Econometric Analysis of Games with Multiple Equilibria

Áureo de Paula

Department of Economics, University College London, London WC1E 6BT, United Kingdom; Center for Microdata Methods and Practice (CeMMAP) and Institute for Fiscal Studies, London WC1E 7AE, United Kingdom; email: a.paula@ucl.ac.uk

Annu. Rev. Econ. 2013. 5:107–31

First published online as a Review in Advance on August 29, 2013

The *Annual Review of Economics* is online at economics.annualreviews.org

This article's doi: 10.1146/annurev-economics-081612-185944

JEL codes: C1, C3, C7

Keywords

identification, multiplicity, social interactions

Abstract

This article reviews the recent literature on the econometric analysis of games in which multiple solutions are possible. Multiplicity does not necessarily preclude the estimation of a particular model (and, in certain cases, even improves its identification), but ignoring it can lead to misspecifications. The review starts with a general characterization of structural models that highlights how multiplicity affects the classical paradigm. Because the information structure is an important guide to identification and estimation strategies, I discuss games of complete and incomplete information separately. Although many of the techniques discussed here can be transported across different information environments, some are specific to particular models. Models of social interactions are also surveyed. I close with a brief discussion of post-estimation issues and research prospects.

1. INTRODUCTION

In this article I review the recent literature on the econometric analysis of games in which multiple solutions are possible. Equilibrium models are a defining ingredient of economics. Game theoretic models, in particular, have played a prominent role in various subfields of the discipline for many decades. When taking these models to data, one endows a sample of games represented by markets, neighborhoods, or economies with an interdependent payoff structure that depends on observable and unobservable variables (to the econometrician and potentially to the players), and participants choose actions. One pervasive feature in many of these models is the existence of multiple solutions for various payoff configurations, and this aspect carries over to estimable versions of such systems.

Although the existence of more than one solution for a given realization of the payoff structure does not preclude the estimation of a particular model (and, in certain cases, even improves their identification), ignoring its possible occurrence can potentially cause severe misspecifications and nonrobustness in the analysis of substantive questions. Fortunately, much has been learned in the recent past about the econometric properties of such models. The tools available benefit from advances in identification analysis, estimation techniques, and computational capabilities, and I discuss some of these below.

Owing to space limitations, the survey is by no means exhaustive. I nevertheless cover some of the main developments thus far. Because most of the literature has concentrated on parametric models, this is also my focus here. As in many other contexts, the parametric and functional restrictions that are imposed deserve careful deliberation, and some of the parametric and functional restrictions in the models I present can be relaxed (e.g., the linearity of the parametric payoff function and the distributional assumptions in theorem 2 of Tamer 2003 and the analysis of social interactions models in Brock & Durlauf 2007). Once point or partial identification has been established, estimation typically proceeds by applying well-understood methods such as maximum likelihood and method of moments (many times with the assistance of simulations) in the case of point-identified models or by carrying out recently developed methods for partially identified models. A thorough discussion of partially identified models would require much more space, and I leave that for other surveys covering those methods in more detail (see, e.g., Tamer 2010). I nevertheless do discuss estimation and computation aspects that are somewhat peculiar to the environments described below.

In the games analyzed here, given a set of payoffs for the economic agents involved, a solution concept defines the (possibly multiple) outcomes that are consistent with the economic environment. The solution concepts I use below essentially consist of mutual best responses (plus consistent beliefs when information is asymmetrically available), and I refer to those as equilibria or solutions indiscriminately (hopefully without much confusion to the reader). In the following sections, the solution concepts are Nash equilibrium for complete-information games, Bayes-Nash or Markov perfect equilibrium for incomplete-information games, and rational expectations equilibrium as defined in the social interactions literature for those types of models. Although these are commonly assumed solution concepts, others exist. Aradillas-Lopez & Tamer (2008), for example, consider rationalizable strategies, and network-formation games rely on pairwise stability or similar concepts. Multiplicity is often an issue for these alternative definitions, and many of the ideas discussed below (e.g., bounds) can be used when those concepts are adopted instead.

