Estimating Static Models of Strategic Interactions

Author(s): Patrick Bajari, Han Hong, John Krainer and Denis Nekipelov

Source: *Journal of Business & Economic Statistics*, October 2010, Vol. 28, No. 4 (October 2010), pp. 469–482

Published by: Taylor & Francis, Ltd. on behalf of American Statistical Association

Stable URL: https://www.jstor.org/stable/20750856

Estimating Static Models of Strategic Interactions

Patrick BAJARI
University of Minnesota and NBER, Minneapolis, MN 55455

Han HONG
Stanford University, Stanford, CA 94305

John KRAINER
Federal Reserve Bank of San Francisco, San Francisco, CA 94105

Denis NEKIPELOV
University of California at Berkeley, Berkeley, CA 94720

We study the estimation of static games of incomplete information with multiple equilibria. A static game is a generalization of a discrete choice model, such as a multinomial logit or probit, which allows the actions of a group of agents to be interdependent. While the estimator we study is quite flexible, in most cases it can be easily implemented using standard statistical packages such as STATA. We also propose an algorithm for simulating the model which finds all equilibria to the game. As an application of our estimator, we study recommendations for high technology stocks between 1998–2003. We find that strategic motives, typically ignored in the empirical literature, appear to be an important consideration in the recommendations submitted by equity analysts.

KEY WORDS: Discrete choice; Stock analyst recommendation; Structural estimation.

1. INTRODUCTION

Game theory is one of the most commonly applied tools in economic theory, with substantive applications in all major fields in economics. In some fields, particularly industrial organization, game theory has not only transformed the analysis of market interactions, but also serves as an important basis for policy recommendations. Given the importance of gaming in economic theory, it is not surprising that the empirical analysis of games has been the focus of a recent literature in econometrics and industrial organization.

In much of the literature, a discrete game is modeled much like a standard discrete choice problem, such as the multinomial logit. An agent's utility is often assumed to be a linear function of covariates and a random preference shock. However, unlike a discrete choice model, utility is also allowed to depend on the actions of other agents. A discrete game strictly generalizes a standard random utility model, but does not impose the often strong assumption that agents act in isolation. Early attempts at the econometric analysis of such games included Bjorn and Vuong (1984) and Bresnahan and Reiss (1991a, 1991b). Recent contributions include Haile, Hortacsu, and Kosenok (2008), Aradillas-Lopez (2008), Ho (2009), Ishii (2008), Pakes et al. (2007), Augereau, Greenstein, and Rysman (2006), Seim (2006), Sweeting (2008), and Tamer (2003). In particular, Aguirregabiria and Mira (2002) propose a two-step method to estimate static games of incomplete information and illustrate it using an example of a static game of market entry.

An important insight in the recent literature is that it is often most straightforward to estimate discrete games in two steps. The static model of strategic interaction with incomplete information is a particular case when the discount rate is zero of the dynamic games considered in Aguirregabiria and Mira (2007), Bajari, Benkard, and Levin (2007), Berry, Pakes, and Ostrovsky (2007), and Pesendorfer and Schmidt-Dengler (2003). In a first step, the economist estimates the reduced forms implied by the model. This often boils down to using standard econometric methods to estimate the probability that one of the finite number of possible choices is observed, conditional on the relevant covariates. In the second step, the economist estimates a single agent random utility model, including as controls the equilibrium beliefs about the behavior of others from the first step.

In this paper, we study semiparametric estimation of static games of strategic interaction with multiple equilibria. Like the two-step approach discussed above, we estimate the reduced form choice probabilities in a first stage, and use them to simplify the estimation of the finite dimensional mean payoff parameters in the second stage.

The two-step approach pioneered in Aguirregabiria and Mira (2007) can be implemented both nonparametrically and semiparametrically. It is closely related to nonparametric identification of the mean utility functions, does not depend on whether the first-stage regressors are discrete or continuous, and does not require a correctly specified first-stage parametric model. The two-step estimator has desirable computational and statistical properties. First, when the regressors are continuous, despite the fact the first stage is nonparametric and might converge at a slow rate, the structural parameters estimated in the second stage have normal asymptotics and converge at a rate

proportional to the square root of the sample size. This follows from arguments based on Newey (1994). Under suitable regularity conditions, the asymptotic variance of the second stage estimator is invariant to whether the first stage nonparametric estimator is implemented using kernel methods or sieve methods. Second, in many cases the two-step nonparametric and semiparametric estimators can be implemented with correct standard errors using a two-stage least squares procedure in a standard statistical package like STATA. The simplicity of this approach makes the estimation of these models accessible to a larger audience of researchers.

In the context of discrete regressors, Pesendorfer and Schmidt-Dengler (2003) demonstrate that exclusion restrictions are sufficient for identification in a particular set of dynamic entry games. A related exclusion restriction, which excludes payoff-relevant covariates for a particular player from the utilities of the other players, is also required when the regressors are continuous. For instance, in an entry model, if the productivity shock of firm i influences its own entry decision, but only indirectly influences the entry decisions of other firms, then the mean payoff function is nonparametrically identified. The condition for nonparametric identification can be formulated as standard rank conditions for an appropriately defined linear system regardless of whether the regressors are continuous or discrete. This identification strategy relies crucially on the assumption that data in each market is generated from a single equilibrium. An alternative identification strategy that is not considered in this paper is to search for events that change which equilibrium to the game is played, but otherwise do not influence payoffs. Sweeting (2008) demonstrates that multiplicity of equilibrium can assist with identification in a symmetric location game.

The assumption that the data come from a single equilibrium has very different implications for discrete and continuous explanatory variables. If the vector of observable explanatory variables has a discrete support and the nonparametric estimator in the first step does not impose any smoothness condition (e.g., an unrestricted frequency estimator), then the assumption needed is that for a given value of explanatory variables, the data come from the same equilibrium. However, when the explanatory variables contain continuous variables, the first-step estimator usually imposes smoothness conditions for the second-step estimator to converge at a parametric rate to a normal distribution. This requires smoothness conditions with respect to continuous state variables in the equilibrium selection mechanism. These smoothness conditions can be stated in terms of pseudo-norms and may not allow for nondifferentiability in a set of measure zero; see, for example, Chen, Linton, and Van Keilegom (2003). In the presence of multiple equilibria, the points of nondifferentiability typically occurs when the equilibrium path bifurcates. The smoothness condition does require the equilibrium paths to not bifurcate for almost all values of the continuous state variable, or a smooth path is chosen at the points of bifurcation. If a substantial amount of discontinuity is present in selecting among multiple equilibria of the game, in which case an alternative approach is to incorporate an equilibrium selection mechanism using exclusion restrictions, either a full solution method or a version of the recursive method proposed by Aguirregabiria and Mira (2007) applied to a static game must be used.

As an application of our methods, we model the determination of stock recommendations (e.g., strong buy, buy, hold, sell) issued by equity analysts for high technology stocks listed in the NASDAQ 100 between 1998 and 2003. The determination of recommendations during this time period is of particular interest in the wake of the sharp stock price declines for technology firms in 2000. Recommended stocks underperformed the market as a whole during this period by a wide margin. Highly publicized allegations of conflicts of interest have called into question whether analysts were more concerned with helping their firms win investment banking business than with producing accurate assessments of the prospects of the firms under scrutiny. While there is a fairly large literature in finance on recommendations, we are not aware of any papers that formally consider the simultaneity of recommendations due to strategic motives.

In our model, recommendations submitted by analysts depend on four factors. First, recommendations must depend on fundamentals and commonly shared expectations about the future profitability of the firm. These expectations will be embedded in the stock price. Second, analysts are heterogeneous, both in terms of talent and perhaps in terms of access to information. We try to capture an individual analyst's private belief about the stock by looking at the difference between the quarterly earnings forecast submitted by the analyst (or the analyst's brokerage firm) and the distribution of forecasts from other firms. Mindful of the large number of inquiries into possible conflicts of interest among research analysts, we include as a third factor a dummy variable for an investment banking relationship between the firm and the analyst's employer.

Finally, we consider the influence of peers on the recommendation decision. Peer effects can impact the recommendation in different ways. Individual analysts have incentive to condition their recommendation on the recommendations of their peers, because even if their recommendations turn out to be unprofitable ex-post, performance evaluation is typically a comparison against the performance of peers. More subtly, recommendations are relative rankings of firms and are not easily quantifiable (or verifiable) objects. As such, ratings scales usually reflect conventions and norms. The phenomenon is similar to the college professor's problem of assigning grades. If a professor were to award the average student with a C while other faculty gave a B+ to the average student, the professor might incorrectly signal his views of student performance. Even while there is heterogeneity in how individual professors feel about grading, most conform to norms if only to communicate clearly with students (and their potential employers) about their performance. Similarly, analysts might have an incentive to benchmark their recommendations against perceived industry norms.

The paper is organized as follows. In Section 2 we outline the general economic environment. For purposes of exposition, we develop many of the key formulae within the context of a simple entry model. In Section 3 we discuss the problem of nonparametric identification of the mean payoff functions. In Section 4 we show how to derive nonparametric and semiparametric estimates of the structural parameters for our class of models. Section 5 contains the empirical application to equity analyst recommendations. Section 6 concludes the paper.

2. THE MODEL

In the model, there are a finite number of players, $i = 1, \ldots, n$ and each player simultaneously chooses an action $a_i \in \{0, 1, \ldots, K\}$ out of a finite set. We restrict players to have the same set of actions for notational simplicity. However, all of our results will generalize to the case where all players have different finite sets of actions. Let $A = \{0, 1, \ldots, K\}^n$ denote the vector of possible actions for all players and let $a = (a_1, \ldots, a_n)$ denote a generic element of A. As is common in the literature, we let $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ denote a vector of strategies for all players, excluding player i. We will abstract from mixed strategies since in our model, with probability one each player will have a unique best response.

