defined by:

$$m_{k,T}(\theta, \beta) = \frac{1}{T} \sum_{t=1}^{T} \left[1(a_t = k) - P(k|x_t, \theta, \beta)\right] w_k(x_t).$$

In practice, $P(k|x_t, \theta, \beta)$ is evaluated by simulation using the importance sampler in Equation 11. For each x_t, we draw a vector of S simulations $u_t^{(s)}$, where $s = \{1, \ldots, S\}$, from the importance density $q(u|x)$. We assume that the simulation draws $u_t^{(s)}$ are independent over both t and s, and are independent of the data. The k-th moment condition is then replaced by its simulation analog:

$$\hat{m}_{k,T}(\theta, \beta) = \frac{1}{T} \sum_{t=1}^{T} \left[1(a_t = k) - \widehat{P}(k|x_t, \theta, \beta)\right] w_k(x_t).$$

Then for a positive definite weighting matrix W_T, the MSM estimator is:

$$\left(\widehat{\theta}, \widehat{\beta}\right) = \arg\min_{(\theta,\beta)} \hat{m}_T(\theta, \beta)' \times W_T \times \hat{m}_T(\theta, \beta). \tag{12}$$

The asymptotic theory for estimating discrete choice models using MSL and MSM is well developed. See McFadden (1989), Pakes and Pollard (1989), or Hajivassiliou and Ruud

[8] In practice we have found that MSL can be useful for finding starting values for MSM. In our experience, the likelihood function is more concave around the maximum than in the MSM estimator.

(1994) detailed discussions.

4 Identification

We next develop two approaches to identification. To be clear, we are interested in conditions under which it is possible to recover the unknown primitives of the model in Section 2: $f(a,x)$ and $\lambda(x)$. In the first approach, we provide sufficient conditions to identify payoffs and the selection mechanism as the support of the covariates grows large. The second approach considers identification based on agent-specific payoff shifters. We first discuss two necessary restrictions on the data-generating process that are familiar from the discrete choice literature. We also provide some negative results on nonparametric identification before discussing the details of our approaches.

4.1 Scale and Location Normalizations

Assumption 1. *The payoffs of one action for each agent are fixed at a known constant.*

This restriction is similar to the argument that we can normalize the mean utility from the outside good equal to a constant, usually zero, in a standard discrete choice model. It is clear from the definition of a Nash equilibrium that adding a constant to all deterministic payoffs does not perturb the set of equilibria, so a location normalization is necessary. Similarly, a scale normalization is also necessary, as multiplying all deterministic payoffs by a positive constant does not alter the set of Nash equilibria either. This restriction is subsumed in the following assumption about the distribution of the error terms.

Assumption 2. *The joint distribution of $\epsilon = (\epsilon_i(a))$ is independent and known to all agents and the econometrician.*

Assumption 2 allows $\epsilon_i(a)$ to be any known joint parametric distribution. However, for expositional clarity, we shall assume that it has a standard normal distribution. This is a necessary condition, as even in the simplest discrete choice models it is not possible to identify both $f_i(a,x)$ and the joint distribution of the $\epsilon_i(a)$ nonparametrically. Consider a standard binary choice model where the dependent variable is 1 if the index $u(x)+\epsilon$ is greater than zero, i.e.

$$y = 1(u(x) + \epsilon > 0) \tag{13}$$

All the population information about this model is contained in the conditional probability $P(y=1|x)$, the probability that the dependent variable is equal to one given the covariates

x. If the CDF of ϵ is G, then Equation (13) implies that:

$$P(y=1|x) = G(u(x)), \tag{14}$$

Obviously, only the composition of $G(u(x))$ can be identified, and it is necessary to make parametric assumptions on one part (e.g. G or u) in order to identify the other part. For instance, if G is the standard normal CDF, we could perfectly rationalize the observed moments in Equation (14) by setting $u(x)$ to the inverse CDF evaluated at $P(y=1|x)$. Therefore, we will assume that the error terms are normally distributed.

We are also making the assumption that the $\epsilon(a)$'s are independently distributed. This assumption is also required for the identification of single agent models if the function $f(a,x)$ is nonparametrically specified. For example, consider a simple single agent multinomial choice model with three options. Denote the possible choices as $a \in \{1,2,3\}$. For $a = 1, 2$ let

$$u(a,x,\epsilon) = f(a,x) + \epsilon(a).$$

Also, the mean utility for the third option is normalized identically equal to 0:

$$u(3,x,\epsilon) \equiv \epsilon(3).$$

In the population, only two conditional probability functions are available to identify the model:

$$P\left(a=1|x\right) \quad \text{and} \quad P\left(a=3|x\right).$$

The last observable probability,

$$P\left(a=2|x\right),$$

is linearly dependent on the other two probabilities and does not help identification.

Holding x fixed, since there are only two moments, the unknowns $f\left(1,x\right)$ and $f\left(2,x\right)$ already exhaust the degrees of freedom available in the population. We can only hope to identify $f\left(1,x\right)$ and $f\left(2,x\right)$ by assuming that the *joint* distribution of the error terms is known. There are no additional degrees of freedom to identify the correlation structure between the error terms. Variation in x does not help because we place no restrictions on how $f(1,x)$ and $f(2,x)$ may vary with x. This is in contrast to a multinomial probit model or a nested logit model, which allows one to estimate the correlation coefficients between $\epsilon(1)$, $\epsilon(2)$ and $\epsilon(3)$. However, this comes at the cost of assuming a parametric functional form for the deterministic utility component $f(a,x)$.

4.2 Nonparametric Identification Results

Even if parametric methods are used, an estimation approach is more appealing if identification does not hinge on functional form assumptions. Therefore, in this section, we consider the nonparametric identification of our model.

A model is said to be identified if the model primitives can be recovered given the probability distributions the economist can observe. In a normal form game, the available population probabilities are $P(a|x)$ for $a \in A$, the probability distribution of the observed actions conditional on the covariates x. The primitives we wish to identify are $f(a, x)$ and $\lambda(x)$. That is, we wish to learn the vector of mean utilities $f(a, x) = (f_i(a, x))_{i=1,\ldots,N}$ and $\lambda(x)$ without making parametric assumptions about these objects.

We can generalize Equation (9) by writing $P(a|x)$ in a way that does not hinge on the specific parametric forms implicitly assumed in Section 2.

$$P(a|x, f, \lambda) = \int \left\{ \sum_{\pi \in \mathcal{E}(u(f,\epsilon))} \lambda\left(\pi; E\left(u\left(f, \epsilon\right), x\right)\right) \left(\prod_{i=1}^{N} \pi(a_i)\right) \right\} g(\epsilon) d\epsilon \tag{15}$$

In Equation (15), we write the vNM utilities as $u(f, \epsilon)$ to remind ourselves that they are a sum of the mean utilities $f(a, x)$ and the shocks ϵ. Holding x fixed we can view Equation (15) as a finite number of equations that depend on the finite number of parameters, $f(a, x)$ and $\lambda(x)$. Denote this system as $P(a|x) = H(f(a, x), \lambda(x))$ where H is the map implicitly defined by Equation (15). When writing H, assume that we drop one choice probability for each player. Since choice probabilities add up to one, this introduces a linear dependence between the rows of this system. We will let $DH_{f,\lambda}(x)$ denote the Jacobian formed by differentiating H with respect to the parameter vectors $f(a, x)$ and $\lambda(x)$.

Definition 1. *Given the probabilities $P(a|x)$, suppose that $f^0(a, x)$ and $\lambda^0(x)$ satisfy Equation (15). We will say that $\left(f^0(a, x), \lambda^0(x)\right)$ are locally identified if there exists an open neighborhood N_x of $\left(f^0(a, x), \lambda^0(x)\right)$ such that there is no other vector $\left(\widetilde{f}(a, x), \widetilde{\lambda}(x)\right) \in N_x$, $\left(\widetilde{f}(a, x), \widetilde{\lambda}(x)\right) \neq \left(f^0(a, x), \lambda^0(x)\right)$, that also satisfies Equation (15).*

In what follows, we shall often invoke the following assumption:

Assumption 3. *The map H is continuously differentiable. Also suppose that for all x, the Jacobian matrix $DH_{f,\lambda}$ has rank that is no less than the minimum of the row dimension and the column dimension.*

Assumption 3 implies that we can check local identification by comparing the number of moments, $P(a|x)$, to the number of free parameters $(f(a, x), \lambda(x))$. If the number of moments is greater than the number of parameters, then the implicit function theorem implies that the parameters are locally identified. While we can directly verify Assumption 3 for certain games, it can be difficult to do so for general games. At a minimum, this would require us to characterize the different sets of all equilibria that can be reached. This can be difficult in games with many players and strategies.

To fix ideas in what follows, consider the following simple game:

	L	R
T	$(\epsilon_1(TL), \epsilon_2(TL))$	$(\epsilon_1(TR), f_2(TR, x) + \epsilon_2(TR))$
B	$(f_1(BL, x) + \epsilon_1(BL), \epsilon_2(BL))$	$(f_1(BR, x) + \epsilon_1(BR), f_2(BR, x) + \epsilon_2(BR))$

In the tabulation of the above payoff matrix, we have implicitly adopted the following normalization:

$$f_1(TL, x) = f_1(TR, x) = f_2(TL, x) = f_2(BL, x) = 0.$$

The first result we note is that, even if the selection mechanism λ is known, a two by two game has more utility parameters that need to be identified than the number of moment conditions that can be observed in the data. Holding a given realization of x fixed, since there are two players with two strategies, the econometrician observes four conditional moments:

$$P(TL|x), P(TR|x), P(BL|x), P(BR|x).$$

However, because the probability of the actions must sum to one, there are effectively three moments that the econometrician observes. On the other hand, we have four utility parameters,

$$f_1(BL, x), f_1(BR, x), f_2(TR, x), f_2(BR, x)$$

that need to be identified. We note that variation in x does not help to reduce the total number of utility parameters that need to be identified because we place no restrictions on how the four utility parameters vary with x.

The result above can easily be generalized to generic games. Consider a game with N players and $\#A_i$ strategies for player i. Holding x fixed, the total number of mean utility parameters $f_i(a, x)$ is equal to

$$N \cdot \prod_i \#A_i - \sum_i \prod_{j \neq i} \#A_j.$$

This is equal to the cardinality of the number of strategies, times the number of players, minus the required normalizations. The number of moments that the economist can observe, conditional on x, is only equal to

$$\prod_i \#A_i - 1.$$

If each player has at least two strategies and if there are at least 2 players in the game, then for each given x the difference between the number of utility parameters, f_i, to estimate and the number of available moment conditions is bounded from below by

$$\left((N-1) - \frac{N}{2}\right) \prod_i \#A_i + 1 \geq 0.$$

Therefore, conditional on x, the number of mean utility parameters is greater than the number of moments available to the econometrician.

While the above calculations suggest that completely nonparametric identification is difficult in our setting, the next two sections do provide conditions that are considerably flexible under which the researcher can recover the underlying primitives. While our approach can be criticized as not being completely nonparametric, it does permit considerable flexibility. To the best of our knowledge, we establish the only results in the literature that allow for identification of the selection mechanism under weak functional form assumptions. At a minimum, we hope that this will be a useful starting point for further work.

4.3 Identification at Infinity

The first approach to providing conditions for the identification of the parameters in the payoffs and equilibrium selection mechanism is based on a strategy of identification at infinity. Suppose that the covariates have full support and that mean utilities are defined by a linear index of the covariates. Suitable sign restrictions are imposed on the covariates that have full support. The identification strategy involves two steps. In the first step, we attempt to identify the mean utilities by focusing on a path of the covariate values that permits a unique equilibrium with probability close to 1. We then vary the covariates locally to identify the parameters of the linear index on utility. In the second step, under an invariance assumption on the equilibrium selection mechanism, the equilibrium selection probabilities are identified from the observable choice probabilities. The basic idea is that if the number of equilibria is less than the number of moments, then identification follows from solving for the equilibrium selection parameters from the observable choice probabilities.

To see how this identification strategy works, consider the following game:

	L	R
T	$(0,0)$	$(0, f_2(TR, x_2) + \epsilon_2(TR))$
B	$(f_1(BL, x_1) + \epsilon_1(BL), 0)$	$(f_1(BR, x_1) + \epsilon_1(BR), f_2(BR, x_2) + \epsilon_2(BR))$

This is identical to the two-by-two game discussed above, save that we have applied the normalization that the payoff to one of the actions is equal to zero for both players. Assume that $X = \{x_1, x_2\}$ has full support on $\mathbb{R}^2$. Then it is possible to find x_2 such that player two will play L with probability approaching one. For this partition of x_2, to a first-order approximation the probability that player one chooses B is:

$$Pr(B) = Pr(f_1(BL, x_1) + \epsilon_1(BL) > 0). \tag{16}$$

Note that this is a single-agent decision problem: x_2 is such that player two is going to play L regardless of player one's decision. Therefore, player one will choose B if and only if the threshold condition in Equation 16 is satisfied. In the space where x_2 is large enough that only L is observed, the econometrician can recover $f_1(BL, x)$ by matching the empirical probabilities of playing B against their theoretical counterparts. To be concrete, suppose that $\epsilon_1(BL)$ is drawn from a standard normal distribution. Then the econometrician finds the value of $f_1(BL, x_1)$ such that:

$$Pr(B|L, x_1) = Pr(\epsilon_1(BL) > -f_1(BL, x_1)) = 1.0 - \Phi(-f_1(BL, x_1) = \Phi(f_1(BL, x_1)), \tag{17}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution, holds for all realizations of x_1. The uniqueness of $f_1(BL, x_1)$ is guaranteed by the monotonicity of $\Phi(\cdot)$.