One important ingredient guiding identification and estimation strategies in these models is the information environment of a game. Whether a game is one of complete or incomplete (i.e., private) information may affect the econometric analysis in a substantive manner. Many techniques discussed below can be transported across these different information environments, but some are specific to particular models. In the next sections, I discuss identification and estimation in games of complete and incomplete information separately. I also separately survey models of social interactions, in which multiplicity occurs as well. For certain specifications, these models coincide with games of incomplete information, and the discussion of that class of games carries forward. Because the number of players in a given economy is typically large in this class of models, additional estimation strategies handling multiplicity may be employed. I end the article with a discussion of postestimation issues and research prospects.

2. PRELIMINARY FOUNDATIONS

Following early analyses of the empirical content of games with multiple equilibria, such as Jovanovic (1989), I begin by casting my discussion in terms of the perspective adopted by Koopmans and coworkers in the mid-twentieth century. I do so because it makes apparent how multiplicity interferes with usual econometric methodologies and lays many of the models on a common ground, making it easier to identify discrepancies and commonalities. Koopmans and colleagues recognized that the analysis of economic phenomena often requires that we go beyond the mere statistical description of the observable probability distributions of interest and pursue primitive parameters, or policy-invariant features, of an economic model:

> In many fields the objective of the investigator's inquisitiveness is not just a "population" in the sense of a distribution of observable variables, but a physical structure projected behind this distribution, by which the latter is thought to be generated. . . . The structure concept is based on the investigator's ideas as to the "explanation" or "formation" of the phenomena studied, briefly, on his theory of these phenomena, whether they are classified as physical in the literal sense, biological, psychological, sociological, economic or otherwise. (Koopmans & Reiersol 1950, p. 165)

The above considerations embody the spirit of what became known as the Cowles Commission research program. The research philosophy typically identified with the Cowles Commission prescribed (*a*) defining an economic model, (*b*) specifying its probabilistic features to take the economic model to data, and (*c*) using statistical methods to estimate policy-invariant parameters and test relevant hypotheses about those parameters. The first step is a prerequisite for the study of any counterfactual intervention in the economy. The economic model essentially postulates how the outcomes and other variables in the model relate to each other. Some of these variables are observed by the econometrician, and others are latent or unobserved by the researcher. When the distribution of observed variables is consistent with only one parameter configuration, the model is point identified. If that is not the case, the model is partially or set identified, and the set of parameters consistent with the observable data may still be informative about the analyst's research question.

One implicit assumption in the above analysis is that the economic model predicts only one value of the observed variables for a given realization of latent variables. This is not necessarily the case for many models of interest. A simple but intriguing early example is that of the entry game depicted by Bresnahan & Reiss (1991) (see also Bjorn & Vuong 1984). The population comprises independent and identically distributed (i.i.d.) copies of a game between two players. In Bresnahan & Reiss (1991), the players are firms (e.g., small businesses in different geographical markets) deciding on whether to enter a particular market. In Bjorn & Vuong (1984), the players are husbands and wives, and the action is whether to participate in the labor force. This model is also related to the precursor works on dummy endogenous variables, although it should be noted that the multiagent, fully simultaneous nature of the problem here creates important differences.

The researcher observes different markets, and in each of these markets there are two firms. It is typically assumed that firms can be labeled. In many applications, this label is natural and refers to the actual identity of the firm (e.g., Delta and United Airlines) or, in the case of a household, the roles of husband and wife. Here I assume that firms are labeled by $i = 1, 2$. A firm's decision depends on its profit, which in turn depends on whether the other firm also entered the market. Let $y_i \in \{0, 1\}$ denote whether firm i enters ($y_i = 1$) or not ($y_i = 0$) and assume that profits are given by $\mathbf{x}_i^\top \beta_i + \Delta_i y_j + u_i, i = 1, 2, j \neq i$, where $\mathbf{x}_i$ are observable variables that affect firm i's profit and u_i is an unobserved (to the econometrician) variable affecting the profit for firm i. The distribution of $\mathbf{u} \equiv (u_1, u_2)$ is denoted by F. In Ciliberto & Tamer's (2009) analysis of the airline industry, for example, the market is a particular route, the firms are different airlines (e.g., American, Delta, and United Airlines), and $\mathbf{x}_i$ comprises market- and firm-specific variables affecting demand (e.g., population size, income) and costs for the firms (e.g., various distance measures between the route end points and the nearest hub airport of the firm). The linear functional form $\mathbf{x}_i^\top \beta_i$ can be relaxed, but additive separability between observed and latent variables in the payoff structure is typically necessary for inference. Because I assume that player identities or roles can be assigned, the parameters β_i and Δ_i are allowed to be label or role specific. If players' roles cannot be distinguished, symmetry in the coefficients would have to be imposed. (This is the case in the social interactions models reviewed in Section 5.) Here I suppose that $\mathbf{x}_i$ includes at least one variable that impacts only firm i. This is a common and powerful exclusion restriction used in the literature, although it may not always be necessary in the analysis of games.