Let $s_i \in S_i$ denote the state variable for player i. Let $S = \prod_i S_i$ and let $s = (s_1, \ldots, s_n) \in S$ denote a vector of state variables for all n players. We will assume that s is common knowledge to all players in the game and in our econometric analysis, we will assume that s is observable to the econometrician. The state variable is assumed to be a real valued vector, but S_i is not required to be a finite set. Much of the previous literature assumes that the state variables in a discrete game lie in a discrete set. While this assumption simplifies the econometric analysis of the estimator and identification, it is a strong assumption that may not be satisfied in many applications.

For each agent, there are also $K + 1$ state variables which we label as $\epsilon_i(a_i)$ which are private information to each agent. These state variables are distributed iid across agents and actions. Let ϵ_i denote the $1 \times (K + 1)$ vector of the individual $\epsilon_i(a_i)$. The density of $\epsilon_i(a_i)$ will be denoted as $f(\epsilon_i(a_i))$. However, we shall sometimes simplify the notation and denote the density for $\varepsilon_i = (\varepsilon_i(0), \ldots, \varepsilon_i(K))$ as $f(\epsilon_i)$.

The period utility function for player i is

$$u_i(a, s, \epsilon_i; \theta) = \pi_i(a_i, a_{-i}, s; \theta) + \epsilon_i(a_i). \tag{1}$$

The utility function in our model is similar to a standard random utility model such as a multinomial logit. Each player i receives a stochastic preference shock, $\epsilon_i(a_i)$, for each possible action a_i. In many applications, this will be drawn from an extreme value distribution as in the logit model. In the literature, the preference shock is alternatively interpreted as an unobserved state variable (see Rust 1994). Utility also depends on the vector of state variables s and actions a through $\Pi_i(a_i, a_{-i}, s; \theta)$. For example, in the literature, this part of utility is frequently parameterized as a simple linear function of actions and states. Unlike a standard discrete choice model, however, note that the actions a_{-i} of other players in the game enter into i's utility. A standard discrete choice model typically assumes that agents i act in isolation in the sense that a_{-i} is omitted from the utility function. In many applications, this is an implausible assumption.

In this model, player i's decision rule is a function $a_i = \delta_i(s, \epsilon_i)$. Note that i's decision does not depend on the ϵ_{-i} since these shocks are private information to the other $-i$ players in the game and, hence, are unobservable to i. Define $\sigma_i(a_i|s)$ as

$$\sigma_i(a_i = k|s) = \int 1\{\delta_i(s, \epsilon_i) = k\} f(\epsilon_i) \, d\epsilon_i.$$

In the above expression, $1\{\delta_i(s, \epsilon_i) = k\}$ is the indicator function that player i's action is k given the vector of state variables (s, ϵ_i). Therefore, $\sigma_i(a_i = k|s)$ is the probability that i

chooses action k conditional on the state variables s that are public information. We will define the distribution of a given s as $\sigma(a|s) = \prod_{i=1}^{n} \sigma(a_i|s)$.

Next, define $U_i(a_i, s, \epsilon_i; \theta)$ as

$$U_i(a_i, s, \epsilon_i; \theta) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, s; \theta)\sigma_{-i}(a_{-i}|s) + \epsilon_i(a_i),$$

$$\text{where } \sigma_{-i}(a_{-i}|s) = \prod_{j \neq i} \sigma_j(a_j|s). \tag{2}$$

In (2), $U_i(a_i, s, \epsilon_i; \theta)$ is player i's expected utility from choosing a_i when the vector of parameters is θ. Since i does not know the private information shocks, ϵ_j for the other players, i's beliefs about their actions are given by $\sigma_{-i}(a_{-i}|s)$. The term $\sum_{a_{-i}} \Pi_i(a_i, a_{-i}, s, \theta)\sigma_{-i}(a_{-i}|s)$ is the expected value of $\Pi_i(a_i, a_{-i}, s; \theta)$, marginalizing out the strategies of the other players using $\sigma_{-i}(a_{-i}|s)$. The structure of payoffs in (2) is quite similar to standard random utility models, except that the probability distribution over other agents' actions enter into the formula for agent i's utility. Note that if the error term has an atomless distribution, then player i's optimal action is unique with probability one. This is an extremely convenient property and eliminates the need to consider mixed strategies as in a standard normal form game.

We also define the deterministic part of the expected payoff as

$$\Pi_i(a_i, s; \theta) = \sum_{a_{-i}} \pi_i(a_i, a_{-i}, s, \theta)\sigma_{-i}(a_{-i}|s). \tag{3}$$

It follows immediately then that the optimal action for player i satisfies

$$\sigma_i(a_i|s) = \text{Prob}\{\epsilon_i|\Pi_i(a_i, s; \theta) + \epsilon_i(a_i)$$
$$> \Pi_i(a_j, s; \theta) + \epsilon_i(a_j) \text{ for } j \neq i\}. \tag{4}$$

2.1 A Simple Example

For expositional clarity, consider a simple example of a discrete game. Perhaps the most commonly studied example of a discrete game in the literature is a static entry game (see Bresnahan and Reiss 1991a, 1991b; Berry 1992; Tamer 2003; Ciliberto and Tamer 2009, and Manuszak and Cohen 2004). In the empirical analysis of entry games, the economist typically has data on a cross section of markets and observes whether a particular firm i chooses to enter a particular market. In Berry (1992) and Ciliberto and Tamer (2009), for example, the firms are major United States airlines such as American, United, and Northwest and the markets are large, metropolitan airports. The state variables, s_i, might include the population in the metropolitan area surrounding the airport and measures of an airline's operating costs. Let $a_i = 1$ denote the decision to enter a particular market and $a_i = 0$ denote the decision not to enter the market. In many applications, $\pi_i(a_i, a_{-i}, s; \theta)$ is assumed to be a linear function, for example,

$$\pi_i(a_i, a_{-i}, s; \theta) = \begin{cases} s' \cdot \beta + \delta \sum_{j \neq i} 1\{a_j = 1\} & \text{if } a_i = 1 \\ 0 & \text{if } a_i = 0. \end{cases} \tag{5}$$

In Equation (5), the mean utility from not entering is set equal to zero. We formally discuss this normalization in our section on identification. The term δ measures the influence of j's choice on i's entry decision. If profits decrease from having another firm enter the market then $\delta < 0$. The parameters β measure the impact of the state variables on $\pi_i(a_i, a_{-i}, s)$.

The random error terms $\varepsilon_i(a_i)$ are thought to capture shocks to the profitability of entry that are private information to firm i. Suppose that the error terms are distributed extreme value. Then, utility maximization by firm i implies that for $i = 1, \ldots, n$

$$
\begin{aligned}
\sigma_i(a_i = 1|s) &= \frac{\exp\left(s' \cdot \beta + \delta\sum_{j \neq i}\sigma_j(a_j = 1|s)\right)}{1 + \exp\left(s' \cdot \beta + \delta\sum_{j \neq i}\sigma_j(a_j = 1|s)\right)} \\
&= \Gamma_i(\beta, \delta, \sigma_j(1|s), \forall j).
\end{aligned}
\tag{6}
$$

In the system of equations above, applying the formula in Equation (3) implies that $\Pi_i(a_i, s; \theta) = s' \cdot \beta + \delta\sum_{j \neq i}\sigma_j(a_j = 1|s)$. Since the error terms are distributed extreme value, Equation (4) implies that the choice probabilities $\sigma_i(a_i = 1|s)$ take a form similar to a single agent multinomial logit model. We note in passing that it can easily be shown using Brouwer's fixed point theorem that an equilibrium to this model exists for any finite s (see McKelvey and Palfrey 1995).

We exploit the convenient representation of equilibrium in Equation (6) in our econometric analysis. Suppose that the econometrician observes $t = 1, \ldots, T$ repetitions of the game. Let $a_{i,t}$ denote the entry decision of firm i in repetition t and let the value of the state variables be equal to s_t. By observing entry behavior in a large number of markets, the econometrician could form a consistent estimate $\hat{\sigma}_i(a_i = 1|s)$ of $\sigma_i(a_i = 1|s)$ for $i = 1, \ldots, n$. In an application, this simply boils down to flexibly estimating the probability that a binary response, a_i, is equal to one, conditional on a given set of covariates. This could be done using any one of a number of standard techniques. Given first-stage estimates of $\hat{\sigma}_i(a_i = 1|s)$, we could then estimate the structural parameters of the payoff, β and δ, by maximizing a pseudo-likelihood function using $\Gamma_i(\beta, \delta, \hat{\sigma}_j(1|s), \forall j)$. There are two attractive features to this strategy. The first is that it is not demanding computationally. First-stage estimates of choice probabilities can be done using a strategy as simple as a linear probability model. The computational burden of the second stage is also light since we only need to estimate a logit model. A second attractive feature is that it allows us to view a game as a generalization of a standard discrete choice model. Thus, techniques from the voluminous econometric literature on discrete choice models can be imported into the study of strategic interaction. While the example considered above is simple, it nonetheless illustrates many of the key ideas that will be essential in what follows.

We can also see a key problem with identification in the simple example above. Both the first-stage estimates $\hat{\sigma}_i(a_i = 1|s)$ and the term $s' \cdot \beta$ depend on the vector of state variables s. This suggests that we will suffer from a collinearity problem when trying to separately identify the effects of β and δ on the observed choices. The standard solution to this type of problem in many settings is to impose an exclusion restriction. Suppose, for instance, a firm specific productivity shock is included in s. In most oligopoly models, absent technology spillovers, the productivity shocks of firms $-i$ will not directly enter into

firm i's profits. These shocks only enter indirectly through the endogenously determined actions of firms $-i$, for example, price, quantity, or entry decisions. Therefore, if we exclude the productivity shocks of other firms from the term $s' \cdot \beta$, we will no longer suffer from a collinearity problem. While this idea is quite simple, as we shall discover in the next section, similar restrictions are required to identify more general models.