An analogous argument can be made to identify all of the unknown payoff parameters by a suitable choice of either x_1 or x_2. Once we have recovered all the payoff parameters, the only unknowns are those governing the equilibrium selection mechanism. Under a regularity condition concerning the dimensionality of the unknowns entering equilibrium selection mechanism, formally discussed below, it is possible to identify the unknown parameters of that mechanism from observed probabilities of outcomes for the values of X where each player does not have a strictly dominant strategy with probability approaching one. We formalize this intuition for two-by-two games with general payoffs in the next section.

4.3.1 Formal Identification in Two-by-Two Games

We first make an assumption about the support of the covariates.

Assumption 4. *For each $i = 1, 2$, $j = T, B$ and $k = L, R$, there exists a set*

$$\mathcal{T}_i^{j(2-i)+k(i-1)}$$

of covariates x such that

$$\lim_{||x||\to\infty, x\in\mathcal{T}_i^{j(2-i)+k(i-1)}} P\left[a_i = j^{2-i}k^{i-1}|x\right] = 1.$$

Assumption 4 requires that for each player i and for each of player i's strategies, we can shift the covariates x along a dimension such that action a_i is a dominant strategy for player i with probability arbitrarily close to 1. For example, for $i = 2$ and $k = L$, Assumption 4 requires that along a path of $||x|| \to \infty$, $x \in \mathcal{T}_2^L$, $P[a_2 = L|x] \to 1$, or

$$P\left[\epsilon_2(TL) > f_2(TR, x) + \epsilon_2(TR), \epsilon_2(BL) > f_2(BR, x) + \epsilon_2(BR)\right] \longrightarrow 1.$$

Assumptions 1, 2, and 4 will basically allow us to identify the mean utilities along this path. The next assumption requires that we can extrapolate knowledge of the deterministic utilities along this path to other values of x on its support.

Assumption 5. *For $i = 1, 2$, $j = T, B$, $k = L, R$, the deterministic payoff functions*

$$f_{3-i}\left((i-1)Bk + (2-i)jR, x\right)$$

can be extrapolated from the path $||x|| \to \infty$, $x \in \mathcal{T}_i^{j(2-i)+k(i-1)}$ to the full support of x. For example, if $f_i((2-i)Bk + (i-1)jR, x) = x\beta_i^\tau$ for $\tau = (2-i)Bk + (i-1)jR$, then a sufficient conditional for global extrapolation of β_i^τ is that there exists some $L_0 > 0$ such that

$$\inf_{L\geq L_0} \min \operatorname{eig} E\left[xx'|x \in \mathcal{T}_{3-i}^{j(i-1)+k(2-i)}, ||x|| \geq L\right] > 0.$$

Theorem 1. *Under Assumptions 1, 2, 4, and 5, $f_i((2-i)Bk + (i-1)jR, x)$ is identified for all $i = 1, 2$, $j = T, B$ and $k = L, R$.*

Proof. The goal is to identify $f_i((2-i)Bk + (i-1)jR, x)$ for $i = 1, k = L, R$ and for $i = 2, j = T, B$. Because of Assumption 2, this is equivalent to identifying the single agent type choice probabilities, for $k = L, R$ and $j = T, B$:

$$P(f_i((2-i)Bk + (i-1)jR, x) + \epsilon_i((2-i)Bk + (i-1)jR) \geq \epsilon_i((2-i)Tk + (i-1)jL)|x).$$

Denote these *unconditional* choice probabilities by:

$$\bar{P}\left((2-i)Bk + (i-1)jR|x\right).$$

From the data we can only identify the conditional choice probabilities:

$$P\left(a_i = (2-i)\,B + (i-1)\,R \middle| a_{3-i} = (2-i)\,k + (i-1)j, x\right).$$

However, because of Assumption 4,

$$\lim_{||x||\to\infty, x\in\mathcal{T}_{3-i}^{(i-1)j+(2-i)k}} \left[P\left(a_i = (2-i)\,B + (i-1)\,R \middle| a_{3-i} = (2-i)\,k + (i-1)j, x\right) \right.$$
$$\left. -\bar{P}\left((2-i)\,Bk + (i-1)\,jR, x\right)\right] = 0.$$

This implies that $\bar{P}\left((2-i)\,B + (i-1)\,R, x\right)$ and hence $f_i\left((2-i)\,Bk + (i-1)\,jR, x\right)$ can be identified along the path of $||x|| \to \infty, x \in \mathcal{T}_{3-i}^{(i-1)j+(2-i)k}$. Assumption 5 then allows for the extrapolation of $f_i\left((2-i)\,Bk + (i-1)\,jR, x\right)$ along this path to the entire support of $x \in \mathcal{X}$. □

A special case of Assumption 4 is when $\epsilon = \left(\epsilon_i\left(jk\right)\right), i = 1, 2, j = T, B, k = L, R$ has finite support but the support for $f_i\left((2-i)\,Bk + (i-1)\,jR, x\right)$ for $i = 1, k = L, R$ and for $i = 2, j = T, B$ is either larger or infinite. Denote by L an upper bound of the absolute value of the support of, for all k and j,

$$\epsilon_i\left((2-i)\,Tk + (i-1)\,jL\right) - \epsilon_i\left((2-i)\,Bk + (i-1)\,jR\right).$$

Then a sufficient condition for Assumption 4 to hold is that

$$P\left[x : f_i\left((2-i)\,Bk + (i-1)\,jR, x\right) > L, \forall k = L, R, \forall j = T, B\right] > 0,$$

and

$$P\left[x : f_i\left((2-i)\,Bk + (i-1)\,jR, x\right) < -L, \forall k = L, R, \forall j = T, B\right] > 0.$$

Then we do not need the requirement that $||x|| \to \infty$. We can define the sets $\mathcal{T}_i^{j(2-i)+k(i-1)}$ to be

$$\mathcal{T}_i^{B(2-i)+R(i-1)} = \left\{x : f_i\left((2-i)\,Bk + (i-1)\,jR, x\right) > L, \forall k = L, R, \forall j = T, B\right\},$$

and

$$\mathcal{T}_i^{T(2-i)+L(i-1)} = \left\{x : f_i\left((2-i)\,Bk + (i-1)\,jR, x\right) < -L, \forall k = L, R, \forall j = T, B\right\}.$$

In this case, a sufficient condition for Assumption 5 to hold is that for all $i = 1$, $j = \{T, B\}$, and $i = 2$, $k = \{L, R\}$, the matrices

$$E\left[xx'|x \in \mathcal{T}_{3-i}^{(i-1)j+(2-i)k}\right] \tag{18}$$

are positive definite and finite. Under the finite support assumption for the error terms, one identifies the sets $\mathcal{T}_{3-i}^{(i-1)j+(2-i)k}$ from the sample, then recovers $f_i\left((2-i)\,Bk + (i-1)\,jR, x\right)$ on this set. This is followed by a linear regression of $f_i\left((2-i)\,Bk + (i-1)\,jR, x\right)$ on the set of $\mathcal{T}_{3-i}^{(i-1)j+(2-i)k}$ to recover β_i^τ for $\tau = (2-i)\,Bk + (i-1)\,jR$.

4.3.2 Identifying the Equilibrium Selection Mechanism

Given that the deterministic utility components are identified in Theorem 1, the next goal is to identify the equilibrium selection mechanism. In the following, with no loss of generality, we can set

$$\epsilon_i\left((2-i)\,Tk + (i-1)\,jL\right) = 0, \quad \text{for} \quad i = 1, k = L, R, i = 2, j = T, B.$$

Otherwise we can replace $\epsilon_i\left((2-i)\,Bk + (i-1)\,jR\right)$ with

$$\epsilon_i\left((2-i)\,Bk + (i-1)\,jR\right) - \epsilon_i\left((2-i)\,Tk + (i-1)\,jL\right).$$

The equilibrium selection probabilities are only needed when there are three equilibria, which can be either (TL, BR, mix) or (BL, TR, mix). The mixing probabilities for these two cases are:

$$P_m\left(R; x, \epsilon\right) = \frac{f_1\left(BL, x\right) + \epsilon_1\left(BL\right)}{f_1\left(BL, x\right) - f_1\left(BR, x\right) + \epsilon_1\left(BL\right) - \epsilon_1\left(BR\right)}, \quad P_m\left(L; x, \epsilon\right) = 1 - P_m\left(R; x, \epsilon\right)$$

and

$$P_m\left(B; x, \epsilon\right) = \frac{f_2\left(TR, x\right) + \epsilon_2\left(TR\right)}{f_2\left(TR, x\right) - f_2\left(BR, x\right) + \epsilon_2\left(TR\right) - \epsilon_2\left(BR\right)}, \quad P_m\left(T; x, \epsilon\right) = 1 - P_m\left(B; x, \epsilon\right).$$

In the ideal case where there are no error terms:

$$\epsilon_i\left((2-i)\,Bk + (i-1)\,jR\right) = 0, i = 1, k = L, R, i = 2, j = T, B.$$

All of $P_m\left(R\right), P_m\left(L\right), P_m\left(T\right)$ and $P_m\left(B\right)$ are functions of the known deterministic payoffs. Let the equilibrium selection probabilities be: $\rho\left(TL, x\right), \rho\left(BR, x\right), 1 - \rho\left(TL, x\right) - \rho\left(BR, x\right)$

in the case of (TL, BR, mix), and be $\rho(BL, x), \rho(TR, x), 1-\rho(BL, x)-\rho(TR, x)$ in the case of (BL, TR, mix), where the dependence on covariates x is made explicit. Then obviously for those values of x where (TL, BR, mix) is realized,

$$\begin{aligned}
P(TL|x) =& \rho(TL, x) + (1 - \rho(TL, x) - \rho(BR, x)) P_m(T) P_m(L) \\
P(TR|x) =& (1 - \rho(TL, x) - \rho(BR, x)) P_m(T) P_m(R) \\
P(BL|x) =& (1 - \rho(TL, x) - \rho(BR, x)) P_m(B) P_m(L).
\end{aligned}$$

These are three equations that identify the two unknown variables $\rho(TL, x)$ and $\rho(BR, x)$. Similarly, for values of x such that (BL, TR, mix) is realized,

$$\begin{aligned}
P(BL|x) =& \rho(BL, x) + (1 - \rho(BL, x) - \rho(TR, x)) P_m(B) P_m(L) \\
P(BR|x) =& (1 - \rho(BL, x) - \rho(TR, x)) P_m(B) P_m(R) \\
P(TL|x) =& (1 - \rho(BL, x) - \rho(TR, x)) P_m(T) P_m(L),
\end{aligned}$$

are the three equations that overidentify the two unknown variables $\rho(BL, x)$ and $\rho(TR, x)$.

In the presence of the unobservable error terms ϵ's, we need to impose certain identification assumptions that isolate the effects of the error terms. Denote

$$\rho(x, \epsilon) = [\rho(TL; x, \epsilon), \rho(BR; x, \epsilon), \rho(BL; x, \epsilon), \rho(TR; x, \epsilon)].$$

Assumption 6. *The equilibrium selection probabilities depend only on the latent utility indices:*

$$\rho(x, \epsilon) = \rho(f_i((2-i)Bk + (i-1)jR, x) + \epsilon_i((2-i)Bk + (i-1)jR, x), \forall i, j, k.)$$

This assumption rules out the possibility that $\rho(x, \epsilon)$ might depend on x and ϵ nonseparably, independent of the latent utility indexes.