I use this simple payoff structure with varying information environments throughout to highlight some of the more important aspects in the econometric analysis of games with possibly multiple equilibria. For now, I assume there is complete information for the players: Realizations of $\mathbf{x}$ and $\mathbf{u}$ are observed by all agents in the game. The payoff matrix for this game is given by **Table 1**.

Much of the literature has focused on pure Nash equilibria as the preferred solution concept (see Berry & Tamer 2006 for a discussion of equilibria in mixed strategies, however). If observable outcomes y correspond to Nash equilibria of the above game, the following econometric model is obtained (see Aradillas-Lopez & Tamer 2008 for examples using other solution concepts):

$$\begin{aligned} y_i^* &= \mathbf{x}_i^\top \beta_i + y_j \Delta_i + u_i, \\ y_i &= \mathbf{1}_{y_i^* \geq 0}. \end{aligned} \tag{1}$$

If $\Delta_i < 0$, this can be analyzed using **Figure 1**.

In the central region of the plane, the model predicts two possible solutions. What is missing here is an equilibrium selection mechanism, and in that sense, the model is incomplete. For payoff realizations within the region of indeterminacy (i.e., the central region), one could imagine circumstances in which (0, 1) is always selected or scenarios in which (1, 0) is always selected. Also

Table 1 Two-by-two game

		Player 2	
		0	1
Player 1	0	$(0, 0)$	$(0, \mathbf{x}_2^\top \beta_2 + u_2)$
	1	$(\mathbf{x}_1^\top \beta_1 + u_1, 0)$	$(\mathbf{x}_1^\top \beta_1 + \Delta_1 + u_1, \mathbf{x}_2^\top \beta_2 + \Delta_2 + u_2)$

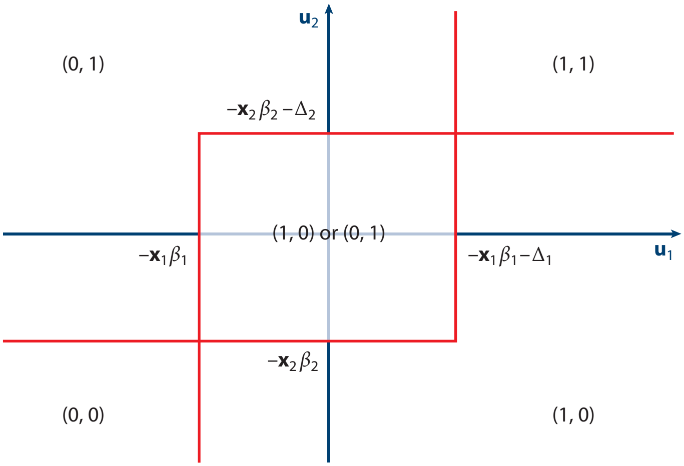

Figure 1

Complete-information two-by-two game ($\Delta_i < 0$, $i = 1, 2$).

possible are intermediary cases in which, with some probability [potentially dependent on the realizations of $\mathbf{x}$ and $\mathbf{u}$ and the parameters in the model $\theta \equiv (\beta_1, \beta_2, \Delta_1, \Delta_2)$], say $\lambda(\mathbf{x}, \mathbf{u}, \theta) \in [0, 1]$, one of the two equilibria is selected whenever payoffs fall in the multiplicity region. Different selection probabilities [i.e., $\lambda(\mathbf{x}, \mathbf{u}, \theta)$ in the example] will induce different distributions over the observable outcomes y_i.