3. IDENTIFICATION

In this section, we consider the problem of identifying the deterministic part of payoffs, without making particular assumptions about its functional form (e.g., that it is a linear index as in the previous example). In the context of this section, we let θ be completely nonparametric and write $\pi_i(a_i, a_{-i}, s)$ instead of $\pi_i(a_i, a_{-i}, s; \theta)$.

Definition 1. We will say that $\pi_i(a_i, a_{-i}, s)$ is identified if $\pi_i(a_i, a_{-i}, s) \neq \pi_i'(a_i, a_{-i}, s)$ for some $i = 1, \ldots, n$ implies that for the corresponding equilibrium choice probabilities $\sigma_i(a_i = 1|s) \neq \sigma_i'(a_i = 1|s)$ for some $i = 1, \ldots, n$.

Formally, identification requires that different values of the primitives generate different choice probabilities. If this condition is not satisfied, then it will be impossible for us to uniquely recover the structural parameters $\pi_i(a_i, a_{-i}, s)$ (for $i = 1, \ldots, n$) from knowledge of the observed choice probabilities, $\sigma_i(a_i = 1|s)$. While the mean payoff function is nonparametric, the model is semiparametric because the distribution of the unobservables is parametrically specified. Even in a single agent problem, it is well known that it is not possible to nonparametrically identify both the mean utility functions and the joint distribution of the error terms $F(\epsilon_i)$ without making strong exclusion and identification at infinity assumptions (see, e.g., Matzkin 1992).

To take the simplest possible example, consider a simple binary response model and assume that the error terms are normally distributed, as in the probit model. Let $\sigma_i(a_i = 1|s)$ denote the probability that the response is equal to one in the data conditional on s. Define $\Pi_i(a_i = 0|s) = 0$ and $\Pi_i(a_i = 1|s) = F^{-1}(\sigma_i(a_i = 1|s))$, where F^{-1} denotes the normal cdf. It can easily be verified that this definition of Π_i perfectly rationalizes any set of choice probabilities $\sigma_i(a_i = 1|s)$. Since even a single agent discrete choice model is not identified without a parametric assumption on the error term, assumptions at least as strong will be required in the more general setup with strategic interactions. In what follows, we will typically impose the assumption that the error terms are distributed iid with a known distribution function, since both independence and parametric form assumption on the error terms are required for identification.

Based on the discussion above, we shall impose the following assumption in order to identify the model:

A1. The error terms $\epsilon_i(a_i)$ are distributed iid across actions a_i and agents i. Furthermore, the parametric form of the distribution F comes from a known family.

Analogous to the notation in the previous section, define $\Pi_i(k, s) = \sum_{a_{-i}} \pi_i(a_i = k, a_{-i}, s)\sigma_{-i}(a_{-i}|s)$. It is straightforward to show that the equilibrium in this model must satisfy

$$
\delta_i(s, \epsilon_i) = k \quad \text{if and only if}
$$
$$
\Pi_i(k, s) + \epsilon_i(k) > \Pi_i(k', s) + \epsilon_i(k') \quad \text{for all } k' \neq k.
\tag{7}
$$

That is, action k is chosen if and only if the deterministic expected payoff and error term associated with action k is greater than the analogous values of $k' \neq k$. An implication of (7) is that the equilibrium choice probabilities $\sigma_i(a|s)$ must satisfy

$$
\begin{aligned}
\sigma_i(a_i|s) = \Pr\{ & \epsilon_i(a_i) + \Pi_i(a_i, s) - \Pi_i(0, s) \\
& > \epsilon_i(k) + \Pi_i(k, s) - \Pi_i(0, s), \\
& \forall k = 0, \ldots, K, k \neq a_i \}.
\end{aligned} \tag{8}
$$

Equation (8) is a simple consequence of (7) where we can subtract $\Pi_i(0, s)$ from both sides of the inequality.

Suppose we generate $\epsilon_i(a_i)$ from an extreme value distribution as in the multinomial logit model. Then (8) implies that $\sigma_i(a_i|s) = \exp(\Pi_i(a_i, s) - \Pi_i(0, s))/(\sum_{k=0}^{K} \exp(\Pi_i(k, s) - \Pi_i(0, s)))$. Alternatively, in an ordered logit model, for the logistic function $\Lambda(\cdot)$, $\sigma_i(a_i = k|s) = \Lambda(\Pi_i(k+1, s)) - \Lambda(\Pi_i(k, s))$. A key insight similar to Hotz and Miller (1993) is that Equation (8) implies that the equilibrium choice probabilities, $\sigma_i(a_i|s)$, have a one-to-one relationship to the "choice specific value functions," $\Pi_i(a_i, s) - \Pi_i(0, s)$. It is obvious that we should expect the one-to-one mapping in any model where the distribution of ϵ_i has full support. We let $\Gamma : \{0, \ldots, K\} \times S \to [0, 1]$ denote the map in general from choice specific value functions to choice probabilities, that is,

$$
\begin{aligned}
(\sigma_i(0|s), & \ldots, \sigma_i(K|s)) \\
& = \Gamma_i(\Pi_i(1, s) - \Pi_i(0, s), \ldots, \Pi_i(K, s) - \Pi_i(0, s)). \quad (9)
\end{aligned}
$$

We will denote the inverse mapping by Γ^{-1}:

$$
\begin{aligned}
(\Pi_i(1, s) - \Pi_i(0, s), & \ldots, \Pi_i(K, s) - \Pi_i(0, s)) \\
& = \Gamma_i^{-1}(\sigma_i(0|s), \ldots, \sigma_i(K|s)). \quad (10)
\end{aligned}
$$

The above analysis implies that we can invert the equilibrium choice probabilities to nonparametrically recover $\Pi_i(1, s) - \Pi_i(0, s), \ldots, \Pi_i(K, s) - \Pi_i(0, s)$. However, the above analysis implies that we will not be able to separately identify $\Pi_i(1, s)$ and $\Pi_i(0, s)$; we can only identify the difference between these two terms. Therefore, we shall impose the following assumption:

A2. For all i and all a_{-i} and s, $\pi_i(a_i = 0, a_{-i}, s) = 0$.

The above assumption is similar to the "outside good" assumption in a single-agent model where the mean utility from a particular choice is set equal to zero. In the context of the entry model, this assumption is satisfied if the profit from not entering the market is equal to zero regardless of the actions of other agents. Just as in the single-agent model, there are alternative normalizations that we could use to identify the $\pi_i(a_i, a_{-i}, s)$. However, for expositional simplicity we shall restrict attention to the normalization in A2.

Given assumption A2 and knowledge of the equilibrium choice probabilities, $\sigma_i(a_i|s)$, we can then apply the mapping in (10) to recover $\Pi_i(a_i, s)$ for all i, a_i, and s. Recall that the definition of $\Pi_i(a_i, s)$ implies that

$$
\Pi_i(a_i, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s) \pi_i(a_i, a_{-i}, s),
$$

$$
\forall i = 1, \ldots, n, a_i = 1, \ldots, K. \tag{11}
$$

Even if we know the values of $\Pi_i(a_i, s)$ and $\sigma_{-i}(a_{-i}|s)$ in the above equation, it is not possible to uniquely determine the values of $\pi_i(a_i, a_{-i}, s)$. To see why, hold the state vector s fixed, then determination of the utilities of all agents involves solving for $n \times K \times (K+1)^{n-1}$ unknowns. That is, there are n agents, for each action $k = 1, \ldots, K$, utility depends on the $(K+1)^{n-1}$ possible actions of the other agents. However, the left-hand side of (11) only contains information about $n \times (K+1)$ scalars holding s fixed. It is clearly not possible to invert this system in order to identify $\pi_i(a_i, a_{-i}, s)$ for all i, all $k = 1, \ldots, K$, and all $a_{-i} \in A_{-i}$.

Obviously, there must be cross-equation restrictions across either i or k in order to identify the system. One way to identify the system is to impose exclusion restrictions. Partition $s = (s_i, s_{-i})$, and suppose $\pi_i(a_i, a_{-i}, s) = \pi_i(a_i, a_{-i}, s_i)$ depends only on the subvector s_i. We can demonstrate this in the context of an entry model. In this type of model, the state is usually a vector of productivity shocks. While we might expect the profit of firm i to depend on the entry decisions of other agents, it should not depend on the productivity shocks of other agents. If such an exclusion restriction is possible, we can then write

$$
\Pi_i(a_i, s_{-i}, s_i) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s_{-i}, s_i) \pi_i(a_i, a_{-i}, s_i). \tag{12}
$$

Clearly, a necessary order condition for identification is that for each s_i, there exists $(K+1)^{n-1}$ points in the support of the conditional distribution of s_{-i} given s_i. Note that this assumption will be satisfied as long as s_{-i} contains a continuously distributed variable with $\Pi_i(a_i, a_{-i}, s_i)$ sufficiently variable. A sufficient rank condition will require that for almost all s_i, the system of equations obtained by varying the values of s_{-i} is nonsingular and invertible.

Theorem 1. Suppose that A1 and A2 hold. A necessary order condition for identifying the latent utilities $\Pi_i(a_i, a_{-i}, s_i)$ is that for almost all s_i, there exists $(K+1)^{n-1}$ points in the support of the conditional distribution of s_{-i} given s_i. A sufficient rank condition for identification is that for almost all values of s_i, the conditional second moment matrix of $E[\sigma_{-i}(a_{-i}|s_{-i}, s_i)\sigma_{-i}(a_{-i}|s_{-i}, s_i)'|s_i]$ is nonsingular.

Note that the rank condition holds regardless of whether the regressors are discrete or continuous. Because the rank condition is stated in terms of the observable reduced form choice probabilities, it is a testable assumption that can be verified from the data. It is analogous to the standard rank condition in a linear regression model. The difference is that the "regressors," $\sigma_{-i}(a_{-i}|s_{-i}, s_i)$ themselves have to be estimated from the data in a first stage. Intuitively, to identify strategic interaction models in which the primitive payoffs depend on the expected action of the opponent, the reduced form choice probabilities are required to depend on the opponent's idiosyncratic states. In the single-agent model with no strategic interactions, the left-hand side of (12) does not depend on s_{-i} and the right hand side does not depend on a_{-i}. The probabilities $\sigma_{-i}(a_{-i}|s_{-i}, s_i)$ sum up to one, and Equation (12) becomes an identity.