Assumption 7. *The equilibrium selection probabilities are scale invariant with respect to the latent utility indexes. For all $\alpha > 0$,*

$$\rho(\alpha U_i((2-i)Bk + (i-1)jR, x), \forall i, j, k) = \rho(U_i((2-i)Bk + (i-1)jR, x), \forall i, j, k).$$

The scale invariance assumption, supplemented by the following support condition on the observables and unobservables, allows us to identify the equilibrium selection probabilities from the variations in the covariates x. In particular, Assumption 7 implies that the determinants for the equilibrium selection probabilities are the same as the determinants for the mixing probabilities, which are

$$\frac{U_i\left((2-i)\,Bk+(i-1)\,jR\right)-U_i\left((2-i)\,Tk+(i-1)\,jL\right)}{U_i\left((2-i)\,Bk'+(i-1)\,j'R\right)-U_i\left((2-i)\,Tk'+(i-1)\,j'L\right)},\quad \text{for}\quad i=1,2,k\neq k',j\neq j'.$$

Assumption 7 allows for a rich class of equilibrium selection mechanisms but does exclude some important ones. For example, it allows for the Pareto efficient equilibrium to be selected with a larger probability and for this probability to depend on the relative efficiency level. However, it does not allow this probability to depend on how much more efficient the efficient equilibrium is compared to the inefficient ones in absolute terms.[9]

Assumption 8. *There exists a set $\mathcal{T}$ such that for all $\epsilon>0$:*

$$\lim_{|x|\to\infty,x\in\mathcal{T}} P\left(\frac{f_i\left((2-i)\,Bk+(i-1)\,jR,x\right)}{U_i\left((2-i)\,Bk+(i-1)\,jR,x,\epsilon\right)}>1-\epsilon\right)=1, \tag{19}$$

for all $i=1,2,j=T,B,k=L,R$ and that for all $\Lambda=R,B,T,L$,

$$\lim_{|x|\to\infty,x\in\mathcal{T}} P\left(\frac{P_m\left(\Lambda;x,\epsilon\right)}{P_m\left(\Lambda;x,0\right)}>1-\epsilon\right)=1.$$

This assumption is satisfied if ϵ has finite support but $f_i\left((2-i)\,Bk+(i-1)\,jR,x\right)$ has infinite support.

Theorem 2. *Under Assumptions 1 to 8, the equilibrium selection probabilities*

$$\rho\left(U_i\left((2-i)\,Bk+(i-1)\,jR,x\right),\forall i,j,k\right)$$

are all identified from the observed choice probabilities.

Proof. Assumptions 1 to 5 identify the payoff functions $f_i\left((2-i)\,Bk+(i-1)\,jR,x\right)$ for all i,j,k. Using Assumption 8, we can approximate the mixing probabilities with arbitrary precision by using larger and larger values of the covariates x. This allows us to recover the equilibrium selection probabilities with arbitrary precision at very large values of the covariates x. The equilibrium selection probabilities with smaller values of the latent utility indexes are obtained by extrapolation along the remote sections of a ray that emanates from the origin and goes through the latent utility indexes. □

[9]This restriction follows intuitively from the idea that if all payoffs were scaled by a constant we would not expect the distribution over outcomes to change, which would be the case if the equilibrium selection mechanism which depended on absolute levels.

4.4 Extension to General Games

The identification results in the previous section for two by two games can be generalized to the more general case with more than two players and more than two strategies. The following assumptions are the direct analog of those in two by two cases.

Assumption 9. *For any* $i = 1, \ldots, N$ *and action profile* $k_{-i} \in A_{-i}$*, there exists a set* $\mathcal{T}_{-i}^{k_{-i}}$ *of covariates* x *such that*

$$\lim_{||x|| \to \infty, x \in \mathcal{T}_{-i}^{k_{-i}}} P\left(a_{-i} = k_{-i} | x\right) = 1. \tag{20}$$

This assumption requires that for each player i, the covariates x can be shifted along a dimension such that each element in k_{-i} is a dominant strategy for each player in $-i$. This assumption allows us to identify $f_i\left(a_i, a_{-i}, x\right)$ as a single agent discrete choice problem holding a_{-i} fixed at these values of the covariates x. The next assumption requires that utilities recovered from this path can be extended to the entire range of covariates.

Assumption 10. *For all* i *and all* $a \in A$*, the deterministic payoff functions* $f_i\left(a, x\right)$ *can be extrapolated from the path* $||x|| \to \infty, x \in \mathcal{T}_{-i}^{a_{-i}}$*, for all* i*, forall* a_{-i}*, to the full support of* x*.*

Theorem 3. *Under Assumptions 9 and 10,* $f_i\left(a, x\right)$ *is identified for all* i *and all* a *up to the normalization in Assumption 1.*

Identification of the equilibrium selection mechanism in the general case similarly requires the invariance assumption:

Assumption 11. *The equilibrium selection probabilities depend only on the latent utility indices:* $\rho\left(x, \epsilon\right) = \rho\left(u\left(a, x, \epsilon\right)\right)$*, and are scale invariant with respect to the latent utility indexes: for all* $\alpha > 0$*,* $\rho\left(\alpha u\left(a, x, \epsilon\right)\right) = \rho\left(u\left(a, x, \epsilon\right)\right)$*.*

Assumption 12. *There exists a set* $\mathcal{T}$ *such that for all* $\delta > 0$*:*

$$\lim_{|x| \to \infty, x \in \mathcal{T}} \min_{i,a} P\left[\left|\frac{f_i\left(a, x\right)}{u_i\left(a, x, \epsilon\right)} - 1\right| < \delta\right] = 1. \tag{21}$$

Theorem 4. *Under Assumptions 9 to 12, the equilibrium selection probabilities* $\rho\left(u\left(x, \epsilon\right)\right)$ *are all identified from the observed choice probabilities whenever the cardinality of* $\mathcal{E}\left(u\left(x, \epsilon\right)\right)$ *is less than or equal to* $\#A - 1$*.*

Remark: Note that the conditions in this theorem depend on the numbers of players and strategies, and generally also on the particular realization of $u(x, \epsilon)$. When the maximum number of equilibria for a game is less than or equal to $\#A - 1$, the condition in the above theorem holds uniformly for all realizations of $u(x, \epsilon)$. Such is the case, for example, for two by two games and for games with two players each equipped with four strategies.

In general, the maximum number of equilibria can be much larger than the number of moment conditions $\#A - 1$, indicating that the equilibrium selection mechanism can not identified for a certain range of the distribution of the random utility $u_i(x, \epsilon)$. On the other hand, the expected number of equilibrium is typically much smaller than the total number of moment conditions, suggesting that the equilibrium selection mechanism can be identified for a certain range of the random utility distribution of $u(x, \epsilon)$. See Appendix A for more details.

4.5 Exclusion Restrictions

The results of the previous section are not surprising in light of the analysis of Bresnahan and Reiss (1991) and Pesendorfer and Schmidt-Dengler (2003), who demonstrate failures of identification in discrete games. As we noted in the introduction, the structure of our models is not unlike treatment effect and sample selection models. The latent utilities f seem analogous to the treatment equation and λ to the selection equation. It is well known that these simpler models cannot be identified without exclusion restrictions. That is, we must search for variables that influence one equation, but not the other. In what follows, we demonstrate that a similar approach is possible in games. The exclusion restrictions that we consider are covariates that shift the utility of agent i but which do not enter as arguments into $u_j(j \neq i)$ or the equilibrium selecction mechanism λ. In many applications, such covariates are not difficult to find.

Assumption 13. *For each agent i, there exists some covariate, x_i that enters the utility of agent i, but not the utility of other agents. That is, i's utility can be written as $f_i(a, x, x_i)$. Furthermore, in addition to assumption 11, $\rho(u(\alpha, x, \epsilon))$ depends on $u(\alpha, x, \epsilon)$ only through a set of sufficient statistics of dimension $M \times (N-1)$ where M is a constant that does not depend on the number of players N.*

The first part of assumption 13 implies that there are agent i specific utility shifters. While this assumption is unlikely to be perfectly satisfied, to a first approximation it does seem reasonable in many applications. The second part of assumption 13 is a weak assumption that will be satisfied, for example, if the equilibrium selection probabilities depends only on the total utilities of all players in each equilibrium.

Theorem 5. *Suppose that Assumptions 1, 2, 6, 7, 11 and 13 hold. If $\#x_i$ are sufficiently large, the model is nonparametrically (locally) identified.*

Proof. The proof follows similarly to the previous section. Hold x fixed. Consider a large, but finite number of values of x_i equal to K for each agent. Consider the all the distinct vectors of the form $x = (x_1, \ldots, x_N)$ that can be formed. The number will be equal to K^N. Consider the moments generated by these K^N distinct covariates. The number of moments is equal to $K^N \cdot \left(\prod_i \#A_i - 1\right)$. The number of mean utility parameters is equal to $\sum_i K(\#A_i - 1)\prod_{j \neq i} \#A_j$ plus the number of parameters required to characterize λ. The maximum number of parameters required to characterize λ depends on the total number of players and the number of strategies for each player, and does not depend directly on x_i. Thus, the second part of assumption 13 implies that the number of equilibrium selection probability parameters grows at most at the rate of K^{N-1}. Since the number of utility parameters depends linearly on K and the number of equilibrium selection probabilities depends on K at the rate of K^{N-1}, but the number of moments grows exponentially with K at the rate of K^N. By choosing sufficiently large values for K the model is identified. □

The intuition behind the theorem above is quite simple. By shifting the individuals' utilities one at a time, it is possible to increase the number of moments at a faster rate than the number of parameters. Thus, identification of the model is possible.

Our results demonstrate that the model is identified if we have covariates that are indexed by the agent's identity i. Such covariates are used in most existing applications of discrete games. For example, consider empirical studies of strategic entry. In the case of an airline deciding whether to serve a particular city-pair, one such shifter could be the number of connecting routes that airline has at both endpoints, or whether one or both of the cities is a hub for that airline. These covariates which influence the payoffs of entering that market that do not directly influence the payoffs of that airline's competitors. In the case of entry by large discount retail chains, such as the entry of Walmart and Kmart as studied by Holmes (2006) and Jia (2006), an analogous payoff shifter is the distance of the market from regional distribution centers and company headquarters.

As a second example, consider technology adoption in the presence of network effects, as in Ryan and Tucker (2006). Here employees within a firm decide whether to adopt a videoconferencing technology which allows them to make video calls to other employees in the firm who have adopted the technology. The benefit to any given employee of adopting this technology depends on the adoption decisions of all the other employees. Furthermore, the benefit of using this technology varies with an employee's rank in the firm, their geographic locale, and their job function. All of these characteristics shift around the benefits of adoption on an individual basis and do not directly influence the payoffs of other employees. For example, senior managing directors in equities are likely to have different payoffs from using the network than a junior administrator in human resources.

We note that the above result demonstrates that our model is identified if we can estimate $f_i(a)$ up to an affine transformation in a first stage. This is commonly done in the literature on two-step estimation of dynamic games. The econometrician typically begins by estimating a differentiated product demand system and static markup following widely used methods such as Berry, Levinsohn, and Pakes (1995). After this demand system is estimated, the remaining "dynamic" parameters take the form of fixed entry costs. See Bajari, Benkard, and Levin (2006), Aguirregabiria and Mira (2004), Pesendorfer and Schmidt-Dengler (2003), and Pakes, Berry, Ostrovsky (2005). A similar approach could obviously be used in our framework. In the application that we consider in the present paper, we examine the decision of the four largest firms in the Californian construction industry to bid on highway paving procurement contracts. In the first step, we flexibly estimate firm level expected profits using the methods proposed by Guerre, Perrigne and Vuong (2000). The remaining parameters of $f_i(a)$ represent the fixed costs of submitting a bid.

Finally, we note that the maintained exogeneity assumptions used in our identification results are quite strong. We assume that the only form of unobserved heterogeneity is an iid shock to payoffs, $\epsilon_{ij}(a_i)$. However, it is quite straightforward to include random effects in our econometric model by modifying the importance sampler to permit correlation between the error terms. For example, in a study of entry, it would be natural to include market specific random effects to control for market specific determinants of entry which are commonly observed by the firms but which are unobserved to the econometrician. In our application, we control for unobserved heterogeneity in the first stage by including contract specific fixed effects in our estimates of expected profits. A similar strategy may be possible in other applications where markups are similarly estimated in a first stage using, for example, the method of Berry, Levinsohn, and Pakes (1995).

5 Monte Carlo

To demonstrate the performance of our estimator in small samples, we conducted a Monte Carlo experiment using a simple entry game with two players. Each player has the following profit function:

$$\pi_i(a) = 1(a_i = 1)\left\{\beta_1 x_{i1} + \beta_2 x_{2i} + \epsilon_i(a)\right\},$$

with the observable covariates defined by $x_{i1} \sim N(1, 1)$ and $x_{2i} = n(a)$, where $N(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2, and $n(a)$ is the number of competitors a firm faces given action profile a. The idiosyncratic error term, which is different for each player for each action profile a, is drawn independently from the standard normal distribution. The choice of unit variance in the random shock satisfies the need for a scale normalization, and assigning payoffs of zero to not entering the market satisfies the location normalization. We think about x_{1i} representing variability in profits to firm i from entering that market, and x_{2i} captures the effects of having a competitor. The true payoff parameters are $\beta_1 = 2$ and $\beta_2 = -10$.

The distributions of the covariates were chosen such that when payoffs are evaluated at their means it is optimal for only one of the two firms to enter the market. Under these circumstances the set of equilibria in this game, denoted by $\mathcal{E}$, has three elements: two pure strategies characterized by one firm or the other entering the market, and one mixed strategy where firms enter with some probability. We specify that the probability of equilibrium $\pi_i \in \mathcal{E}$ being played as:

$$Pr(\pi_i) = \frac{\exp(\theta_1 MIXED_i)}{\sum_{\pi_j \in \mathcal{E}} \exp(\theta_1 MIXED_j)},$$

where $MIXED_i$ is an indicator variable equal to one if equilibrium π_i is in mixed strategies. When $\theta_1 = 0$ one of the three equilibria is picked with equal chance. As that parameter tends to either negative or positive infinity, the mixed strategy is played with probability approaching zero or one, respectively. The true selection parameter is $\theta_1 = 1$.