One could (and in many examples below does) include the equilibrium selection mechanism into the structure. Nevertheless, one must bear in mind that modeling the equilibrium selection process requires extra assumptions. This opens up an additional avenue for misspecification. Moreover, an estimated equilibrium selection mechanism is more likely to be policy sensitive. This is because

> in a game with multiple equilibria, anything that tends to focus the players' attention on one equilibrium may make them all expect it and hence fulfill it, like a self-fulfilling prophecy. … The question of which equilibrium would be focused on and played by real individuals in a given situation can be answered only with reference to the psychology of human perception and the cultural background of the players. (Myerson 1991, pp. 108, 113)

Consequently, a question that permeates much of the literature on empirical games with possibly many equilibria is whether one can be agnostic about the equilibrium selection rule, economizing on identifying assumptions, and still be able recover the parameters of interest (i.e., θ in this example) or functions of these parameters. Of course, many authors adopt a less pessimistic view than Myerson's and point out that certain equilibria (such as Pareto-efficient, risk-dominant, and pure-strategy equilibria) are more salient and more likely to be played than others. In the husband and wife example, for instance, important papers in the intrahousehold allocation literature support the view that Pareto-inefficient equilibria should not be selected. Inference methods that incorporate the equilibrium selection mechanism may be informative about when and how certain equilibria are likely to be played.

Jovanovic (1989) stresses that multiplicity may or may not preclude point identification. (Without further restrictions, the above example is not point identified.) Certainly, even when the model is not point identified, it may still be partially (i.e., set) identifiable and informative about many issues of interest. What is perhaps more intriguing is that, in some instances, multiplicity can help establish point identification or make set-identified models more informative partly because it introduces additional variation in the data (see, e.g., Manski 1993, section 3.1).

3. GAMES OF COMPLETE INFORMATION

In Section 2 I present a two-player game in which information is assumed to be complete: Both players know the realized values of all observable and latent variables (i.e., there is no private information). Bresnahan & Reiss (1991, proposition 1) show that multiplicity (of pure-strategy Nash equilibria) will happen with positive probability in similar models of simultaneous moves with more than two actions and two players if the support of the latent variables $\mathbf{u}$ is large enough [e.g., $u_i \in (-\infty, \infty)$]. For simplicity, I stay with the example above with two actions (0 or 1) and two players. In what follows I discuss various approaches to inference in games of complete information.

3.1. Pooling Multiple Equilibria Outcomes

If $\Delta_i < 0, i = 1, 2$, the entry of a firm in the market affects the other firm's profits negatively, and all equilibria in pure strategies for a given realization of $\mathbf{u}$ involve a unique number $y_1 + y_2$ of players choosing 1. Let $\mathbf{1}(N, \mathbf{x}, \mathbf{u}, \theta)$ be 1 if the number of players choosing 1 is N for $\mathbf{x}$, $\mathbf{u}$, and θ, and 0 otherwise. In this case, the likelihood conditional on $\mathbf{x}$ of observing N players choosing 1 is given by

$$\mathbb{P}(N|\mathbf{x};\theta) = \int \mathbf{1}(N, \mathbf{x}, \mathbf{u}; \theta)\mathbb{P}(d\mathbf{u}|\mathbf{x}).$$