4. ESTIMATION

In the previous section, we demonstrated that there is a nonparametric inversion between choice probabilities and the choice-specific value functions, $\Pi(a_i, s)$. Furthermore, we demonstrated that the structural parameters of our model are identified if appropriate exclusion restrictions are made on payoffs. In this section, we exploit this inversion to construct nonparametric and semiparametric estimates of our structural parameters.

Step 1: Estimation of Choice Probabilities. Suppose the economist has access to data on $t = 1, \ldots, T$ repetitions of the game. For each repetition, the economist observes the actions and state variables for each agent $(a_{i,t}, s_{i,t})$. In the first step we form an estimate $\hat{\sigma}_i(k|s)$ of $\sigma_i(k|s)$ using sieve series expansions (see Newey 1990 and Ai and Chen 2003). We note, however, that we could alternatively estimate the first stage using other nonparametric regression methods such as kernel smoothing or local polynomial regressions.

The usual approach in the nested fixed point algorithm is to discretize the state space, which is only required to be precise enough subject to the constraints imposed by the computing facility. However, increasing the number of grids in a nonparametric or two-stage semiparametric method has two offsetting effects. It reduces the bias in the first-stage estimation but increases the variance. In fact, when the dimension of the continuous state variables is larger than four, it can be shown that it is not possible to obtain $\sqrt{T}$ consistency through discretization and asymptotically normal parameter estimates in the second stage, where T is the sample size. Therefore, discretizing the state space does not provide a solution to continuous state variables, which requires a more refined econometric analysis.

Let $\{q_l(s), l = 1, 2, \ldots\}$ denote a sequence of known basis functions that can approximate a real valued measurable function of s arbitrarily well for a sufficiently large value of l. The sieve could be formed using splines, Fourier Series, or orthogonal polynomials. We let the basis become increasingly flexible as the number of repetitions of the game T becomes large. Let $\kappa(T)$ denote the number of basis functions to be used when the sample size is T. We shall assume that $\kappa(T) \to \infty$, $\kappa(T)/T \to 0$ at an appropriate rate to be specified below. Denote the $1 \times \kappa(T)$ vector of basis functions as $q^{\kappa(T)}(s) = (q_1(s), \ldots, q_{\kappa(T)}(s))$, and its collection into a regressor data matrix as $Q_T = (q^{\kappa(T)}(s_1), \ldots, q^{\kappa(T)}(s_T))$.

One potential sieve estimator for $\hat{\sigma}_i(k|s), k = 1, \ldots, K$, is a linear probability model, that is,

$$\hat{\sigma}_i(k|s) = \sum_{t=1}^{T} 1(a_{it} = k) q^{\kappa(T)}(s_t)(Q_T' Q_T)^{-1} q^{\kappa(T)}(s). \quad (13)$$

Equation (13) is the standard formula for a linear probability model where the regressors are the sieve functions $q^{\kappa(T)}(s)$. We note that in the presence of continuous state variables, the sieve estimator $\hat{\sigma}_i(k|s)$ will converge to the true $\sigma_i(k|s)$ at a nonparametric rate slower than $\sqrt{T}$.

Second Step: Inversion. In our second step, we take as given our estimates $\hat{\sigma}_i(k|s)$ of the equilibrium choice probabilities. We then form an estimate of the expected deterministic utility functions, $\hat{\Pi}_i(k, s_t) - \hat{\Pi}_i(0, s_t)$ for $k = 1, \ldots, K$ and $t = 1, \ldots, T$.

This can be done by evaluating (10) using $\hat{\sigma}_i(k|s)$ in place of $\sigma_i(k|s)$. That is,

$$\left(\hat{\Pi}_i(1, s_t) - \hat{\Pi}_i(0, s_t), \ldots, \hat{\Pi}_i(K, s_t) - \hat{\Pi}_i(0, s_t)\right)$$
$$= \Gamma_i^{-1}(\hat{\sigma}_i(0|s_t), \ldots, \hat{\sigma}_i(K|s_t)).$$

In the specific case of the binary logit model, this inversion would simply be

$$\hat{\Pi}_i(1, s_t) - \hat{\Pi}_i(0, s_t) = \log(\hat{\sigma}_i(1|s_t)) - \log(\hat{\sigma}_i(0|s_t)).$$

In an alternative model, such as one with normal shocks, we would need to solve a nonlinear system. In what follows, we shall impose A2 so that $\hat{\Pi}_i(0, s) = \hat{\Pi}_i(0, a_{-i}, s) = 0$ for all a_{-i}.

Third Step: Recovering the Structural Parameters. In the first step we recovered an estimate of $\hat{\sigma}_i(a_i, s)$ and in the second step we recovered an estimate of the choice specific value function $\hat{\Pi}_i(k, s)$. In our third step, we use the empirical analog of (11) to form an estimate of $\pi(a_i, a_{-i}, s_i)$. We shall assume that we have made a sufficient number of exclusion restrictions, as discussed in the previous section, so that the model is identified. For a given value of s_i, for a given $a = (a_i, a_{-i})$, we estimate $\pi_i(a_i, a_{-i}, s_i)$ by minimizing the following weighted least square function:

$\Pi_i(a_i, a_{-i}, s_i)$, which are taken to be a vector of coefficients

$$\sum_{t=1}^{T}\left(\hat{\Pi}_i(a_i, s_{-it}, s_i) - \sum_{a_{-i}} \hat{\sigma}_{-i}(a_{-i}|s_{-it}, s_i)\pi_i(a_i, a_{-i}, s_i)\right)^2$$
$$\times w(t, s_i),$$

where the nonparametric weights $w(t, s_i)$ can take a variety of forms. For example,

$$w(t, s_i) = k\left(\frac{s_{it} - s_i}{h}\right) \Big/ \sum_{\tau=1}^{T} k\left(\frac{s_{i\tau} - s_i}{h}\right)$$

uses kernel weights, and other local weights are also possible. The identification condition in the previous section ensures that the regressor matrix in this weighted least squares regression is nonsingular asymptotically.

4.1 A Linear Model of Utility

The nonparametric estimation procedure described in the previous section follows the identification arguments closely and offers the advantage of flexibility and robustness against misspecification. However, without a huge amount of data, nonparametric estimation methods can be subject to a severe curse of dimensionality when we intend to control for a large dimension of state variables s. Also, in small samples, different implementations of nonparametric procedures may lead to drastically different point estimates. Therefore, in the following we consider a semiparametric estimation where the deterministic utility components $\pi_i(a_i, a_{-i}, s)$ are specified to be a linear function of a finite dimensional parameter vector θ. This is the typical econometric specification that is commonly used in the empirical literature. In this section we describe a straightforward estimation and inference procedure for this model.

The mean utility is assumed to take the form of $\pi_i(a_i, a_{-i}, s_i) = \Phi_i(a_i, a_{-i}, s_i)'\theta$. In the above expression, the deterministic part of utility is a linear combination of a vector of basis functions, $\Phi_i(a_i, a_{-i}, s_i)$. For instance, we might let utility be a linear index as in our simple entry game example of the previous section. Alternatively, we might choose $\Phi_i(a_i, a_{-i}, s_i)$ to be a standard flexible functional form, such as a high-order polynomial, spline function, or orthogonal polynomial. The estimator we discuss below can easily be generalized to allow for the possibility that θ enters the utility nonlinearly. However, the exposition of the properties of the estimator is facilitated by the linearity assumption. Also, most applications of discrete choice models and discrete games usually are linear in the structural parameters of interest.

This linearity assumption implies that the choice specific value function, given a_i and s, takes the convenient form $\Pi_i(a_i, s) = E[\pi_i(a_i, a_{-i}, s_i)|s, a_i] = \Phi_i(a_i, s)'\theta$, where $\Phi_i(a_i, s)$ is defined as

$$\Phi_i(a_i, s) = E[\Phi_i(a_i, a_{-i}, s_i)|a_i, s]$$
$$= \sum_{a_{-i}} \Phi_i(a_i, a_{-i}, s_i) \prod_{j \neq i} \sigma(a_j = k_j|s).$$

Equation (4) implies that each $\sigma_i(a_i|s)$ depends on $\sigma_j(a_j|s)$, $j \neq i$, through (14). We denote this mapping as

$$\sigma_i(a_i|s) = \Gamma_{i,a_i}(s, \sigma_j(k|s), j \neq i, k = 1, \ldots, K). \quad (14)$$

If we define $\sigma(s)$ to be the stacked vector of choice probabilities $\sigma_i(k|s)$ for all $k = 1, \ldots, K, i = 1, \ldots, n$, then we can collect (14) into a fixed point mapping: $\sigma(s) = \Gamma(\sigma(s))$. To emphasize the dependence on the parameter θ, we can also write

$$\sigma(s; \theta) = \Gamma(s, \theta; \sigma(s; \theta)). \quad (15)$$

4.2 Semiparametric Estimation

Step 1: Estimation of Choice Probabilities. The simple semiparametric procedure we propose proceeds in two steps. We begin by forming a nonparametric estimate of the choice probabilities, $\hat{\sigma}_i(k|s)$. We will do this like above using a sieve approach, though one could alternatively use kernels or a local polynomial method.

$$\hat{\sigma}_i(k|s) = q^{\kappa(T)}(s)'(Q_T'Q_T)^{-1} \sum_{\tau=1}^{T} q^{\kappa(T)}(s_\tau) 1(a_i = k). \quad (16)$$

Given our estimates of the choice probabilities, we can then estimate $\Phi_i(k, s)$ correspondingly by

$$\hat{\Phi}_i(k, s) = \sum_{a_{-i}} \Phi_i(a_i = k, a_{-i}, s_i) \prod_{j \neq i} \hat{\sigma}(a_j|s).$$

For instance, take the example presented in (5). In this example, $\Pi_i(a_i = 1, a_{-i}, s) = (s, \sum_{j \neq i} 1\{a_j = 1\}) \cdot (\beta, \delta)$, where "$\cdot$" denotes an inner product. Thus, in the above formula, $\Phi_i'(a_i = 1, a_{-i}, s) = (s, \sum_{j \neq i} 1\{a_j = 1\})$ and $\theta = (\beta, \delta)$. Then, given our first-stage estimates of the choice probabilities, $\hat{\Phi}'(a_i = 1, a_{-i}, s) = (s, \sum_{j \neq i} 1\{a_j = 1\}\sigma_j(a_j|s))$. For each parameter value θ, we can evaluate the empirical analogue of (15). For example, in the binary logit case, it is denoted as $\sigma_i(a_i = 1|s, \hat{\Phi}, \theta)$.