Our game has three unknown parameters: β_1, β_2, and θ_1. The game generates moments from the probabilities of observing the four possible combinations of entry choices. Only three of these moments are linearly independent, as the probabilities must sum to one, implying our model is exactly identified. Consequently, we use the identity matrix as our GMM weighting matrix without loss of efficiency. We generated 500 samples of size $n =$ 25, 50, 100, 200, and 400 to assess the finite sample properties of our estimator. We set the number of importance games to be equal to the sample size, and generated new importance games for each replication. Standard nonlinear optimization techniques were used to find the estimated parameters. The results of our Monte Carlo are reported in Table 1.

The results are encouraging even in the smallest samples sizes. The payoff parameters are tightly estimated near their true values, while the mixed strategy shifter is estimated with considerably lower precision. The standard deviation of the estimates of all three parameters shrinks as the sample size increases, as does the mean and median absolute deviations. Significantly, the decrease in the standard deviation for the payoff parameters

Table 1: Monte Carlo Results

Parameter	Mean	Median	Standard Deviation	Mean Bias	Median Bias	MSE	Mean AD	Median AD
n = 25								
β_1	2.0071	2.0034	0.0149	0.0071	0.0034	0.0003	0.0091	0.0048
β_2	-10.0016	-9.9964	0.0534	-0.0016	0.0036	0.0028	0.0309	0.0208
θ_1	1.6866	0.8142	4.7465	0.6866	-0.1858	22.9559	2.2876	0.8235
n = 50								
β_1	2.0051	2.0036	0.0060	0.0051	0.0036	0.0001	0.0058	0.0040
β_2	-9.9996	-9.9981	0.0233	0.0004	0.0019	0.0005	0.0170	0.0123
θ_1	1.0456	0.7950	2.0455	0.0456	-0.2050	4.1778	0.9818	0.6008
n = 100								
β_1	2.0037	2.0032	0.0035	0.0037	0.0032	0.0000	0.0040	0.0032
β_2	-9.9990	-9.9975	0.0152	0.0010	0.0025	0.0002	0.0119	0.0101
θ_1	0.9394	0.8592	0.6194	-0.0606	-0.1408	0.3866	0.4789	0.3591
n = 200								
β_1	2.0036	2.0033	0.0024	0.0036	0.0033	0.0000	0.0036	0.0033
β_2	-9.9992	-9.9988	0.0095	0.0008	0.0012	0.0001	0.0076	0.0063
θ_1	0.9782	0.9555	0.4127	-0.0218	-0.0445	0.1705	0.3117	0.2639
n = 400								
β_1	2.0034	2.0032	0.0016	0.0034	0.0032	0.0000	0.0034	0.0032
β_2	-9.9986	-9.9983	0.0062	0.0014	0.0017	0.0000	0.0051	0.0043
Mixed Strategy	0.9810	0.9800	0.2855	-0.0190	-0.0200	0.0817	0.2256	0.1790

The true parameter vector is $\beta_1 = 2$, $\beta_2 = -10$, and $\theta_1 = 1$. Each sample size was evaluated 500 times.

is close to $\sqrt{n}$, as theory would imply. The rate of convergence of the equilibrium selection parameter is much more dramatic as the sample size gets increasingly larger. The precision of the estimated payoff coefficients intuitively follows from the fact that the payoff-relevant covariates define the thresholds at which firms are willing to enter a market, and thus enter into the likelihood of every observation directly. It is also of interest to note the relative precision of β_1 and β_2. Recall that the β_1 multiplies a covariate with continuous support, and enters into every entry decision, conditional on the other player's action, while β_2 influences a discrete shifter which only enters the payoff function when the other firm has entered the market. The combination of these factors implies that the first payoff coefficient is estimated more precisely than the second payoff coefficient.

The third parameter, θ_1, governs how often mixed strategies are played relative to pure strategies. This parameter is estimated with some slight upward bias in the smallest samples, and is estimated less precisely than the payoff parameters across all sample sizes. This is to be expected, as the selection parameter is identified off coordination failures between the firms due to the mixed strategy equilibrium. To illustrate, suppose that all payoffs, including idiosyncratic shocks, were observed by the econometrician. For some realizations of the covariates, the model would predict two pure strategies, with one or the other of the firms entering the market, and a single mixed strategy. If the mixed strategy equilibrium is played, there is a chance of either no firms entering the market or both firms entering the market. It is only when these mistakes are observed that the econometrician knows that the mixed strategy was played. This is a subtle and complex interplay among the components of the game, as the probability of observing a mistake is a function of both θ_1, which controls how often a mixed strategy occurs, and the payoffs of the game, which determine the probability of observing a mistake conditional on playing a mixed strategy.

This interplay illustrates a more general point, which is that although the parameters are theoretically identified, the estimation of some parameters may depend on a relatively small subset of outcomes. It should be emphasized that this is true even in the extreme case that the payoff functions, including the idiosyncratic shocks, are known with certainty, since the model itself generates probabilistic outcomes through both the equilibrium selection device and the nature of mixed strategies. In light of this, the results here are very positive, as we are able to recover unbiased estimates of the true parameters with acceptable precision in moderate sample sizes.

There is one caveat to our procedure that researchers have to address in practice. In each Monte Carlo, we knew the true parameters of the game, and we were able to generate importance games using these. With real data, of course, these parameters are initially unknown. The importance sampler can generate imprecise parameter estimates if given poor initial guesses, so it is necessary to derive starting parameters from another source. We demonstrate one possible solution to this problem in our application below by using a related game of private information to generate initial starting values.

6 Application

Next we use our estimator to model strategic entry by bidders into highway procurement auctions conducted by the California Department of Transportation (CalTrans) between 1999 and 2000. Econometric modeling of entry has been of considerable interest in empirical industrial organization; see Bresnahan and Reiss (1990, 1991), Berry (1992), Mazzeo (2002), Tamer (2002), and Ciliberto and Tamer (2003). Bajari and Hortacsu (2003), Li and Zheng (2006), Athey, Levin and Seira (2006) and Krasnokutskaya and Seim (2005) have studied entry in bidding markets.

Bidder entry in highway procurements is an attractive application for our estimator for three reasons. First, CalTrans awards its contracts using an open competitive bidding system. For each highway contract, there is a fixed and publicly announced deadline for submitting bids. Any communication between bidders about entry or other bidding decisions would be considered collusion and could lead to civil and criminal penalties. Therefore, the assumption of a simultaneous move game is plausibly satisfied in our application.

Second, there is a well developed empirical literature for estimating structural models of bidding for highway procurement contracts, see Porter and Zona (1999), Bajari and Ye (2003), Pesendorfer and Jofre-Bonet (2003), Krasnokutskaya and Seim (2005) and Li (2006). The flexible econometric methods proposed by Guerre, Perrigne and Vuong (2000) are commonly used in this literature, and in the empirical auctions literature more generally, to estimate the structural parameters of the model. In a first step, we use these methods to precisely estimate the expected payoffs to each player for all possible configurations of entry, conditioning on observable characteristics. In a second step, we estimate the fixed costs of bidding and the parameters of our equilibrium selection mechanism.

Finally, in our data set, the dependent variable is a decision by a contractor to submit a bid to complete a single and indivisible construction project. We focus on paving contracts, instead of all contracts awarded by CalTrans, as in Pesendorfer and Jofre-Bonet (2003), in order to reduce the importance of dynamics in our application. Most of the existing entry literature considers the decision by a firm to enter a spatially separated retail or service market and compete for an indefinite length of time. We believe that a static model is more plausible in our application than in much previous work on entry.

Our model of entry in auctions is similar to Athey, Levin and Serra (2006). In the first stage, contractors simultaneously choose whether to incur a fixed cost in order participate. In the second stage, participating contractors submit sealed bids in a first-price auction and the contract is awarded to the low bidder. Our model of entry often has multiple equilibria, and there is no clear criterion from economic theory that can be used to select a unique equilibrium to our game. Previous empirical research on entry in auctions abstracts from the multiplicity problem by imposing modeling assumptions that guarantee a unique equilibrium. We contribute to the literature on entry in auctions by estimating λ, the probability of selecting a particular equilibrium. We parameterize λ to allow for four criteria that have been suggested in the literature as potentially influencing equilibrium selection: the equilibrium is in pure strategies, the equilibrium maximizes joint profits, the equilibrium is dominated, and the equilibrium is in pure strategies and has the highest Nash product among pure strategy equilibria. To the best of our knowledge, this is the first empirical analysis of equilibrium selection in a normal form game.

6.0.1 The Bidding Game

In the model, there are $i = 1, \ldots, N$ potential bidders who bid on $t = 1, \ldots, T$ highway paving contracts. Following previous researchers, we model bidding in this industry as an asymmetric first-price auction with independent private values (see Porter and Zona (1999), Bajari and Ye (2003), Pesendorfer and Jofre-Bonet (2003), Krasnokutskaya and Seim (2005) and Li (2006)). Let $N(t) \subseteq \{1, \ldots, N\}$ denote the set of contractors who submit bids on project t. We assume that the set of bidders is common knowledge at the time bids are submitted.[10]

Before submitting a bid, bidder i will prepare a cost estimate $c_{i,t}$. The cost estimate of bidder i is private information which has a distribution $F_i(x_{i,t})$ where $x_{i,t}$ are publicly observable covariates which influence bidder i's cost distribution. We follow previous research and include in $x_{i,t}$ an engineering cost estimate, the distance of contractor i to project t, a measure of i's backlog, contractor fixed effects and project fixed effects. We assume that the cost distribution has a common support for all bidders and satisfies the regularity conditions

[10] In principal, it is possible to consider a model where bidders are uncertain about which firms will participate. Changing our estimator to allow for this possibility would be straightforward. However, existence and uniqueness of equilibrium bidding functions to the first price asymmetric auction with random entry has not yet been established to the best of our knowledge.

Also, we believe that allowing the set of bidders to be common knowledge corresponds most closely to what happens in this industry. Bidders that we have spoken with feel like they are quite knowledgeable about which other contractors will submit bids. Typically, the closest firms and firms with the lowest backlogs of outstanding work are most likely to bid. Also, CalTrans provides a list of plan holders for the project shortly before bids are due which allows the contractors to learn about which competing firms are interested in the project. A similar modeling assumption is made in Athey, Levin and Serra (2006) and Kransnokutskaya and Seim (2005).

discussed in LeBrun (1996) and Maskin and Riley (2000) so that an equilibrium exists, is unique, and is strictly increasing in a bidder's private information.

Let $\mathbf{b}_{i,t}(c_{i,t})$ be the bidding strategy used by bidder i in auction t, and let $\phi_{i,t}(b_{i,t})$ denote the inverse bid function. Bidders are assumed to be risk neutral. The expected profit to bidder i from bidding $b_{i,t}$ is:

$$(b_{i,t} - c_{i,t}) \prod_{j \in N(t), j \neq i} \left(1 - F_j\left(\phi_{j,t}(b_{i,t})|x_{i,t}\right)\right)$$

Expected profit is the product of two terms. The first term is a markup, $(b_{i,t} - c_{i,t})$, which reflects bidder i's profits conditional on winning the job. Since the bid functions are strictly increasing, the term $1 - F_j\left(\phi_{j,t}(b_{i,t})|x_{i,t}\right)$ is the probability that firm j's bid is greater than i's bid $b_{i,t}$. As a result, $\prod_{j \in N(t), j \neq i} \left(1 - F_j\left(\phi_{i,t}(b_{i,t})|x_{i,t}\right)\right)$ is the probability that bidder i wins the contract with a bid of $b_{i,t}$. Thus, expected profits are the product of a markup times the probability that firm i wins the contract.

Following Guerre, Perrigne, and Vuong (2000), we rewrite bidder i's profit maximization problem as:

$$\max_{b_{i,t}} \ (b_{i,t} - c_{i,t}) \prod_{j \in N(t), j \neq i} (1 - G_j(b_{i,t}|x_t)).$$

We let $G_j(b_{i,t}|x_t)$ denote the equilibrium distribution of bids submitted by firm j conditional on the publicly observed information $x_t = (x_{i,t})_{i \in N(t)}$. The first order conditions for profit maximization imply that:

$$c_{i,t} = b_{i,t} - \left[\sum_{j \in N(t), j \neq i} \frac{g_j(b_{i,t}|x_t)}{(1 - G_j(b_{i,t}|x_t)} \right]^{-1} \tag{22}$$

Note that the right hand side of the above equation is a function of $b_{i,t}$ and the distribution of bids, which can be estimated by pooling bidding data across contracts $t = 1, ..., T$. The left hand side is the structural parameter $c_{i,t}$ which is unobserved to the econometrician. Following Guerre, Perrigne and Vuong (2000), we will evaluate the empirical analogue of the right hand side of the above expression in order to recover the structural cost parameter $c_{i,t}$.[11]

[11] In Bajari, Houghton and Tadelis (2006), we argue that bidders payoffs are somewhat more complicated than in the above model because of change orders and cost overruns. However, we find that the method of Guerre, Perrigne and Vuong (2000) estimates bidder profits quite well. As we report below, our estimates seem sensible given what is known about bidder markups and other structural parameters.