For the illustrative example in the previous section, we obtain

$$\begin{aligned}
\mathbb{P}_\theta(N=0|\mathbf{x}) &= F_{u_1,u_2}(-\mathbf{x}_1^\top\beta_1,\ -\mathbf{x}_2^\top\beta_2),\\
\mathbb{P}_\theta(N=2|\mathbf{x}) &= 1 - F_{u_1}(-\mathbf{x}_1^\top\beta_1 - \Delta_1) - F_{u_1}(-\mathbf{x}_2^\top\beta_2 - \Delta_2)\\
&\quad + F_{u_1,u_2}(-\mathbf{x}_1^\top\beta_1 - \Delta_1,\ -\mathbf{x}_2^\top\beta_2 - \Delta_2),\\
\mathbb{P}_\theta(N=1|\mathbf{x}) &= 1 - \mathbb{P}_\theta(N=0|\mathbf{x}) - \mathbb{P}_\theta(N=2|\mathbf{x}),
\end{aligned}$$

where $\mathbf{u}$ is assumed to be independent of $\mathbf{x}$, and F_{u_1,u_2} is its known cumulative distribution function (CDF). Given the assumption of independence between $\mathbf{x}$ and $\mathbf{u}$ and the known CDF F_{u_1,u_2}, point identification can be ascertained using arguments similar to those employed for parametric discrete choice models. Once point identification of θ is demonstrated, the model can be estimated via maximum likelihood, under the assumption that $\Delta_i < 0$, with a random sample of games. This strategy is pursued, for example, by Bresnahan & Reiss (1990) to identify and estimate a model of firm entry in automobile retail markets and by Berry (1992) for an entry model in the airline industry. It also highlights that multiplicity is not necessarily an impediment to econometric analysis. This happens because certain quantities (i.e., the number of entrants) are invariant across equilibria when more than one solution is possible.

More generally, the key insight is that in this model, certain outcomes can only occur as unique equilibria. When $\Delta_i < 0$, $i = 1, 2$, this happens for $\mathbf{y} = (0, 0)$ and $\mathbf{y} = (1, 1)$. This provides an avenue for identification and estimation of the model. In a coordination game, where $\Delta_i > 0$, $i = 1, 2$, multiple (pure-strategy) Nash equilibria occur whenever

$\mathbf{u} = (u_1, u_2) \in \left[-\mathbf{x}_1^\top \beta_1 - \Delta_1, -\mathbf{x}_1^\top \beta_1\right] \times \left[-\mathbf{x}_2^\top \beta_2 - \Delta_2, -\mathbf{x}_2^\top \beta_2\right]$. In this case, both $\mathbf{y} = (0, 0)$ and $\mathbf{y} = (1, 1)$ are possible equilibria. [$\mathbf{y} = (0, 0)$ is a unique equilibrium when $u_i < -\mathbf{x}_i^\top \beta_i - \Delta_i$, $i = 1$, 2, and $\mathbf{y} = (1, 1)$ is a unique equilibrium when $u_i > -\mathbf{x}_i^\top \beta_i$, $i = 1, 2$.] The number of players choosing 1 is no longer the same across equilibria, but one could nonetheless mimic the previous strategy and consider the probability of events $\{(0, 1)\}$, $\{(1, 0)\}$, and $\{(1, 1), (0, 0)\}$, where one pools together any two outcomes that are both equilibria for some given $\mathbf{x}$ and $\mathbf{u}$. Once this is done, singleton events correspond to outcomes that occur only as unique equilibria regardless of the values taken by $\mathbf{x}$ and $\mathbf{u}$.

Honoré & de Paula (2010) pursue the idea of identifying certain (nonexhaustive) outcomes with the occurrence of multiple equilibria for the identification of a complete-information timing game. The outcomes $\mathbf{y}$ are duration variables chosen by individuals to model circumstances in which timing decisions by one individual affect the payoff of another individual (e.g., joint migration, joint retirement, technology adoption). The authors show that multiplicity occurs only when duration spells terminate simultaneously for both players (i.e., $y_1 = y_2$), and sequential spell terminations (i.e., $y_1 \neq y_2$) occur only as unique equilibria. Identification and estimation can then be achieved by considering those events separately using standard arguments in the duration literature.