Step 2: Parameter Estimation. In the second stage, a variety of estimators can be used to recover the value of θ. Most of these estimators can be written as GMM estimators with a properly defined set of instruments. To describe the second stage, define $y_{ikt} = 1$ if $a_{it} = k$ and $y_{ikt} = 0$ otherwise, for $k = 0, \ldots, K$. Define $y_{it} = (y_{i1t}, \ldots, y_{iKt})$ and the vector

$$\sigma_i(s_t, \hat{\Phi}, \theta) = (\sigma_i(k|s_t, \hat{\Phi}, \theta), k = 1, \ldots, K).$$

Furthermore, collect $y_{it}, i = 1, \ldots, n$, into a long vector y_t with $n \times K$ elements, and similarly collect $\sigma_i(s_t, \hat{\Phi}, \theta), i = 1, \ldots, n$, into a long vector $\sigma(s_t, \hat{\Phi}, \theta)$ with corresponding $n \times K$ elements. Then for any dimension $\dim(\theta) \times (nK)$ matrix of instruments $\hat{A}(s_t)$, a GMM estimator $\hat{\theta}$ can be defined by solving the sample equations

$$\frac{1}{T} \sum_{t=1}^{T} \hat{A}(s_t)(y_t - \sigma(s_t, \hat{\Phi}, \hat{\theta})) = 0. \quad (17)$$

The instrument matrix $\hat{A}(s_t)$, known as $A(s_t)$, may be estimated in the first stage (such as two-step optimally weighted GMM), or may be estimated simultaneously (such as pseudo MLE). It is well known that the estimation errors in $\hat{A}(s_t)$ will not affect the asymptotic distribution of $\hat{\theta}$ defined by (17), regardless of whether $\hat{A}(s_t)$ is estimated in a preliminary step or is estimated simultaneously with $\hat{\theta}$. Therefore, next we will focus on deriving the large sample properties of $\hat{\theta}$ defined by (17) where $A(s_t)$ is known.

The estimator that we consider falls within the class of semiparametric estimators considered by Newey (1994). A somewhat surprising conclusion is that even though the first stage is estimated nonparametrically and can be expected to converge at a rate slower than $\sqrt{T}$, the structural parameters will be asymptotically normal and will converge at a rate of $\sqrt{T}$. Moreover, under appropriate regularity conditions, the second stage asymptotic variance will be independent of the particular choice of nonparametric method used to estimate the first stage (e.g., sieve or kernel). As a practical matter, these results justify the use of the bootstrap to calculate standard errors for our model.

In the Appendix, we derive the following result, applying the general framework developed by Newey (1990). Under appropriate regularity conditions, the asymptotic distribution of $\hat{\theta}$ defined in (17) satisfies

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, G^{-1}\Omega G^{-1'}),$$

where $G = EA(s_t)\frac{\partial}{\partial\theta}\sigma(s_t, \Phi_0, \theta_0)$, and Ω is the asymptotic variance of $A(s_t)(y_t - \sigma(s_t, \hat{\Phi}, \theta_0))$. In the Appendix, we also compare the asymptotic variance of alternative estimators.

4.3 Market Specific Payoff Models

If a large panel data with a large time dimension for each market is available, both the nonparametric and semiparametric estimators can be implemented market by market to allow for a substantial amount of unobserved heterogeneity. Even in the absence of such rich datasets, market specific payoff effects can still be introduced into the two-step estimation method if we are willing to impose a somewhat strong assumption on the market specific payoffs, α_t, which is observed by all the players in that market but not by the econometrician. We will assume

that α_t is an unknown but smooth function of the state variables $s_t = (s_{t1}, \ldots, s_{tn})$ in that market, which we will denote as $\alpha(s_t)$. In principal, we would prefer a model where the fixed effect was not required to be a function of the observables. However, in highly nonlinear models, such as ours, similar assumptions are commonly made. See, for example, Newey (1994). Strictly speaking, our assumption is stronger than and implies Newey (1994), who only assumes that sum of α_t and the idiosyncratic errors is homoscedastic and normal conditional on the observed state variables. This assumption, albeit strong, is convenient technically since it implies that the equilibrium choice probabilities, σ_i, can still be written as a function of the state s_t.

With the inclusion of a market-specific component, the mean period utility function in (1) for player i in market t is now modified to $\pi_i(a_i, a_{-i}, s) = \alpha(a_i, s) + \tilde{\pi}_i(a_i, a_{-i}, s)$. In the above, and what follows, we drop the market specific subscript t for notational simplicity.

Under the normalization assumption that $\Pi_i(0, a_{-i}, s) \equiv 0$ for all $i = 1, \ldots, n$, our previous results show that, as in (11), the choice-specific value functions $\Pi_i(a_i, s)$ are nonparametrically identified. Note that the choice specific value functions must satisfy $\forall i = 1, \ldots, n, a_i = 1, \ldots, K$:

$$\Pi_i(a_i, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\pi_i(a_i, a_{-i}, s)$$

$$= \alpha(a_i, s) + \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\tilde{\pi}_i(a_i, a_{-i}, s_i).$$

Obviously, since $\alpha(a_i, s)$ is unknown but is the same function across all market participants, it can be differenced out by looking at the difference of $\Pi_i(k, s)$ and $\Pi_j(k, s)$ between different players i and j. By differencing (18) between i and j one obtains

$$\Pi_i(k, s) - \Pi_j(k, s) = \sum_{a_{-i}} \sigma_{-i}(a_{-i}|s)\tilde{\pi}_i(a_i, a_{-i}, s_i)$$

$$- \sum_{a_{-j}} \sigma_{-j}(a_{-j}|s)\tilde{\pi}_j(a_j, a_{-j}, s_j).$$

Here we can treat $\tilde{\pi}_i(a_i, a_{-i}, s_i)$ and $\tilde{\pi}_j(a_j, a_{-j}, s_j)$ as coefficients, and $\sigma_{-i}(a_{-i}|s)$ and $\sigma_{-j}(a_{-j}|s)$ as regressors in a linear regression. Identification follows as in Theorem 1. As long as there is sufficient variation in the state variables s_i, s_j, the coefficients $\tilde{\pi}_i(a_i, a_{-i}, s_i)$ and $\tilde{\pi}_j(a_j, a_{-j}, s_j)$ can be nonparametrically identified.

We could nonparametrically estimate $\tilde{\pi}_j(a_j, a_{-j}, s_j)$ using an approach analogous to the nonparametric approach discussed in section Section 4. However, in practice, semiparametric estimation will typically be a more useful alternative. Denote the mean utility (less the market specific fixed effect) as: $\tilde{\pi}_i(a_i, a_{-i}, s_i) = \Phi_i(a_i, a_{-i}, s_i)'\theta$. In practice, we imagine estimating the structural model in two steps. In the first step, we estimate the equilibrium choice probabilities nonparametrically. In the second stage, we estimate $\tilde{\pi}_i$ treating $\alpha(s_t)$ as a fixed effect in a discrete choice model. Estimating discrete choice models with fixed effects is quite straightforward in many cases.

For instance, consider a model of entry and suppose that the error terms are distributed extreme value. In the first step, we nonparametrically estimate $\hat{\Phi}_i(1, s_{it})$, the probability of entry by firm i when the state is s_{it}. As in the previous section, we

could do this using a sieve linear probability model. In the second stage, we can form a conditional likelihood function as in Chamberlain (1984). This allows us to consistently estimate θ when market specific fixed effects $\alpha(s_t)$ are present. Alternatively, we can also apply a panel data rank estimation type procedure as in Manski (1987), which is free of distributional assumptions on the error term. It is worth emphasizing that the assumption of market specific payoff being a smooth function of observed state variables is a very strong one that is unlikely to hold in many important applications. In these cases a more general approach of coping with unobserved heterogeneity, as developed in Aguirregabiria and Mira (2007), is required.

5. APPLICATION TO STOCK MARKET ANALYSTS' RECOMMENDATIONS AND PEER EFFECTS

Next, we discuss an application of our estimators to the problem of analyzing the behavior of equity market analysts and the stock recommendations that they issue (e.g., strong buy, buy, hold, sell). There is a fairly sizeable empirical literature on this topic. However, the literature does not allow for strategic interactions between analysts. We believe that this is an important oversight. Accurate forecasts and recommendations are highly valued, of course. But the penalty for issuing a poor recommendation depends on whether competitor analysts also made the same poor recommendation. Therefore, the utility an analyst receives from issuing a recommendation is a function of the recommendations issued by other analysts. Therefore, we apply the framework discussed in the previous sections to allow payoffs to be interdependent.