6.0.2 The Entry Game

In the first stage of our model, bidders simultaneously and independently decide whether to bid for contract t. Submitting a bid is a costly decision. Based on extensive industry experience, Park and Chapin (1992) report that the costs of preparing a bid for projects in this industry is typically one percent of the total bid $b_{i,t}$. Publicly traded firms in the construction industry typically report profit margins of one to five percent. This implies that the fixed costs of bidding are nontrivial compared to a firm's profit margins and hence bidders should selectively submit bids on projects they are most likely to win. Let θ_i denote the cost to firm i of submitting a bid. Note that it is useful to allow the costs of bidding to vary across firms i in order to rationalize differences in participation rates across firms. As we shall discuss in the next section, the size distribution of firms in our data is quite skewed. While there are 271 firms that submit bids, a small number of these firms account for the majority of total output. In our application, we shall focus on the entry decisions of the four largest firm that each have a market share of at least five percent, as measured by winning bids. We shall denote these firms as $i = 1, 2, 3, 4$. We shall take the entry decisions of the other bidders $N(t) - \{1, 2, 3, 4\}$ as predetermined. It would obviously be preferable to endogenize the entry decisions of all bidders. However, repeatedly solving for all of the Nash equilibrium to a game with approximately three hundred players is not computationally tractable. Nonetheless, we believe that it is fairly innocuous to take the entry decisions of small, fringe firms as exogenous. Such firms rarely win large CalTrans contracts, since they lack the capital and managerial expertise to complete these large projects at a competitive price. Fringe firms typically win much smaller jobs in the public sector, such as resurfacing streets for a mid-sized California city, or smaller private sector jobs, such as resurfacing parking lots for small businesses. In our CalTrans data, these fringe firms have little influence on the winning bid and hence on profits at the margin. We believe that it is much more important to carefully model the entry decisions by the largest firms in our data set and this is where we focus our attention.

Let $a_{i,t} = 1$ if firm i decides to submit a bid on project t and $a_{i,t} = 0$ otherwise. Given $a_{i,t}$ $i = 1, \ldots, 4$ for the largest bidders, the set of bidders who participates will be denoted as $N(t|a)$. This set includes all the fringe firms observed to participate in the data and those firms $i = 1, \ldots, 4$ for which $a_{i,t} = 1$. If one of our four largest firms i enters, then conditional on a and x_t, i's profit will be:

$$u_i(a; x_t, \theta_i) = \int (\mathbf{b}_{i,t}(c_{i,t}; x_t, N(t|a)) - c_{i,t}) \prod_{j \neq i} (1 - G_j(b_{i,t}|x_t, N(t|a))) dF(c_{i,t}|x_{i,t}) - \theta_i \quad (23)$$

The above expression implicitly assumes that the timeline for the game is as follows. First, all large firms simultaneously decide whether or not to enter. The project characteristics, x_t, and the entry decisions of the fringe firms are common knowledge. Second, after entering, each of the four largest firms observes which other large firms have entered the market. Third, all participating bidders independently make their cost draws $c_{i,t}$. Finally, firms submit sealed bids and the lowest bidder wins. In the above equation, $u_i(a; x_t, \theta_i)$ is i's profits conditional on the entry decisions of the other large firms, the publicly observed data x_t and the parameter θ_i. Given $u_i(a; x_t, \theta_i)$, we can specify a normal form game as in the framework of Section 2.

6.1 Estimation

The estimation procedure that we propose is done in two steps. In the first step, we form an estimate of the term $\int(\mathbf{b}_{i,t}(c_{i,t}; x_t, N(t|a)) - c_{i,t}) \prod_{j \neq i}(1 - G_j(b_{i,t}|x_t, N(t|a)))dF(c_{i,t}|x_{i,t})$ in Equation 23 by adapting the approach proposed in Guerre, Perrigne, and Vuong (2000). In the second step, we take the estimates from the first stage and estimate θ_i, the fixed cost of preparing a bid, and λ, the selection of equilibrium, using the methods from Sections 2 and 3.

6.1.1 Markup Estimation

The basic idea behind Guerre, Perrigne and Vuong's estimator is quite simple. The left hand size of Equation 22 is the bidder's private information, $c_{i,t}$, which is unobserved to the econometrician. The right hand side is a function of the bid, $b_{i,t}$, the density of bids, $g_j(b_{i,t}|x_t)$, and the CDF of bids, $G_j(b_{i,t}|x_t)$. By pooling observations from contracts $t = 1, \ldots, T$, we construct an estimate $\widehat{g}_j(b_{i,t}|x_t)$ and $\widehat{G}_j(b_{i,t}|x_t)$ using standard nonparametric techniques. We then construct an estimate firm i's private information $\widehat{c}_{i,t}$ by evaluating the empirical analogue of the right hand side of Equation 22. Once we have recovered the distribution of a firm's private information, it is possible to compute the ex-post entry profits in Equation 23.

6.1.2 Equilibrium Selection

In the second step, we estimate the fixed costs of bidding, θ_i, and the probability that a particular equilibrium is selected, taking the expected entry profits in Equation 23 as given. We follow the approach outlined in Section 2 and use a conditional logit as a parsimonious specification of λ. Following the previous literature on entry games, we have found four criteria discussed in the literature that might influence the selection of equilibrium. First, in empirical papers such as Tamer (2002), Ciliberto and Tamer (2003) and Andrews, Berry and Jia (2005) and Jia (2006) it is usually assumed that only pure strategies are used in the entry game. The authors argue that mixed strategy equilibria are a priori implausible in these markets. There is also an experimental literature on entry in auctions. A paper by Levin and Smith (2001) argues that experimental evidence suggests that the mixed strategy equilibrium seems most reasonable in auction entry experiments. Therefore, we construct a dummy variable $MIXED(\pi)$ if the equilibrium π involves mixed strategies.

Second, we allow the selection of equilibrium to depend if the equilibrium is efficient in the sense that it maximizes joint payoffs. Therefore, we include the term from Equation 4. Maximizing joint payoffs has commonly been used to select an equilibrium in economic theory and has also been proposed in the empirical entry literature by Ciliberto and Tamer (2003). Since the firms in our data interact repeated, they obviously have incentives to tacitly collude on an equilibrium that maximizes industry surplus. Although we label this equilibrium as "efficient," it is not necessarily the equilibrium that the social planner would choose as does not account for either the revenues generated by the auction or the other fringe participants.

Third, we include a dummy variable that is equal to one if an equilibrium is Pareto dominated. In the game theory literature, it is common to assume that Pareto dominated equilibrium are less plausible and are unlikely to be observed in the data. Finally, we include the Nash product of player's utilities for pure strategy equilibria. Harsanyi and Selten (1988) argue that risk dominant equilibrium are more plausible. When an equilibrium has a large Nash product, this implies that deviating from the observed equilibrium behavior is especially costly and hence, the equilibrium is more likely to be self reinforcing.

6.2 The Data

We have constructed a unique data set of bidding by highway contractors in the State of California from 1999-2000. The data includes 414 contracts awarded by the California Department of Transportation (CalTrans) during this time period.[12] The total value of winning bids in this data was $369.2 million. There are a total of 1,938 bids and 271 general contractors in our sample. Highway improvement projects are awarded using open

[12] The data contains contracts for paving and excludes other contracts such as bridge repair. We partition the data by filtering out contracts where asphalt costs accounted for less than 1/3 of the winning bid. We focus on paving contracts since capacity constraints and the dynamics emphasized in Pesendorfer and Jofre-Bonet (2003) are less important for this set of contracts. In Bajari, Houghton and Tadelis (2006) we produce closely related structural estimates. In this paper, we adjust our estimates to allow for the presence of dynamics through non-trivial capacity constraints. We find that the inclusion of such capacity constraints has little effect on estimated markups. In order to simplify the presentation of the results, we focus on a static model of profits, although it would be quite straightforward to extend the analysis to allow for capacity to influence profit margins and markups.

Table 2: Bidder Identities and Summary Statistics

Company	Share	No. Wins	No. Bids Entered	Participation Rate	Total Bids for Contracts Awarded
Granite Construction Company	27.2%	76	244	58.9%	343,987,526
E. L. Yeager Construction Co Inc	10.4%	13	31	7.5%	132,790,460
Kiewit Pacific Co	6.6%	5	30	7.2%	112,057,627
M. C. M. Construction Inc	6.5%	2	6	1.4%	89,344,972
J. F. Shea Co Inc	3.3%	9	40	9.7%	43,030,861
Teichert Construction	3.3%	16	43	10.4%	40,177,076
W. Jaxon Baker Inc	2.9%	13	65	15.7%	37,702,631
All American Asphalt	2.2%	14	33	8.0%	30,764,962
Tullis And Heller Inc	2.1%	10	16	3.9%	27,809,535
Sully Miller Contracting Co	1.9%	17	49	11.8%	27,889,186

competitive bidding. Any qualified contractor can submit a bid and contracts are awarded to the lowest qualified bidder.[13] This data set is described in detail in Bajari, Houghton and Tadelis (2006). We will describe some of the highlights of the data and the industry in this section. The reader interested in a more complete description can consult this reference.

Let $i = 1, \ldots, N$ denote a bidder in our data set and $t = 1, \ldots, T$ denote a contract. For each contract, we observe a detailed list of covariates including $b_{i,t}$, the bid of contractor i on project t, EST_t, the engineer's estimate for project t, $DIST_{i,t}$ the distance in miles of firm i to project t, $CAP_{i,t}$, the capacity utilization of firm i at the time of bidding for project t, and $FRINGE_{i,t}$, a dummy variable equal to one if firm i is a fringe firm, defined as firms with market shares of less than one percent. The data set includes the bids for all contractors, not just the winning bids. The engineer's estimate, EST_t, is constructed by CalTrans and is meant to represent a fair market value for completing the work. The project plans and specifications contain an exhaustive list of work items. The estimate is formed using blue book prices for specific work items and local material prices.

In Table 2, we summarize the market shares of the 10 largest firms in the industry, where share is defined using the winning bids.[14] The market shares in this industry are quite skewed. The largest firm, Granite Construction Company, has a share of 27.2 percent compared to a share of 1.9 percent for the 10th largest firm, Union Asphalt. This skewed distribution suggests that productivity varies across firms and hence it is important to include firm fixed effects in our estimates of g_j and G_j.

Table 2 demonstrates that the largest firms tend to bid more often as measured by their participation rate. However, we note that the second largest firm only submits bids for 7.5

[13]In about 5 percent of the projects in our sample, CalTrans rejects all bids and lets the contracts at a later date. We do not include these contracts in our sample.

[14]Market share is defined using the total bids which may differ from final payments for reasons that we describe in Bajari, Houghton and Tadelis (2006).

percent of the jobs compared to Granite which submits bids for 58.9 percent of the jobs. This suggests that in modeling entry, it will be important to account for firm specific differences in the costs of bidding, θ_i.

In Table 3, we provide some summary statistics about the bids. In our data, a typical winning bid is \$3.2 million dollars and is about 6 percent lower than the engineer's estimate. Comparing the winning bid to the second highest bid, the average money left on the table is about 6 percent of the estimate. This suggests that there is asymmetric information in this market. If the low bidder knew the cost of the second lowest bidder, then in a Nash equilibrium we would expect the low bidder to shade just under the cost of the second lowest bidder. Leaving money on the table does not increase the probability of winning and only decreases the profit of the low bidder.

In Table 4, we see that the the ranking of the bids corresponds closely to the distance of the participating contractors from the project. For instance, $DIST1$, the distance of the closest contractor is smaller than $DIST2$, the distance of the second closest contractor. The average time of the closest contractor to the project is 47 miles compared to 73 miles for the second closest contractor. This is consistent with the presence of substantial transportation costs. The closest contract will, all else held constant, have a lower cost of hauling asphalt to the project site and is therefore more likely to win the project.

In Table 5, we regress the bids on the various cost controls in our data set. In the first column, a regression of all total bids on the engineer's estimate has an R^2 of 0.987 with a coefficient of 1.02 suggesting that the estimate accounts for much of the observed variation in bids and is therefore an extremely useful regressor. In the second column, we change the dependent variable to $b_{i,t}/EST_t$ since the variance of the errors in the bid regressions are likely to be proportional to EST_t. The next set of regressions demonstrate that distance, an identifier for fringe firms, project fixed effects, and firm fixed effects for the largest four firms are significant determinants of the observed bids. The expected signs are as anticipated. The impact of distance on the bid is positive, reflecting the higher transportation cost for firms that are farther away from the project. The average distance of a firm from the project in our data is 72 miles with a standard deviation of 92 miles. The regression results imply that increasing the distance by a standard deviation will raise the bid, relative to the engineer's estimate, by about 2.3 percent. Highway contracting is widely viewed as quite competitive. The largest firm in our sample, Granite Construction, reports a profit margin of 3.31 percent and an operating margin of 5.15 percent. Therefore, a two percent increase in bids is fairly substantial. Relative to the engineer's estimate, fringe firms on average bid 4.2 percent higher than non-fringe firms.