Unfortunately, this strategy is not always feasible in more general games as most of the possible outcomes (if not all) might arise as equilibria alongside many other possible solutions for a set of $\mathbf{x}$ and $\mathbf{u}$ with high probability. This would limit the ability of this insight to achieve point identification (although it might still allow for set identification). Let us take, for instance, a simple example from Jovanovic (1989). This example corresponds to the two-action, two-player example from Section 2 with $\beta_1 = \beta_2 = 0$, $\Delta \equiv \Delta_1 = \Delta_2 \in [0, 1)$, and the PDF for $\mathbf{u}$ is $f_{\mathbf{u}}(u_1, u_2) = 1_{(u_1,u_2)\in[-1,0]^2}$ (i.e., the latent variables are independently and uniformly distributed on $[-1, 0]$). In this example, $\mathbf{y} = (0, 0)$ is a Nash equilibrium for any realization of $\mathbf{u}$. In addition, $\mathbf{y} = (1, 1)$ is also a Nash equilibrium for $(u_1, u_2) \in [-\Delta, 0]^2$. Moreover, $(0, 1)$ and $(1, 0)$ are never equilibria. Notice that $(0, 0)$ arises as a unique equilibrium when $(u_1, u_2) \notin [-\Delta, 0]^2$ but also occurs as an equilibrium alongside $(1, 1)$. Observation of $\mathbf{y} = (0, 0)$ does not allow one to determine whether (u_1, u_2) were drawn in the region of multiplicity (i.e., $[-\Delta, 0]^2$). Pooling the two outcomes as in the discussion above is hopeless for (either point or set) identification (and consequently estimation) as $\mathbb{P}_\theta(\{(0, 0), (1, 1)\}) = 1$. This is not an isolated example or a mere consequence of the lack of covariates: Bresnahan & Reiss (1991, proposition 3) show that this problem is generally present for a wide class of discrete action games. In this case, pooling outcomes that are equilibria for the same realizations of observed and latent covariates provides no identification leverage.

3.2. Large Support and Identification

Adapting similar ideas from the individual discrete choice literature (e.g., Manski 1988, Heckman 1990), Tamer (2003) provides an alternative strategy for point identification of the parameters of interest in the presence of multiplicity. Consider models in which $\Delta_1 \times \Delta_2 < 0$, for example. **Figure 2** depicts the possible equilibria on the space of latent covariates (u_1, u_2). In contrast to the case in which $\Delta_1 \times \Delta_2 > 0$, any equilibrium is unique. Nonetheless, there are no equilibria in pure strategies in the central region of the figure. Tamer assumes that any outcome is possible in that area. This can be rationalized as either that players are thought to choose one of the outcomes in ways that are not prescribed by the solution concept adopted here (Nash in pure strategies) or that the (unique) equilibrium in mixed strategy places positive probability on every outcome $\mathbf{y}$ (although the mixed-strategy equilibrium probabilities are not incorporated into the econometric model). Either way, we end up with an incomplete econometric model in which the outcome of the game in the central region is not resolved within the economic model. In this case, the strategy from

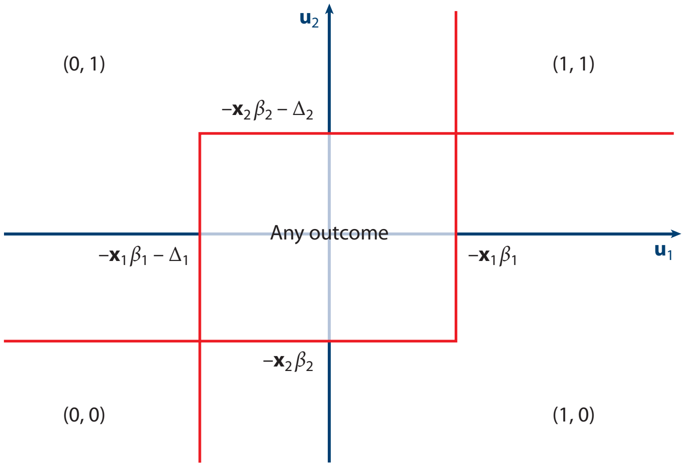

Figure 2

Complete-information two-by-two game ($\Delta_1 > 0$, $\Delta_2 < 0$).

the previous subsection would be inconclusive, as in Jovanovic's example, because it would bundle together all possible outcomes, given that they are all allowed in the central area of **Figure 2**.