The focus in this paper is on the recommendations generated for firms in the high-tech sector during the run-up and subsequent collapse of the NASDAQ in 2000; see also Barber et al. (2003), Chan, Karceski, and Lakonishok (2007), Womack (1996), Lin and McNichols (1998), and Michaely and Womack (1999). Given the great uncertainty surrounding the demand for new products and new business models, the late 1990s would seem to have been the perfect environment for equity analysts to add value. Yet analyst recommendations were not particularly helpful or profitable during this period. For example, the analysts were extremely slow to downgrade stocks, even as it was apparent that the market had substantially revised its expectations about the technology sector's earnings potential. The remarkably poor performance of the analysts during this time naturally led to questions that the recommendations were tainted by agency problems (see Barber et al. 2003). Allegedly, analysts faced a conflict of interest that would lead them to keep recommendations on stocks high in order to appease firms, which would then reward the analyst's company by granting it underwriting business or other investment advisory fees. Indeed, these suspicions came to a head when then New York State Attorney General Eliot Spitzer launched an investigation into conflicts of interest in the securities research business.

In this application we develop an empirical model of the recommendations generated by stock analysts from the framework outlined in Section 1. We quantify the relative importance of four factors influencing the production of recommendations in a sample of high technology stocks during the time period between 1998 and 2003.

5.1 Data

Our data consist of the set of recommendations on firms that made up the NASDAQ 100 index as of year-end 2001. The recommendations were collected from Thomson I/B/E/S. The I/B/E/S data is one of the most comprehensive historical data sources for analysts' recommendations and earnings forecasts, containing recommendations and forecasts from hundreds of analysts for a large segment of the set of publicly traded firms. It is common for analysts to rate firms on a 5-point scale, with 1 denoting the best recommendation and 5 denoting the worst. When this is not the case, these nonstandard recommendations are converted by Thomson to the 5-point scale.

We have 51,194 recommendations from analysts at 297 brokerage firms (see Table 1) submitted between March 1993 and June 2006 for firms in the NASDAQ 100. In a given quarter, for a given stock, we also merge a quarterly earnings forecast with a recommendation from the same brokerage firm. When there were multiple recommendations by the same analyst within a quarter, we chose to use the last recommendation in the results that we report. This merge will allow us to determine if analysts that are more optimistic than the consensus tend to give higher recommendations. In the I/B/E/S data, quarterly earnings forecasts are frequently made more than a year in advance. In order to have a consistent time frame, we limit analysis to forecasts that were made within the quarter that the forecast applies.

We chose to merge the brokerage field, instead of the analysts field, because the names and codes in the analysts field were not recorded consistently across I/B/E/S data sets for recommendations. It was possible to merge at the level of the brokerage. Note that not every recommendation can be paired with an earnings forecast made in the contemporaneous quarter. However, qualitatively similar results were found for a dataset where this censoring was not performed. We choose not to report these results in the interests of brevity. The variables in our data include numerical recommendations (REC) for stocks in the NASDAQ 100, the brokerage firm (BROKERAGE) employing the analyst, an accompanying earnings per share forecast (EPS) for each company with a recommendation, an indicator stating whether the brokerage firm has an investment banking relationship (RELATION) with the firm being recommended, and an indicator stating whether the brokerage firm has any investment banking relationship with a NASDAQ 100 company (IBANK).

The investment banking relationship was identified from several different sources. First, we checked form 424 filings in the U.S. Securities and Exchange Commission's (SEC) database for information on the lead underwriters and syndicate members of debt issues. When available, we used SEC form S-1 for information on financial advisors in mergers. We also gathered information on underwriters of seasoned equity issues from Securities Data Corporation's Platinum database. To be sure,

transaction advisory services (mergers) and debt and equity issuance are not the only services that investment banks provide. However, these sources contribute the most to total profitability of the investment banking side of a brokerage firm.

The average recommendation in our dataset is 2.2, which is approximately a buy recommendation (see Table 1). About 6% of the analyst–company pairs in the sample were identified as having a potential conflict of interest due to some kind of investment banking activity for the stock in question. A full 78% of the recommendations come from brokerage firms that had an investment banking relationship with at least one firm in the NASDAQ 100. Both of these variables are potentially useful measures of potential conflict of interest. The variable RELATION is more direct, since it indicates that the brokerage is engaged in investment banking with the same company it is making recommendations about, during the same quarter the recommendation was issued. However, brokerages might view any company it is giving a recommendation to as a potential client, particularly in the NASDAQ 100, where many of the companies generated considerable investment banking fees.

We also make use of analyst earnings forecasts. In a given quarter, for a given stock, we merge the quarterly earnings forecast with the recommendation from the same brokerage. This allows us to determine if analysts that are more optimistic than the consensus tend to give higher recommendations.

5.2 Empirical Model

An observation is a recommendation submitted for a particular stock during a specific quarter. We will let $t = 1, \ldots, T$ denote a quarter, $j = 1, \ldots, J$ a stock, and $i = 1, \ldots, I$ an analyst. We will denote a particular recommendation by $a_{i,j,t}$. The recommendation can take on integer values between 1 and 5, where 1 is the highest recommendation and 5 the lowest. Since the dependent variable can be naturally ranked from highest to lowest, we will assume that the utilities come from an ordered logit. Let $s(i, j, t)$ denote a set of covariates that influence the recommendation for analyst i for stock j during quarter t. Let $s(j, t)$ denote a vector of $(s(i, j, t))$ of payoff relevant covariates that enter into the utility of all the analysts who submit a recommendation for stock j during quarter q. Let $z(j, t)$ denote a set of covariates that shift the equilibrium, but which do not influence payoffs.

Define the utility or payoff to analyst i for a recommendation on stock j in quarter t to be,

$$\pi_{i,j,t} = \beta' s(i, j, t) + \eta E(a|s(j, t), z(j, t)) + \varepsilon_{i,j,t}. \quad (18)$$

In Equation (18), the term $E(a|s(j, t), z(j, t))$ is the expected recommendation for stock j during quarter t and $\varepsilon_{i,j,t}$ is an error term drawn from an extreme value model. Thus, conforming to the expected actions of peers enters into an individual analyst's utility. The model is the familiar ordered logit, where the probability that a particular recommendation is observed is determined as follows, where we let $\mu_0 = 0$

$$P(a = 1) = \Lambda\big(-\beta' s(i, j, t) - \eta E(a|s(j, t), z(j, t))\big),$$

$$P(a = k) = \Lambda\big(\mu_{k-1} - \beta' s(i, j, t) - \eta E(a|s(j, t), z(j, t))\big)$$
$$\qquad - \Lambda\big(\mu_{k-2} - \beta' s(i, j, t) \qquad\qquad (19)$$
$$\qquad - \eta E(a|s(j, t), z(j, t))\big), \quad k = 2, 3, 4,$$

$$P(a = 5) = 1 - \Lambda\big(\mu_3 - \beta' s(i, j, t) - \eta E(a|s(j, t), z(j, t))\big).$$

Table 1. Summary statistics

	Mean	Standard deviation	Min	Max
Recommendation	2.225	0.946	1	5
Relation	0.069	0.236	0	1
IBANK	0.778	0.416	0	1
Observations	51,194			

Table 2. Ordered logit estimates of the effect of fundamentals

Variable	Coefficient	Coefficient	Coefficient	Coefficient
%DEV	−0.0038	−0.0044		
	(−1.12)	(−1.24)		
ABS DEV			−0.0194	−0.0114
			(−6.33)	(−3.73)
Fixed effects	none	quarterly, stock	none	quarterly, stock
Log likelihood	−16,082.6	−15,211.7	−14,861.2	−14,861.4
Pseudo-R^2	0.0	0.054	0.001	0.054
Observations	51,194	51,194	51,194	51,194

In Equation (19), the likelihood that determines the probability that the recommendation is a depends on the latent estimated covariates β and η along with the cut points $\mu_1 - \mu_3$.

The analysis of the previous section suggests that identification depends crucially on having appropriate exclusion restrictions. First, we need covariates that influence the payoffs of one particular agent, but not other agents. In our analysis, the covariates will include IBANK and RELATION. This assumption would imply, for instance, that the amount of investment banking done by Merrill Lynch should not directly influence the recommendations submitted by analysts working for Goldman Sachs. We believe that this is a reasonable assumption.

In addition, we have attempted to control for unobserved heterogeneity in several ways. First, in many specifications, we include a full set of stock and quarter fixed effects to control for factors that remain fixed in a quarter that influence recommendations. Second, we have controlled for unobserved heterogeneity using both a fixed effects and random effects specification.

5.3 Results

The first question that we ask is the extent to which recommendations were determined by publicly observable information about the stocks. In our data, these fundamentals correspond to time fixed effects, stock fixed effects, and the difference between an individual analyst's beliefs about earnings and beliefs in the market as whole. In Table 2, we run an ordered logit to explore these questions. The variable %DEV is the percentage deviation of an analyst's earnings forecast from the average earnings forecast in the current quarter. DEV is the algebraic difference. In both cases, a more optimistic earnings forecast has the anticipated sign; a better earnings outlook is associated with a lower (i.e., better) recommendation. However,

the estimated coefficients are not significant at conventional levels in any of the specifications that we have tried. On the other hand, quarterly and stock fixed effects are almost all statistically significant (not reported in this table). If quarter and stock fixed effects proxy for publicly available information about the stock, then this information is considerably more important than measures of an individual analyst's optimism. In an earlier version of the paper we reported results on the correlation between our estimated quarterly effects and both the NASDAQ index and the QQQ. These results show that the quarterly effects can reasonably be interpreted as reflecting publicly observed information about the firms that is embedded in the share prices, as opposed to some other latent effects. These results are available from the authors on request.

In Table 3 we run an ordered logit model of recommendations as a function of our conflict of interest measures. The coefficient on RELATION indicates that potential conflicts of interest are statistically significant at conventional levels, except for the third column where quarterly and stock fixed effects are included, and the fourth column where the full set of fixed effects are included along with the more inclusive IBANK variable. The coefficient sign on RELATION is also consistent with our a priori beliefs that conflicts of interest could lead to the issuance of more favorable recommendations. However, these results must be interpreted with some caution. Since brokerage firms are expected to cover companies with whom they have significant investment banking business, the firms have an incentive to select brokerages that already view them favorably. It would be hard to imagine that a rational manager would want to hire an investment banking firm that views her company in an unfavorable manner.