In Table 6, we estimate a logit model of entry for the four largest firms in the industry.

Table 3: Bidding Summary Statistics

	Obs	Mean	Std. Dev.	Min	Max
Winning Bid	414	3,203,130	7,384,337	70,723	86,396,096
Markup: (Winning Bid-Estimate)/Estimate	414	-0.0617	0.1763	-0.6166	0.7851
Normalized Bid: Winning Bid/Estimate	414	0.9383	0.1763	0.3834	1.7851
Second Lowest Bid	414	3,394,646	7,793,310	84,572	92,395,000
Money on the Table: Second Bid-First Bid	414	191,516	477,578	68	5,998,904
Normalized Money on the Table: (Second Bid-First Bid)/Estimate	414	0.0679	0.0596	0.0002	0.3476
Number of Bidders	414	4.68	2.30	2	19
Distance of the Winning Bidder	414	47.47	60.19	0.27	413.18
Travel Time of the Winning Bidder	414	56.95	64.28	1.00	411.00
Utilization Rate of the Winning Bidder	414	0.1206	0.1951	0.0000	0.9457
Distance of the Second Lowest Bidder	414	73.55	100.38	0.19	679.14
Travel Time of the Second Lowest Bidder	414	82.51	97.51	1.00	614.00
Utilization Rate of the Second Lowest Bidder	414	0.1401	0.2337	0.000	0.9959

Table 4: Distance to Job Site

	Mean	Std. Dev.	Min	Max
DIST1	47.47	60.19	0.27	413.18
DIST2	73.55	100.38	70.19	679.14
DIST3	75.47	95.56	0.13	594.16
DIST4	84.38	89.87	1.45	494.08
DIST5	76.12	86.33	1.25	513.31

Table 5: Bid Function Regressions

Variable	$b_{i,t}$	$b_{i,t}/EST_t$	$b_{i,t}/EST_t$	$b_{i,t}/EST_t$
EST_t	1.025			
	(56.26)			
$DIST_{i,t}$		.000246	.000223	
		(5.66)	(5.01)	
$UTIL_{i,t}$			0.02539	
			(0.93)	
$FRINGE_{i,t}$				0.4288
				(4.65)
Constant	-25686	1.19	1.001	
	(0.56)	(94.9)	(79.98)	
Fixed Effects	No	Project	Project	Project/Firm
R^2	0.989	0.5245	0.5292	0.5321

Number of observations = 1938; t-statistics are reported in parentheses.

Table 6: Logit Model of Entry

	I	II	III
Constant	-.9067	-1.6811	
	(7.91)	(7.53)	
$DIST_{i,t}$	-0.00218	-0.00322	-0.00854
	(5.42)	(5.66)	(4.85)
Granite		2.889	4.4537
		(13.28)	(7.31)
E. L. Yeager		-	-
Kiewit Pacific		-0.1527	1.1969
		(0.57)	(2.1)
M. C. M.		-1.786	-.70779
		(3.94)	(1.12)
Fixed Effects	No	No	Project
Observations	1656	1656	1068
Number of Groups			261
Log-Likelihood	-784.20	-511.86	-101.0728

The dependent variable is whether one of the four largest firms in the industry decides to submit a bid in a particular procurement; t-statistics are reported in parentheses.

For each of these firms, we calculate their distance to each project t even the firm does not submit a bid. We find that participation is a decreasing function of the firm's distance to the project. Also, there is heterogeneity across the firms in terms of their participation decisions suggesting that inclusion of firm level effects θ_i is important in modeling entry.

6.2.1 Estimates of Profits

We estimate bidder markups using the approach by Guerre, Perrigne and Vuong (2000) as described above. Given the number of covariates in our application, it is not feasible to non-parametrically estimate the distribution of bids g_j and G_j. Instead, we use a semiparametric approach. We first run a regression, as in Table 5:

$$\frac{b_{i,t}}{EST_t} = x'_{i,t}\alpha + u^{(t)} + \epsilon_{i,t},$$

where the dependent variable is normalized by dividing through by the engineer's estimate. We also include an auction specific fixed effect, $u^{(t)}$. Let $\widehat{\alpha}$ denote the estimated value of α and let $\widehat{\epsilon}_{i,t}$ denote the fitted residual. We will assume that the residuals to this regression are iid. Let $\widehat{H}$ denote the Kaplan-Meier estimate of the CDF of the fitted residuals.[15]

Under these assumptions, the estimated bid distributions satisfy:

$$\begin{aligned}\widehat{G}_i(b|z_{j,t}, N(t)) &= \Pr\left(\frac{b_{i,t}}{EST_t} \leq \frac{b}{EST_t}\right) \\ &= \Pr\left(x'_{i,t}\widehat{\alpha} + \widehat{u}^{(t)} + \widehat{\epsilon}_{i,t} \leq \frac{b}{EST_t}\right) \\ &= \widehat{H}\left(\widehat{\epsilon}_{i,t} \leq \frac{b}{EST_t} - x'_{i,t}\widehat{\alpha} - \widehat{u}^{(t)}\right).\end{aligned}$$

That is, the distribution of the fitted residuals, $\widehat{\epsilon}_{i,t}$ can be used to to infer the distribution of the bids. As the estimates in Table 5 suggest, variation in the estimated bid distribution will be driven by three factors. The first is the auction fixed effects, $\widehat{u}^{(t)}$. The second is the distance of each firm from the project. The further a particular firm is away from the project, the higher its bid will be. Finally, the firm fixed effects are important. The largest four firms will bid more aggressively than the smaller fringe firms.[16] Recall that earlier

[15]We estimate the density h of the fitted residuals using kernel density estimation with an estimated optimal bandwidth. Since there are 1938 fitted residuals, the estimates of H and h are quite precise given θ. Ideally, our estimates would take account of the first stage estimation error in θ. However, the computational burden of performing a resampling procedure such as the bootstrap is considerable and beyond the scope of this research.

[16]We note that we must estimate the distribution of each firm i's bid even if it does not participate in a particular auction. We have therefore computed the distance of each of the four largest firms from all

Table 7: Margin Estimates

Variable	Num. Obs.	Mean	Std. Dev.	Median	25th Percentile	75th Percentile
Profit Margin	1938	0.0644	0.1379	0.0271	0.0151	0.0520

The markup is defined as 1 minus the ratio of the estimated cost, which is private information, to the actual bid.

we demonstrated that firm specific profit shifters are sufficient to identify our model under fairly mild parametric assumptions. In our analysis, distance and firm fixed effects will be the primary shifters of individual firms profits. Each firm has a unique distance to a particular contract t. The variation in commuting times across projects generates shifts in the payoffs of individual firms and therefore allows us to identify our model.

In Table 7, we summarize the distribution of estimated markups on the 1,938 bids in our data set. The average markup is about 6 percent. However, the distribution of markups is skewed. The median markup is estimated at 2.71 percent and the 75th percentile is 5.2 percent. These margins seem sensible given the publicly available information about this market. The largest firm in our data set, Granite Construction is publicly traded under the symbol GVA. In the first quarter of 2006, they reported a profit margin of 3.31 percent and an operating margin of 5.15 percent. Margins reported in quarterly statements and our margin estimates are not directly comparable since accounting costs and economic costs differ. However, the publicly available information suggest that a reasonable answer should involve a fairly low profit margin. The above margin estimates are very similar to estimates constructed in Bajari, Houghton and Tadelis (2006). For a discussion of robustness checks on the margin estimates to alternative covariates and alternative econometric modeling assumptions, the interested reader can consult this reference.

6.2.2 Equilibrium Selection Parameters and Bid Costs

As mentioned in Section 5, a good choice of the importance density is vitally important for the success of our estimation procedure. To this end, we first estimated a private-information version of the entry game to obtain starting values for bid preparation costs and the profit scale parameter. While these parameters will generally not be consistent estimates of the complete information game parameters, they will be roughly in the correct neighborhood, which greatly enhances the convergence properties of our importance sampling MSM procedure. All equilibrium selection parameters were initially set at zero. We performed several

$t = 1, \ldots, 414$ projects even if they did not submit a bid. We use Equation 22 to infer the distribution of firm i's bid in this case. The estimates of Table 6 suggest that bidder i will have a low chance of winning a particular procurement t if it is a long distance from the project site.

Table 8: Games Estimation Results

Variable	Mean	Median	Std.Dev.	95% Confidence Interval	
		Equilibrium Selection Parameters (λ)			
Pure Strategy	-1.3524	-1.5345	0.7979	-2.4903	0.1954
Joint Profit Maximizing	6.4365	6.4226	0.5321	5.6151	7.5149
Dominated	-5.3841	-5.3316	0.7002	-6.7164	-4.0986
Nash Product	4.4143	4.2025	1.1017	2.9651	6.4836
		Profit Scale			
Profit Scale	0.0965	0.0954	0.0015	0.0954	0.0984
		Bid Preparation Costs (θ_i)			
Granite Construction	0.2341	0.2393	0.0977	0.0679	0.4271
E. L. Yeager	1.4583	1.4757	0.0941	1.2563	1.6227
Kiewit Pacific	1.6751	1.6720	0.0511	1.5775	1.7789
M. C. M. Construction	2.4490	2.4360	0.1144	2.2547	2.6966

Estimation and inference was performed using the LTE method of Chernozhukov and Hong (2003). A Markov chain was generated with 500 draws for each parameter. 409 importance games were used in the importance sampler for the 409 observations.

sequential estimations, using 409 importance games that were initialized with the last iteration's final value. When the parameter values were consistent after several iterations, we ran the Laplace-type estimator of Chernozhukov and Hong (2003) to generate standard errors and to ensure that the estimated coefficients were robust to the optimization method used in the initial steps. The results are reported in Table 8.

The first parameter to interpret is the coefficient on profit scale. The expected entry profits for a firm are denominated in tens of thousands of dollars. These expected profits are then multiplied by the profit scale parameter, which is equal to 0.0965. Therefore we should interpret each unit of the fixed costs of bidding as representing about $96,500. For example, this means that Kiewit Pacific faces, on average, a cost of approximately $161,650 to prepare a bid, or roughly five percent of the average winning bid. This amount varies across the four firms, from $22,590 for Granite Construction to $235,460 for M. C. M. Construction. Given that winning bids are drawn from the left tail of the bid distribution, these numbers are roughly consistent with Park and Chapin (1992), who argue that bid preparation costs are approximately equal to one percent of the total bid in magnitude.

A second check on the validity of these estimates is to consider that the bid preparation costs have to reconcile the rates of entry into these auctions, suggesting they should be significant relative to the size of the expected profits. The typical expected profit if a firm is the only one of the big four to enter is about $50,000. In a simple probit entry model, we would expect a firm with a participation rate of 50 percent to have fixed costs equal to this value. Indeed, the Granite Construction Company, with the the highest participation rate by far at over 58 percent, has fixed costs slightly lower than this. The fact that the bid costs are lower than $50,000 reflects both that the entry rate is above 50 percent and that in most auctions several other large firms are going to enter, which decreases the expected profits below $50,000. A third check on the validity of our results is that bid costs are monotonically and inversely related to participation rates: the two firms with similar participation rates, E. L. Yeager and Kiewit Pacific, have almost identical bid preparation costs, and the bid costs for M. C. M. Construction, with a low entry rate of 1.4 percent, are correspondingly high, over ten times that of Granite Construction.

Turning to the parameters of the equilibrium selection mechanism, we have several interesting results. First of these is that mixed strategy equilibria are more preferred to pure strategy equilibria, all else equal. In entry games, this result is somewhat intuitive; supposing that there is room for one firm in a game with two potential entrants, it would be surprising to see one of the two firms consistently conceding the market when committing to the mixed strategy strictly increases expected profits. This also dovetails with the results in Levin and Smith (2001), who find support for mixed strategy equilibria in an experimental setting. This results is particularly important in light of the number of papers which assume that only pure strategies are played in equilibrium; our findings suggest that it may be common to obtain biased parameter estimates under these assumptions.

Also surprising is the strong effect that efficiency has on the probability of an equilibrium being chosen. Given that there are typically many pure and mixed strategies in a given game, this shifter is by far the most influential factor in deciding which equilibrium is played. This finding also has potentially collusive implications, as one would expect firms coordinating their entry actions to operate in such a fashion that the equilibrium played in the data maximized joint profits.

A less surprising result is that the coefficient on dominated equilibria is strongly negative. This is also in line with intuition, as we would expect firms to stay away from equilibria where everyone could do at least as well in expectation by playing a different equilibrium. Finally, the coefficient on the equilibrium with the highest Nash product is strongly positive. This indicates that among pure strategy equilibria, the one with the highest Nash product has a strong tendency to be played in the data.

7 Conclusion

Estimating models that are consistent with Nash equilibrium behavior is an important empirical problem. In this paper, we have developed algorithms that can be used to estimate both the utilities and the equilibrium selection parameters for static, discrete games. Our algorithms, unlike previous research, can be applied to general normal form games, not just specific examples such as entry games. The algorithms use computationally efficient methods and our Monte Carlo work suggests that they may work well even with a moderate number of observations.