Tamer (2003, theorem 2) nevertheless demonstrates the point identification of the parameter vector θ using exclusion restrictions and large support conditions on the observable covariates. Without loss of generality, assume that $\Delta_1 > 0$ and $\Delta_2 < 0$. In this case, given that games are i.i.d. and assuming that $\mathbf{u}$ is independent of other covariates and has a known CDF, the following result holds.

> **Theorem (Tamer 2003):** Assume that for $i = 1$ or $i = 2$, there exists a regressor x_{ik} with $\beta_k \neq 0$ such that $x_{ik} \notin \mathbf{x}_j$, $j \neq i$, and such that the distribution of $x_{ik}|\mathbf{x}_{-ik}$ has everywhere positive density, where $\mathbf{x}_{-ik} = (x_{i1}, \ldots, x_{ik-1}, x_{ik+1}, \ldots, x_{iK})$ and K is the dimension of β_i, $i = 1, 2$. Then the parameter vector $\theta = (\beta_1, \beta_2, \Delta_1, \Delta_2)$ is identified if the matrices $\mathbb{E}\left[\mathbf{x}_1\mathbf{x}_1^\top\right]$ and $\mathbb{E}\left[\mathbf{x}_2\mathbf{x}_2^\top\right]$ are nonsingular.

Here I sketch only the idea behind this result. Without loss of generality, assume that $\beta_{ik} > 0$ (an analogous argument can be made for $\beta_{ik} < 0$). Because of the large support condition, one can find extreme values of x_{ik} such that only unique equilibria in pure strategies are realized [because $\mathbb{P}\left(u_i > -\Delta - \mathbf{x}_i^\top\beta_i\right) \to 1$ and $\mathbb{P}\left(u_i > -\mathbf{x}_i^\top\beta_i\right) \to 1$ as $x_{ik} \to \infty$]. When $\beta_j \neq \beta_j'$, $j \neq i$, the full rank condition on $\mathbb{E}\left[\mathbf{x}_j\mathbf{x}_j^\top\right]$ guarantees that $\mathbf{x}_j^\top\beta_j \neq \mathbf{x}_j^\top\beta_j'$ with positive probability. This implies that in the limit (as $x_{ik} \to \infty$), $\mathbb{P}_{\beta_j}(y_j = 1|\mathbf{x}) \neq \mathbb{P}_{\beta_j'}(y_j = 1|\mathbf{x})$, and the model identifies the parameter β_j. Similar arguments can be employed to demonstrate the identifiability of the remaining parameters. This point-identification strategy can be extended to other contexts. Grieco (2012), for example, extends it to a game of incomplete information in which some variables are publicly observed by the participants but not by the econometrician. I also note that the result uses the excluded regressor x_{ik} in conjunction with the large support assumption. This is nonetheless not necessary in more general contexts, for example, when the payoff is not a linear function of $\mathbf{x}$ for every player (see, e.g., the generalization in Bajari et al. 2010b).

The essence of this result is to find values of the regressor for which the actions of all but one player are dictated by dominant strategies (regardless of the realizations of the latent unobservable variables). This turns the problem into one of a discrete choice by the single agent who does not play dominant strategies (given the covariates). The existence of a regressor value that induces some players to always choose one action may be hard to ascertain in many empirical contexts, but one may get reasonably close to that ideal.

Ciliberto & Tamer (2009), for example, estimate a firm entry model for the airline industry in which the distance between two airports affects differentially the decision of traditional carriers and discount airlines to operate a route between those two airports. They find that distance carries a very negative effect on the payoff for Southwest Airlines and a mildly positive effect on the payoff for traditional carriers. Ciliberto & Tamer (2009, p. 1818) point out that "this is consistent with anecdotal evidence that Southwest serves shorter markets than the larger national carriers." One could then think of markets in which the distance between airports is large as inducing a small probability of entry by Southwest and other discount airlines but still providing variation in entry decisions by other, larger, carriers.