Our results suggest that even though investment banking relationships may generate potential conflicts of interest for equity analysts, the magnitude of the effects on recommendations

Table 3. Ordered logit estimates of the effect of conflicts of interest

Variable	Coefficient	Coefficient	Coefficient	Coefficient
RELATION	−0.2787	−0.2215	−0.1695	−0.2083
	(−8.37)	(−6.51)	(−4.66)	(−5.68)
IBANK				0.1849
				(8.71)
Fixed effects	none	quarterly	quarterly, stock	quarterly, stock
Log likelihood	−65,716.7	−62,683.5	−61,465.4	−61,427.5
Pseudo-R^2	0.001	0.047	0.065	0.066
Observations	51,194	51,194	51,194	51,194

Table 4. Ordered logit estimates of peer effects (parametric first stage)

Variable	Coefficient	Coefficient	Coefficient	Coefficient
IVBELIEF	1.5937	1.5807	0.1437	0.1507
	(71.33)	(66.6)	2.98	3.2
RELATION				−0.2036
				(−4.84)
IBANK				0.1845
				(8.16)
Fixed effects	none	stock	quarterly, stock	quarterly, stock
Log likelihood	−62,371.0	−62,078.4	−61,315.8	−61,267.9
Pseudo-R^2	0.049	0.053	0.065	0.066
Observations	51,194	51,194	51,194	51,194

may be small in practice. Notice that measures of the goodness of fit are very low when only investment banking relationship is included. This overall finding is not consistent with the prosecutor's belief that "unbiased" research, separate from investment banking, will generate recommendations less tainted by potential conflicts of interest.

The final question we consider is whether peer effects come into play when analysts submit their recommendations. We explore this question in Tables 4–6 by using the two-stage procedure described in the previous sections. First, we regress the recommendations on a broker fixed effect, a full set of stock and quarterly dummies, and IBANK. In Table 4, these first-stage regressions are done using linear regression, while in the later tables we include stock-time interactions as a more flexible first stage. We experimented with other functional forms, such as a 3rd-order spline, and the results were little changed. We will let IVBELIEF for an analyst-broker i denote the expected average recommendation from the first-stage model, where the average excludes the predicted recommendation of that broker i. If the coefficient on IVBELIEF is positive, this means that broker i has an incentive to conform to the recommendations of the other brokers. If it is negative, it means there is a return from submitting a dissenting recommendation.

In all of the specifications that we examine in Table 4, peer effects seem to be important. An individual analyst will raise his recommendation proportionally to the recommendation that he expects from other analysts. This is intuitive. A recommendation does not make sense in isolation, but only in comparison to the recommendations of other analysts. If no one else in the

market is issuing recommendations of "market underperform" or "sell," an individual analysts may give the wrong signal by issuing such a recommendation even if he believes the recommendation is literally true. It is worth noting that the results for our measure of peer effects are not only statistically significant, but peer effects also explain the results quite well compared to the other covariates. The Pseudo-R^2 suggests that quarterly dummies, stock dummies, and IVBELIEF explain most of the variation in the data. Adding the additional conflict of interest variables does not do much to improving the model fit.

We note that the presence of the peer effect is robust to allowing for a more flexible first stage (see Tables 5 and 6). Also, the peer effect remains significant, allowing for unobserved heterogeneity in the form of a stock/quarter-specific random effect in Table 6. For these specifications, the investment banking relationship coefficient is no longer significant. In the random effect specification, the individual effect component is assumed to be drawn from a normal distribution with mean zero and a constant variance. The validity of the random effect model requires the strong assumption that the random effects are orthogonal to the regressors and the errors.

Note that it is possible to extend this analysis by using the parameter estimates obtained above and simulating the model in order to find all the equilibria in the recommendation game. Bajari et al. (2009) develop an algorithm to compute all solutions to this recommendation game (as well as to other applications). Perhaps not surprisingly, we find evidence of two equilibria in the pre-Spitzer era, one of which was characterized by an across-the-board tendency for analysts to grant more

Table 5. Ordered logit estimates of peer effects (semiparametric first stage)

Variable	Coefficient	Coefficient	Coefficient	Coefficient
IVBELIEF	1.5982	1.5872	0.1867	0.1938
	(81.82)	(62.73)	(4.58)	(3.95)
RELATION				−0.2033
				(−5.22)
IBANK				0.1851
				(7.16)
Fixed effects	none	stock	quarterly, stock	quarterly, stock
Log likelihood	−62,344.2	−62,050.6	−61,311.8	−61,263.7
Pseudo-R^2	0.049	0.054	0.065	0.066
Observations	51,194	51,194	51,194	51,194

Table 6. Random effect estimates of peer effects (semiparametric first stage)

Variable	Coefficient	Coefficient	Coefficient	Coefficient	
IVBELIEF	1.1717			1.1732	1.1961
	(39.43)			(40.35)	(43.96)
RELATION			−0.007	0.0681	−0.0281
			(−0.07)	(1.74)	(−0.57)
IBANK		0.3229			0.3832
		(8.67)			(18.67)
Log likelihood	−62,344.2	−62,050.6	−61,311.8	−61,263.7	
Observations	51,194	51,194	51,194	51,194	

optimistic ratings than in the single equilibrium that prevailed in the aftermath of the dot-com crash.

6. CONCLUSION

In this paper we propose a method for estimating static games of incomplete information. The method we propose is semiparametric and does not require the covariates to lie in a discrete set. Perhaps most importantly, the method is both flexible and easy to implement using standard statistical packages. We apply these methods to the problem of determining the factors that govern the assignment of stock recommendations by equity analysts for a set of high-tech stocks between 1998 and 2003. Two factors seem to be most important for explaining the production of stock recommendations. First, publicly observable information about the stocks under recommendation, as reflected in our time and quarter dummies, plays a large role in explaining the distribution of recommendations. Simply put, recommendations improved in 1999–2000 as the stock market rose, and then deteriorated as the market fell in the ensuing years. The second and most important factor for explaining recommendations is the peer group effect. Individual analysts appear to raise their recommendations proportionally to the recommendations they expect from their peers. Investment banking relationships are shown to be statistically significant in the recommendations regressions, but the economic effect of the investment banking relationship is estimated to be small. Additionally, when the investment banking relationship variables are included alongside our measure of peer effects, the banking relationships tend to be insignificant.

APPENDIX: SEMIPARAMETRIC VARIANCE

To derive Ω, we need to follow Newey (1990) and derive the asymptotic linear influence function of the left-hand side of the above relation. For this purpose, note that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(y_t - \sigma(s_t, \hat{\Phi}, \theta_0))$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(y_t - \sigma(s_t, \Phi_0, \theta_0))$$

$$- \frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(\sigma(s_t, \hat{\Phi}, \theta_0) - \sigma(s_t, \Phi_0, \theta_0)).$$

Since $\hat{\Phi}$ depends only on the nonparametric estimates of choice probabilities $\hat{\sigma}_j(k|s), j = 1, \ldots, n, k = 1, \ldots, K$, in (16) through (14), the second part can also be written as

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)\big(\Gamma(s_t, \theta_0; \hat{\sigma}(s)) - \Gamma(s_t, \theta_0; \sigma_0(s))\big),$$

where $\hat{\sigma}(s)$ is the collection of all $\hat{\sigma}_j(k|s)$ for $j = 1, \ldots, n$ and $k = 1, \ldots, K$, and the function $\Gamma(\cdot)$ is defined in (15). Then using the semiparametric influence function representation of Newey (1994), as long as $\Gamma(s_t, \theta, \sigma(s))$ is sufficiently smooth in $\sigma(s)$ and as long as the nonparametric first-stage estimates satisfy certain regularity conditions regarding the choice of the smoothing parameters, we can write this second part as

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)\big(\Gamma(s_t, \theta_0; \hat{\sigma}(s)) - \Gamma(s_t, \theta_0; \sigma_0(s))\big)$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)\frac{\partial}{\partial\sigma}\Gamma(s_t, \theta_0; \sigma_0(s))(y_t - \sigma(s_t, \theta_0))$$

$$+ o_p(1).$$

In other words, if we write $\Gamma_\sigma(s) = \frac{\partial}{\partial\sigma}\Gamma(s_t, \theta_0; \sigma_0(s))$, we can write

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(y_t - \sigma(s_t, \hat{\Phi}, \theta_0))$$

$$= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(I - \Gamma_\sigma(s_t))(y_t - \sigma(s_t, \theta_0)) + o_p(1).$$

Therefore, two-step semiparametric $\hat{\theta}$ has the following representation:

$$\sqrt{T}(\hat{\theta} - \theta_0) = -(EA(s_t)\Gamma_\theta(s_t))^{-1}$$

$$\times \frac{1}{\sqrt{T}}\sum_{t=1}^{T}A(s_t)(I - \Gamma_\sigma(s_t))(y_t - \sigma(s_t, \theta_0))$$

$$+ o_p(1).$$

Hence $\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \Sigma)$, where Σ is equal to

$$E(A(s_t)\Gamma_\theta(s_t))^{-1}\big[EA(s_t)(I - \Gamma_\sigma(s_t))\Omega(s_t)$$

$$\times (I - \Gamma_\sigma(s_t))'A(s_t)'\big]E(\Gamma_\theta(s_t)'A(s_t)')^{-1}.$$

Using the previous definitions of $\Gamma_\theta(s_t) = \frac{\partial}{\partial\theta_1}\Gamma(s_t, \theta_1, \sigma(s_t; \theta_2))|_{\theta_1=\theta_2=\theta_0}$, and $\Omega(s_t) = \text{Var}(y_t - \sigma(s_t, \theta_0)|s_t)$, the efficient choice of the instrument matrix (which can be feasibly estimated in preliminary steps without affecting the asymptotic variance) is given by

$$A(s_t) = \Gamma_\theta(s_t)'(I - \Gamma_\sigma(s_t))^{-1}\Omega(s_t)^{-1}(I - \Gamma_\sigma(s_t))^{-1'}.$$

With this efficient choice of the instrument matrix, the asymptotic variance of $\hat\theta$ becomes

$$\begin{aligned}\big(E\Gamma_\theta(s_t)'(I - \Gamma_\sigma(s_t))^{-1}\Omega(s_t)^{-1}\\ \times (I - \Gamma_\sigma(s_t))^{-1'}\Gamma_\theta(s_t)\big)^{-1}. \quad \text{(A.1)}\end{aligned}$$

A.1 Efficiency Considerations

We present two efficiency results in this section. First of all, we show that with the above efficient choice of the instrument matrix $A(s_t)$, the semiparametric two-step estimation procedure above is as efficient as the full maximum likelihood estimator where the fixed point mapping in (15) is solved for every parameter value θ which is then nested inside maximum likelihood optimization to obtain choice probabilities as a function of θ. Secondly, we show that estimating $\hat\sigma(s_t)$ may even improve efficiency over the hypothetical case where $\sigma(s_t)$ is known and an infeasible pseudo MLE which uses Φ_0 instead of $\hat\Phi$ is used to estimate θ.