We also study the nonparametric identification of these games. We provide approaches for identification based on sufficient variation of payoff covariates and identification through two types of exclusion restrictions. First, we can obtain identification if there are variables that influence equilibrium selection but do not directly enter into payoffs. Second, we can obtain identification if there are variables that: a.) shift a specific agent's utility but which do not enter into the utility of other players, and b.) can be excluded from the equilibrium selection mechanism.

As an application of our methods, we studied the decision of four large construction firms to enter into procurement auctions in California. We recovered fixed bid preparation costs for each of the four firms which rationalize their entry rates into these auctions. The application also highlighted one of the strengths of our approach: the ability to estimate an equilibrium selection mechanism. Our estimates indicate that mixed strategy equilibria are selected with a greater probability than pure strategy equilibria, an important finding given the prevalence of empirical models which exclude the possibility of mixed strategy equilibria. We also find that the equilibrium mechanism favors joint profit maximizing and non-dominated equilibria. Among pure strategy equilibria, the one with the largest Nash product is selected with higher probability. The estimation methods we propose to the best of our knowledge are the only current approach capable of accommodating both multiplicity and mixed strategy equilibria.

References

Ackerberg, D. and G. Gowrisankaran (2006). Quantifying equilibrium network externalities in the ach banking industry. Forthcoming in RAND.

Ackerberg, D. A. (2004). Importance sampling and the method of simulated moments. UCLA Working Paper.

Aguirregabiria, V. and P. Mira (2004). Sequential simulation-based estimation of dynamic discrete games. Boston University Working Paper.

Andrews, D. W. K., S. Berry, and P. Jia (2005). Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location. Yale University Working Paper.

Bajari, P., C. L. Benkard, and J. Levin (2006). Estimating dynamic models of imperfect competition. Forthcoming in Econometrica.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2006). Estimating static models of strategic interactions. University of Minnesota Working Paper.

Bajari, P. and A. Hortacsu (2003, Summer). The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics 34*, 329–55.

Bajari, P., S. Houghton, and S. Tadelis (2006, February). Bidding for incomplete contracts: An empirical analysis. NBER Working Papers 12051, National Bureau of Economic Research, Inc.

Bajari, P. and J. Krainer (2004). An empirical model of analysts' stock recommendations: Market fundamentals, conflicts of interest, and peer effects. Duke University Working Paper.

Bajari, P. and L. Ye (2003, November). Deciding between competition and collusion. *The Review of Economics and Statistics 85*(4), 971–989.

Berg, J. and A. McLennan (2002). The asymptotic expected number of nash equilibria of two player normal form games. University of Minnesota Working Paper.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in equilibrium. *Econometrica 63*(4), 841–90.

Berry, S., A. Pakes, and M. Ostrovsky (2005, June). Simple estimators for the parameters of dynamic games (with entry/exit examples). Harvard University Working Paper.

Berry, S. and P. Reiss (2006). Empirical models of entry and exit. working paper, for Vol 3 of the Handbook of IO.

Berry, S. and E. Tamer (2006). Identification in models of oligopoly entry. In Blundell, Newey, and Persson (Eds.), *Advances in Economics and Econometrics, Vol 2, Ninth World Congress, 46-85.* Cambridge University Press.

Berry, S. T. (1992, July). Estimation of a model of entry in the airline industry. *Econometrica 60*, 889–917.

Bjorn, P. and Q. Vuong (1984). Simultaneous equations models for dummy endogenous variables: A game theoretic formulation with application to labor force participation. SSWP No. 537, California Institute of Technology.

Bresnahan, T. F. and P. C. Reiss (1990). Entry in monopoly markets. *Review of Economic Studies 57*, 531–553.

Bresnahan, T. F. and P. C. Reiss (1991). Empirical models of discrete games. *Journal of Econometrics 48*, 57–81.

Brock, W. and S. Durlauf (2001). Discrete choice with social interactions. *Review of Economic Studies 62*(2), 235–260.

Brock, W. and S. Durlauf (2003). Multinomial choice with social interactions. University of Wisconsin Working Paper.

Chernozhukov, V. and H. Hong (2003). A MCMC approach to classical estimation. *Journal of Econometrics 115*(2), 293–346.

Chiappori, P.-A., S. Levitt, and T. Groseclose (2002). Testing mixed strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review 92*(4), 1138–1151.

Ciliberto, F. and E. Tamer (2003). Market structure and multiple equilibria in airline markets. Princeton and North Carolina State Working Paper.

Davis, P. (2005). Estimation of quantity games in the presence of indivisibilities and heterogeneous firms. Forthcoming in Journal of Econometrics.

El-Gamal, M. A. and D. M. Grether (1995). Are people bayesian? uncovering behavioral strategies. *Journal of the American Statistical Association 90*(432), 1137–1145.

Guerre, E., I. Perrigne, and Q. Vuong (2000, May). Optimal nonparametric estimation of first-price auctions. *Econometrica 68*(3), 525–574.

Harsanyi, J. and R. Selten (1988). *A General Theory of Equilibrium Selection in Games.* MIT Press, Cambridge, MA.

Heckman, J. (1990). Varieties of selection bias. *American Economic Review 80*(2), 313–318.

Holmes, T. J. (2006). The diffusion of wal-mart and economies of density. University of Minnesota Working Paper.

Jia, P. (2006). What happens when wal-mart comes to town: An empirical analysis of the discount retail industry. MIT Working Paper.

Jofre-Bonet, M. and M. Pesendorfer (2003). Estimation of a dynamic auction game. *Econometrica 71*(5), 1443–1489.

Keane, M. and K. Wolpin (1997). The career decisions of young men. *The Journal of Political Economy 105*(3), 473–522.

Keane, M. and K. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), 1051–1103.

Kooreman, P. and A. Soetevent (2006). A discrete choice model with social interactions, with an application to high school teen behavior. forthcoming, Journal of Applied Econometrics.

Krasnokutskaya, E. and K. Seim (2005). Bid preference programs and participation in highway procurement auctions. University of Pennsylvania Working Paper.

Lebrun, B. (1996). Existence of an equilibrium in first-price auctions. *Economic Theory 7*, 421–423.

Levin, J., S. Athey, and E. Seira (2006). Comparing open and sealed bid auctions: Theory and evidence from timber auctions. Stanford University Working Paper.

Li, T. and X. Zheng (2006). Entry and competition effects in first-price auctions: Theory and evidence from procurement auctions. Vanderbilt University Working Paper.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*(3), 531–42.

Manuszak, M. and A. Cohen (2004). Endogenous market structure with discrete product differentiation and multiple equilibria: An empirical analysis of the competition between banks and thrifts. Carnegie Mellon Working Paper.

Maskin, E. and J. Riley (2000). Equilibrium in sealed high bid auctions. *Review of Economic Studies 67*, 439–454.

Mazzeo, M. (2002, Summer). Product choice and oligopoly market structure. *RAND Journal of Economics 33*(2), 221–242.

McKelvey, R. and A. McLennan (1997). The maximal number of regular totally mixed nash equilibria. *Journal of Economic Theory 72*, 411–425.

McKelvey, R. and T. Palfrey (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior 10*, 6–38.

McLennan, A. (1997). The maximal generic number of pure nash equilibria. *Journal of Economic Theory 72*, 408–410.

McLennan, A. (2002). The expected number of nash equilibria of a normal form game. Forthcoming in Econometrica.

McLennan, A. and R. McKelvey (1996). The computation of equilibrium in finite games. In H. Amman, D. A. Kendrick, and J. Rust (Eds.), *The Handbook of Computational Economics*, pp. 87–142. Elsevier.

Nash, J. (1951). Non-cooperative games. *The Annals of Mathematics 54*(2), 286–295.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2005). Moment inequalities and their application. working paper, Harvard University.

Park, W. R. and W. B. Chapin (1992, October). *Construction Bidding: Strategic Pricing for Profit* (2nd ed.). Wiley.

Pesendorfer, M. and P. Schmidt-Dengler (2003). Identification and estimation of dynamic games. LSE Working Paper.

Porter, R. H. and J. D. Zona (1999, Summer). Ohio school milk markets: An analysis of bidding. *RAND Journal of Economics 30*(2), 263–288.

Rojas, M. (1996). Roots of certain random sparse polynomial systems. In J. Reneger, M. Shub, and S. Smale (Eds.), *The Mathematics of Numerical Analysis*, Volume 32 of *Lectures in Applied Mathematics*, pp. 689–699. American Mathematical Society.

Ryan, S. P. and C. E. Tucker (2006). Diversification and the dynamics of technology adoption. MIT Working Paper.

Seim, K. (2001). Spatial differentiation and market structure: the video retail industry. Yale University Ph.D. Dissertation.

Shachat, J. and M. Walker (2004). Unobserved heterogeneity and equilibrium: an experimental study of bayesian and adaptive learning in normal form games. *Journal of Economic Theory 114*, 280–309.

Smith, J. L. and D. Levin (2001). Entry coordination in auctions and social welfare: An experimental investigation. *International Journal of Game Theory 30*, 321–350.

Stern, S. and M. Engers (2002). Long-term care and family bargaining. *International Economic Review 43*(1), 73–114.

Stern, S. and B. Heidemann (1999). Strategic play among family members when making long-term care decisions. *Journal of Economic Behavior and Organization 40*(1), 29–57.

Sweeting, A. (2006). Coordination games, multiple equilibria, and the timing of radio commercials. Northwestern University Working Paper.

Tamer, E. (2002). Incomplete bivariate discrete response model with multiple equilibria. *Review of Economic Studies 70*, 147–167.

A Maximal and Expected Number of Nash Equilibria

When considering the nonparametric identification of the equilibrium selection mechanism, it is useful to know some generic properties of the set of Nash equilibria in a game. Briefly we review results in the literature on the maximum number and expected number of equilibria to normal form games of the class considered here.

A.1 Maximum Number of Equilibria

Solutions to normal form games can be characterized using polynomial equalities and inequalities. Therefore before considering games we review some general important results discussing the solutions of the systems of polynomials. Let $F = \{f_i(x)\}_{i=1}^n$ be the system of n polynomials of n variables and we are looking for the set of all common roots of this system. A polynomial $f_i(x) = \sum_{j=1}^{J} a_j\, x_1^{e_1^{ij}}\, x_2^{e_2^{ij}} \ldots x_n^{e_n^{ij}} = \sum_{j=1}^{J} a_j \prod_{k=1}^{n} x_k^{e_k^{ij}}$. Here the powers e_k^{ij} are in general integers and index i refers to the number of equation, index j refers to the number of monomial in the polynomial i and index k refers to the specific variable x_k. The points $e^{ij} = \left(e_1^{ij}, \ldots, e_n^{ij}\right)$ form the finite sets $E_i = (e^{ij}, j = 1, \ldots, J)$ indicate which monomial terms appear in f_i. For instance, in the polynomial $f_i(x_1, x_2) = x_1^2\, x_2^3 + 2x_1^2$ the support set is $E_i = \{(2,\, 3)\,,\; (2,\, 0)\}$.

The collection of sets $E = (E_1,\, E_2, \ldots, E_n)$ is called the support of the system of polynomials. The convex hulls $Conv\,(E_i)$ are called Newton polytopes of f_i. For example the Newton polytope of the polynomial $f(x_1, x_2) = x_1\, x_2 + x_1 + x_2 + 1$ is the unit square with vertices in $(1, 1)$, $(1, 0)$, $(0, 1)$, and $(0, 0)$.

The degree of the polynomial i is $d_i = \max_j \sum_{k=1}^{n} e_k^{ij}$. One of the most important theorems describing the behavior of zeros of F in the complex space $\mathbb{C}^n$ is Bézout's theorem which says that the total number of common complex roots of F is at most $\prod_{i=1}^{n} d_i$. Bézout's theorem gives an upper bound on the number of common roots in the system. A drawback of Bézout's theorem is that it gives little information about the polynomials that are sparse. In fact for sparse systems the number of common roots of the polynomial system can be significantly less than the bound set by this theorem. A universal and powerful tool for root counting in case of sparse polynomial is the Bernstein's theorem.

Let P_i be Newton polytopes of equations $f_i(x)$ in the system F defined previously. The mixed volume of the system of polytopes is defined as:

$$\mathcal{M}\left(P_1,\ldots,P_n\right)=\sum_{S\subseteq\{1,\ldots,n\}}(-1)^{|S|}\,\mathrm{Vol}\left(\sum_{i\in S}P_i\right), \tag{24}$$

where S are all subsets of $\{1,2,\ldots,n\}$, $|S|$ is the cardinality (number of elements) of a particular subset, while $\mathrm{Vol}\left(\cdot\right)$ is the conventional geometric volume. The sum of the polytopes is defined for two polytopes P and Q as $P+Q=\{p+q\,|\,p\in P,\ q\in Q\}$.

Theorem 6 (Bernstein)**.** *The number of common roots in the system F is equal to the mixed volume of the n Newton polytopes of this system.*

This is an extremely powerful result because the mixed volume is easy to compute. A general problem with the complex roots though is that it is not invariant with respect to the group of polynomial transformations of F. For example, if the polynomial $f(x)$ has degree d and thus has d distinct complex roots then the polynomial $f(x)^2$ can have $2d$ distinct complex roots. This is not the case with the real roots of the system of polynomials and thus power transformations have no effect on the number of distinct real roots. This effect is captured by Khovanskii's theorem which sets the upper bound on the number of common real roots of the polynomial system which does not depend on the degrees of the equations in the system.

Theorem 7 (Khovanskii)**.** *If m is the number of all monomials in F (or equivalently $m=|E|=\sum_{i=1}^{n}J-i$ - cardinality of the support) and in F there are n polynomials then the maximum number of real solutions of the system is $2^{\left(\binom{m}{2}\right)}\left(n+1\right)^m$.*

In many cases, however, the so-called Kouchnirenko's conjecture holds: if the number of terms in f_i is at most m_i then the number of isolated real roots is at most $\prod_{i=1}^{n}(m_i-1)$. This conjecture is violated for some generic (although quite complex) counterexamples.

Consider an arbitrary N-person game where the player i has n_i strategies. Using the Lagrangian multiplier techniques it can be reduced to the system of $n+\sum_i n_i$ polynomial equations with $n+\sum_i n_i$ unknowns. Let the variable $x_k^{(i)}$ denote the strategy k of the player i, $\xi^{(i)}_{j_1,j_2,\ldots,j_{i-1},k,j_{i+1},\ldots,j_N}$ be the payoff function, representing the payoff of player i when she plays the pure strategy k and the other players are playing $j_1,\ldots,j_N$, and $\pi^{(i)}$ be the expected payoff of player i. Let $\lambda_{k0}^{(i)}$ be the lagrange multiplier for the constraint $x_k^{(i)}\geq 0$, and $\widetilde{\lambda}^{(i)}$ be the lagrange multiplier for the constraint $\sum_{k=1}^{n_i}x_k^{(i)}=1$. The Lagrangian for the bidder i can be written as:

$$\mathcal{L}^{(i)} = \sum_{k=1}^{n_i} x_k^{(i)} \sum_{j_{-i}} \xi^{(i)}_{j_1,j_2,\dots,j_{i-1},k,j_{i+1},\dots,j_N} x^{(1)}_{j_1} \dots x^{(i-1)}_{j_{i-1}} x^{(i+1)}_{j_{i+1}} \dots x^{(N)}_{j_N}$$
$$- \sum_{k=1}^{n_i} \lambda^{(i)}_{k0} x_k^{(i)} + \widetilde{\lambda}^{(i)} \left(1 - \sum_{k=1}^{n_i} x_k^{(i)}\right).$$

The first order condition for the Lagrangian is complemented by the complementary slackness conditions for the multipliers $\lambda^{(i)}_{k0}$:

$$\frac{\partial \mathcal{L}^{(i)}}{\partial x_k^{(i)}} = \sum_{j_{-i}} \xi^{(i)}_{j_1,j_2,\dots,j_{i-1},k,j_{i+1},\dots,j_N} x^{(1)}_{j_1} \dots x^{(i-1)}_{j_{i-1}} x^{(i+1)}_{j_{i+1}} \dots x^{(N)}_{j_N} - \lambda^{(i)}_{k0} - \lambda^{(i)} = 0, \tag{25}$$

$$1 - \sum_{k=1}^{n_i} x_k^{(i)} = 0, \tag{26}$$

$$x_k^{(i)} \lambda^{(i)}_{k0} = 0, \quad k = 1, \dots, n_i. \tag{27}$$

If we multiply the first equation by x_k^i then it reduces to:

$$x_k^{(i)} \sum_{j_1,j_2,\dots,j_{i-1},j_{i+1},\dots,j_N} \xi^{(i)}_{j_1,j_2,\dots,j_{i-1},k,j_{i+1},\dots,j_N} x^{(1)}_{j_1} \dots x^{(i-1)}_{j_{i-1}} x^{(i+1)}_{j_{i+1}} \dots x^{(N)}_{j_N} - x_k^{(i)} \widetilde{\lambda}^{(i)} = 0, \tag{28}$$

$$1 - \sum_{k=1}^{n_i} x_k^{(i)} = 0, \tag{29}$$

$$x_k^{(i)} \lambda^{(i)}_{k0} = 0, \quad k = 1, \dots, n_i. \tag{30}$$

Summing the first equation over all k we obtain that $\pi^{(i)} = \widetilde{\lambda}^{(i)}$. Then each bidder is characterized by the system of equations:

$$x_k^{(i)} \left(\pi^i - \sum_{j_1,j_2,\dots,j_{i-1},j_{i+1},\dots,j_N} \xi^{(i)}_{j_1,j_2,\dots,j_{i-1},k,j_{i+1},\dots,j_N} x^{(1)}_{j_1} \dots x^{(i-1)}_{j_{i-1}} x^{(i+1)}_{j_{i+1}} \dots x^{(N)}_{j_N} \right) = 0, \tag{31}$$

$$\sum_{k=1}^{n_i} x_k^{(i)} - 1 = 0, \quad k = 1, \dots, n_i, \quad i = 1, \dots, N. \tag{32}$$

For each player thus we have $n_i + 1$ equations and $n_i + 1$ unknown parameters (n_i mixed strategies and the expected payoff). The individual equation has $\prod_{j \neq i} n_j + 1$ terms (the number of strategies of the other players when the strategy of the player i is fixed plus the expected payoff of player i). In addition, the linear equations limiting the mixed strategies to the simplex have n_i+1 terms each. In total there are $\sum_i n_i + N$ equations and unknowns. The total number of terms is $\sum_i \prod_{j\neq i} n_j + \sum_i n_i + 2N$. If we consider purely mixed strategies then $x_k^{(i)} > 0$, and thus the system can be rewritten as:

$$\sum_{j_{-i}} \left(\xi^{(i)}_{j_1,j_2,\ldots,j_{i-1},k,j_{i+1},\ldots,j_N} - \xi^{(i)}_{j_1,j_2,\ldots,j_{i-1},n_i,j_{i+1},\ldots,j_N} \right) x^{(1)}_{j_1} \ldots x^{(i-1)}_{j_{i-1}} x^{(i+1)}_{j_{i+1}} \ldots x^{(N)}_{j_N} = 0, \quad (33)$$

$$k = 1, \ldots, n_i - 1, \quad i = 1, \ldots, N. \quad (34)$$

This system has $n_i - 1$ unknowns for player i and $\sum_i n_i - N$ unknowns in total. The number of terms for each equation is $\prod_{j\neq i} n_j$ according to the number of strategies of the rival players when the strategy of the given player is fixed. The total number of terms is then given by the sum $\sum_i \left(\prod_{j\neq i} n_j\right)(n_i - 1)$.

McKelvey and McLennan directly apply Bernstein's theorem to the given system of equations and express the number of solutions in terms of the mixed volume of Newton polytopes for the case of totally mixed solutions (the case with possible pure strategies needs specific consideration for each payoff structure). We can also apply Khovanskii's result to this system. First, by Kouchnirenko's conjecture the maximum number of solutions to this system is

$$\prod_{i=1}^{N} \left(\prod_{j\neq i} n_j - 1 \right)^{n_i - 1} \quad (35)$$

which gives an approximate formula for the upper bound on the number of solutions. By the exact application of Khovanskii's formula as $m = \sum_i \left(\prod_{j\neq i} n_j\right)(n_i - 1)$ the maximum number of solutions is $2^{\frac{m!}{2(m-2)!}} \left(\sum_i n_i - N + 1\right)^m$.

For a particular case of k equal number of strategies of players Kouchnirenko's formula gives $\left(k^{N-1} - 1\right)^{N(k-1)}$ for the number of equilibria while Khovanskii's bound is

$$2^{\binom{Nk^{(N-1)}(k-1)}{2}} \left[N(k-1)+1\right]^{Nk^{(N-1)}(k-1)}.$$

The number of moments which arise from the game with N players where each player has k strategies is $k^N - 1$. The corresponding numbers of moments are tabulated in Table 10. The number of moments is significantly smaller than Kouchnirenko's bounds.

A.2 Expected number of equilibria

McLennan applies a general formula obtained in Rojas (1996) which characterizes the expected number of solutions to sparse systems of polynomials with random coefficients. It is important to note that the result in Rojas (1996) refers to the number of complex roots.

Table 9: Tabulation of Kouchnirenko's formula

	Number of Strategies (k)				
N	2	3	4	5	6
2	1	16	729	65536	9765625
3	27	262144	38443359375	3.65203E+16	1.44884E+23
4	2401	2.08827E+11	3.90919E+21	3.12426E+33	4.45419E+46
5	759375	1.07374E+19	1.25344E+36	8.01109E+55	6.40832E+77
6	887503681	4.03445E+28	1.50578E+54	7.4656E+83	5.2603E+116

Table 10: Tabulation of the Number of Available Moments

	Number of Strategies (k)				
N	2	3	4	5	6
2	3	8	15	24	35
3	7	26	63	124	215
4	15	80	255	624	1295
5	31	242	1023	3124	7775
6	63	728	4095	15624	46655

As it was discussed in the previous literature, the number of solutions to the sparse systems of polynomials is described by the Bernstein's theorem which characterizes the number of solutions by the mixed volume of a corresponding Newton polyhedron of the system. The result in Rojas (1996) is obtained from the fact that if the coefficients are symmetrically distributed, then the computation of the mixed volume becomes symmetric as well. As a result, through a non-singular transformation, we can relate the distribution of the mixed volume for any distribution of polynomial coefficients to the standard normal distribution.

In McLennan and Berg (2002) and McLennan (2002) the authors consider a game with the payoffs of the players distributed normally. They provide an independent asymptotic analysis of this expected number of equilibria with large number of strategies. In particular, the authors show that the number of equilibria in the 2-person games grow exponentially (as $O\left(\exp(\mathcal{M}k)\right)$ for some constant $\mathcal{M}$) if the number of pure strategies of both players, k, grows at the same rate. However, if the number of strategies of one player is fixed, then the growth rate is slower than exponential (as $O\left(\{\log k\}^{\frac{M-1}{2}}\right)$, where M is the fixed number of strategies). Their analysis relies on the insight that degenerate equilibria arise only with probability zero. The expected number of equilibria in the asymptotic case can be expressed by the formula:

$$E\{N_{eq}\} = \frac{1}{\sqrt{M}}\left(\frac{\sqrt{\pi}\ \log\ k}{2}\right)^{M-1} + o\left([\log\ k]^{(M-1)/2}\right), \tag{36}$$

when the number of pure strategies of one player is fixed and equal to M and the number of strategies of the other player is equal to $k \to \infty$. The relative error of this asymptotic approximation becomes smaller than 2% only if the number of strategies N is greater than 100. The number of moments available from the game and the number of expected equilibria are tabulated in Tables 11 and 12. The number of moments significantly exceeds the expected number of equilibria.

Table 11: Expected number of equilibria for different numbers of pure strategies

	Number of strategies (k)											
M	10	20	50	200	300	400	500	600	700	800	900	1000
2	0.95	1.08	1.23	1.44	1.49	1.53	1.56	1.58	1.60	1.62	1.63	1.64
3	1.04	1.35	1.77	2.40	2.58	2.71	2.81	2.90	2.97	3.03	3.08	3.13
4	1.21	1.80	2.69	4.24	4.74	5.10	5.39	5.63	5.83	6.01	6.17	6.3
5	1.46	2.47	4.22	7.74	8.97	9.90	10.6	11.2	11.8	12.3	12.7	13.1
6	1.79	3.46	6.75	14.4	17.3	19.6	21.4	23.0	24.5	25.7	26.9	27.9
7	2.23	4.92	10.9	27.2	33.9	39.3	43.9	47.9	51.4	54.6	57.6	60.3
8	2.81	7.06	17.9	51.9	67.2	79.9	90.8	100	109	117	124	131
9	3.56	10.2	29.7	99.9	134	163	189	212	233	253	271	288
10	4.54	14.8	49.3	193	269	336	396	451	502	550	595	638
11	5.83	21.7	82.5	376	543	695	835	965	1087	1202	1312	1417

Table 12: The number of moments from the two-player game

	N											
M	10	20	50	200	300	400	500	600	700	800	900	1000
2	19	39	99	399	599	799	999	1199	1399	1599	1799	1999
3	29	59	149	599	899	1199	1499	1799	2099	2399	2699	2999
4	39	79	199	799	1199	1599	1999	2399	2799	3199	3599	3999
5	49	99	249	999	1499	1999	2499	2999	3499	3999	4499	4999
6	59	119	299	1199	1799	2399	2999	3599	4199	4799	5399	5999
7	69	139	349	1399	2099	2799	3499	4199	4899	5599	6299	6999
8	79	159	399	1599	2399	3199	3999	4799	5599	6399	7199	7999
9	89	179	449	1799	2699	3599	4499	5399	6299	7199	8099	8999
10	99	199	499	1999	2999	3999	4999	5999	6999	7999	8999	9999
11	109	219	549	2199	3299	4399	5499	6599	7699	8799	9899	10999