A.1.1 Efficiency Comparison With Full Maximum Likelihood. Consider a full maximum likelihood approach where a fixed point calculation (assuming the solution is unique) of (15) is nested inside the likelihood optimization. For each θ, (15) is solved to obtain $\sigma(s_t, \theta)$ as a function of θ, which is then used to form the likelihood function. Define the total derivative of (15) as

$$\begin{aligned}\frac{d}{d\theta}\sigma(s_t, \theta_0) &= \frac{d}{d\theta}\Gamma(s_t, \theta, \sigma(s_t; \theta))\bigg|_{\theta=\theta_0}\\ &= \Gamma_\theta(s_t) + \Gamma_\sigma(s_t)\frac{d}{d\theta}\sigma(s_t, \theta_0)\end{aligned}$$

which can be used to solve for

$$\frac{d}{d\theta}\sigma(s_t, \theta_0) = (I - \Gamma_\sigma(s_t))^{-1}\Gamma_\theta(s_t). \quad \text{(A.2)}$$

Following the same logic as the discussions of pseudo MLE, it is easy to show that the asymptotic distribution of the full maximum likelihood estimator, which is the same as an IV estimator with the instruments chosen optimally, satisfies $\sqrt{T}(\hat\theta_{\text{FMLE}} - \theta_0) \xrightarrow{d} N(0, \Sigma_{\text{FMLE}})$, where

$$\Sigma_{\text{FMLE}} = \left(E\frac{d}{d\theta}\sigma(s_t, \theta_0)'\Omega(s)^{-1}\frac{d}{d\theta}\sigma(s_t, \theta_0)'\right)^{-1}.$$

Using (A.2), we can also write

$$\Sigma_{\text{FMLE}} = \left[E\Gamma_\theta(s_t)'(I - \Gamma_\sigma(s_t))^{-1}\Omega(s_t)^{-1} \right.$$
$$\left. \times (I - \Gamma_\sigma(s_t))^{-1}\Gamma_\theta(s_t)\right]^{-1}.$$

This is identical to (A.1) for the asymptotic variance of the two-step semiparametric IV estimator when the instrument matrix is chosen optimally.

A.1.2 Efficiency Comparison With Infeasible Pseudo MLE. Consider an infeasible pseudo MLE, with $\hat\Phi$ replaced by the true but unknown Φ_0:

$$\sum_{t=1}^{T}\sum_{i=1}^{n}\left[\sum_{k=1}^{K} y_{ikt}\log\sigma_i(k|s_t, \Phi_0, \theta) + \left(1 - \sum_{k=1}^{K} y_{ikt}\right)\right.$$
$$\left. \times \log\left(1 - \sum_{k=1}^{K}\sigma_i(k|s_t, \Phi_0, \theta)\right)\right]. \quad \text{(A.3)}$$

The asymptotic variance of this estimator is similar to that of Σ_{FMLE} except with $\frac{d}{d\theta}\sigma(s_t, \theta_0)'$ replaced by $\Gamma_\theta(s_t)$. In other words,

$$\Sigma_{\text{IPMLE}} = [E\Gamma_\theta(s_t)'\Omega(s_t)^{-1}\Gamma_\theta(s_t)]^{-1}$$

where IPMLE stands for infeasible pseudo MLE.

The relation between Σ_{FMLE} and Σ_{IPMLE} is obviously ambiguous and depends on the response matrix $\Gamma_\sigma(s_t)$. It is clearly possible that $\Sigma_{\text{FMLE}} < \Sigma_{\text{IPMLE}}$, in which case estimating $\hat\Phi$ may improve efficiency over the case where Φ_0 is known.

ACKNOWLEDGMENTS

REFERENCES

Aguirregabiria, V., and Mira, P. (2002), "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70 (4), 1519–1543. [469]

—— (2007), "Sequential Simulation-Based Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1–53. [469,470,476]

Ai, C., and Chen, X. (2003), "Efficient Estimation of Models With Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71 (6), 1795–1843. [474]

Aradillas-Lopez, A. (2008), "Semiparametric Estimation of a Simultaneous Game With Incomplete Information," working paper, Princeton University. [469]

Augereau, A., Greenstein, S., and Rysman, M. (2006), "Coordination vs. Differentiation in a Standards War: 56K Modems," *RAND Journal of Economics*, 37, 887–909. [469]

Bajari, P., Benkard, L., and Levin, J. (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75 (5), 1331–1370. [469]

Bajari, P. J., Hong, H., Krainer, J., and Nekipelov, D. (2009), "Estimating Games With Multiple Equilibria," working paper, Stanford University. [479]

Barber, B., Lehavey, R., McNichols, M., and Trueman, B. (2003), "Reassessing the Returns to Analysts' Stock Recommendations," *Financial Analysts Journal*, 59 (2), 88–96. [476]

Berry, S. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60 (4), 889–917. [471]

Berry, S., Pakes, A., and Ostrovsky, M. (2007), "Simple Estimators for the Parameters of Dynamic Games, With Entry/Exit Examples," *RAND Journal of Economics*, 38 (2), 373–399. [469]

Bjorn, P. A., and Vuong, Q. (1984), "Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation With an Application to Labor Force Participation," SSWP No. 537, Caltech. [469]

Bresnahan, T., and Reiss, P. (1991a), "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81. [469,471]

———— (1991b), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009. [469,471]

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics*, Vol. 2, eds. Z. Griliches and M. Intriligator, Amsterdam, Boston: Elsevier, North-Holland. [476]

Chan, L., Karceski, J., and Lakonishok, J. (2007), "Analysts' Conflict of Interest and Biases in Earnings Forecasts," *Journal of Financial and Quantitative Analysis*, 42 (4), 893–913. [476]

Chen, X., Linton, O., and Van Keilegom, I. (2003), "Estimation of Semiparametric Models When the Criterion Function Is Not Smooth," *Econometrica*, 71 (5), 1591–1608. [470]

Ciliberto, F., and Tamer, E. (2009), "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, to appear. [471]

Haile, P., Hortacsu, A., and Kosenok, G. (2008), "On the Empirical Content of Quantal Response Equilibrium," *American Economic Review*, 98 (1), 180–200. [469]

Ho, K. E. (2009), "Insurer–Provider Networks in the Medical Care Market," *American Economic Review*, 99 (1), 393–430. [469]

Hotz, J., and Miller, R. (1993), "Conditional Choice Probabilties and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529. [473]

Ishii, J. (2008), "Interconnection Pricing, Compatibility, and Investment in Network Industries: ATM Networks in the Banking Industry," working paper, Stanford University. [469]

Lin, H., and McNichols, M. (1998), "Underwriting Relationships, Analysts' Earnings Forecasts, and Investment Recommendations," *Journal of Accounting Economics*, 25, 1–34. [476]

Manski, C. F. (1987), "Semiparametric Analysis of Random Effect Linear Models From Binary Panel Data," *Econometrica*, 55 (2), 357–362. [476]

Manuszak, M., and Cohen, A. (2004), "Endogenous Market Structure With Discrete Product Differentiation and Multiple Equilibria: An Empirical Analysis of Competition Between Banks and Thrifts," working paper, Federal Reserve Board of Governors. [471]

Matzkin, R. (1992), "A Nonparametric and Distribution Free Estimator of the Binary Choice and the Threshold Crossing Models," *Econometrica*, 60 (2), 239–270. [472]

McKelvey, R., and Palfrey, T. (1995), "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10 (1), 6–38. [472]

Michaely, R., and Womack, K. (1999), "Conflict of Interest and Credibility of Underwriter Analyst Recommendations," *Review of Financial Studies*, 12, 653–686. [476]

Newey, W. (1990), "Semiparametric Efficiency Bounds," *Journal of Applied Econometrics*, 5 (2), 99–135. [474,475,480]

———— (1994), "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 62, 1349–1382. [470,475,476,480]

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2007), "Moment Inequalities and Their Application," working paper, Harvard University. [469]

Pesendorfer, M., and Schmidt-Dengler, P. (2003), "Identification and Estimation of Dynamic Games," Working Paper w9726, NBER. [469,470]

Rust, J. (1994), "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, Vol. 4, eds. R. Engle and D. McFadden, Amsterdam, Boston: Elsevier, North-Holland, pp. 3082–3146. [471]

Seim, K. (2006), "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *RAND Journal of Economics*, 37 (3), 619–640. [469]

Sweeting, A. (2008), "The Strategic Timing of Radio Commercials: An Empirical Analysis Using Multiple Equilibria," *RAND Journal of Economics*, to appear. [469,470]

Tamer, E. (2003), "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70 (1), 147–165. [469, 471]

Womack, K. (1996), "Do Brokerage Analyst Recommendations Have Investment Value," *Journal of Finance*, 51 (1), 137–167. [476]