When the regressor values guaranteeing a dominant strategy for a subset of players cannot be obtained in an empirical application, other identification ideas need to be employed. One feasible identification plan relies on bounds, focusing on a (possibly nonsingleton) set of parameters that are consistent with the model (see Section 3.3). In that case, if the probability that certain players adopt dominant strategies gets larger, one approaches the conditions for the point-identification strategy above, and the identified set of parameters becomes smaller.

Because the point-identification result above relies on extreme values of the covariates, even with the ideal conditions for Tamer's result satisfied, this identification-at-infinity strategy will have important consequences for inference, as pointed out by Khan & Tamer (2010), leading to asymptotic convergence rates that are slower than parametric rates as the sample size (i.e., number of games) increases (see also the discussion in Bajari et al. 2011).[1]

Bajari et al. (2010b) incorporate the equilibrium selection mechanism into the problem and demonstrate how large support conditions can help establish necessary conditions for point identification of the selection process. In expanding the model to include the equilibrium selection, they generalize the paper by Bjorn & Vuong (1984) in which equilibria are selected with a nondegenerate probability that is estimated along the other payoff-relevant parameters. Bajari et al. point out that "estimating the selection mechanism allows the researcher to simulate the model, which is central to performing counterfactuals." Some caution is nonetheless warranted in justifying the policy invariance of the equilibrium selection mechanism in relation to the particular counterfactual of interest, as indicated above (and also below).

To maintain consistency with the discussion above, I assume that there are always pure-strategy equilibria and remain in the two-player, two-action environment introduced above. Bajari et al. (2010b) do allow for equilibria in mixed strategies (see also Berry & Tamer 2006), which are guaranteed to exist under more general conditions. Let $\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta) \subset \{0, 1\}^2$ be the set of equilibrium profiles given particular realizations of $\mathbf{x}$ and $\mathbf{u}$ and the value of θ. Denote by $\lambda(\mathbf{y}|\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta), \mathbf{x}, \mathbf{u}, \theta)$ the probability that $\mathbf{y}$ is selected when the equilibrium set is $\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta)$ given covariates at $\mathbf{x}$ and $\mathbf{u}$ and the parameter θ. The distribution of latent variables F is known. When $\mathbf{y}$ is the unique equilibrium for the realizations of $\mathbf{x}$ and $\mathbf{u}$ and the parameter vector θ, $\lambda(\mathbf{y}|\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta), \mathbf{x}, \mathbf{u}, \theta) = 1$. When it is not an equilibrium, $\lambda(\mathbf{y}|\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta), \mathbf{x}, \mathbf{u}, \theta) = 0$. Finally, when there are other equilibria, $\lambda(\mathbf{y}|\mathcal{E}(\mathbf{x}, \mathbf{u}, \theta), \mathbf{x}, \mathbf{u}, \theta) \in [0, 1]$. The conditional probabilities of actions are then

[1]The statement of theorem 3 in Tamer (2003) is inaccurate. An additional term in the asymptotic variance is missing, potentially invalidating the efficiency claim in the result, as noted by Hahn & Tamer (2004).

$$\mathbb{P}\big(\mathbf{y}|\mathbf{x};(\theta,F_{\mathbf{u}})\big)=\int\lambda\left(\mathbf{y}|\mathcal{E}(\mathbf{x},\mathbf{u},\theta),\mathbf{x},\mathbf{u},\theta\right)\mathbf{1}_{\mathbf{y}\in\mathcal{E}(\mathbf{x},\mathbf{u},\theta)}f_{\mathbf{u}}(\mathbf{u})d\mathbf{u}, \tag{2}$$

where $f_{\mathbf{u}}(\cdot)$ is the PDF for $\mathbf{u}$, which is assumed to be independent of $\mathbf{x}$.

Given the specification summarized in Equation 2, it is unclear whether the model identifies the equilibrium selection mechanism $\lambda(\cdot)$. In fact, unless further restrictions are imposed, it does not. To take an extreme, but simple illustration, consider again the example in Jovanovic (1989) with no covariates. In this case, let λ denote the probability that (1, 1) is selected whenever there are multiple equilibria (i.e., $(u_1, u_2) \in [-\Delta, 0]^2$). Then

$$\mathbb{P}\big(\mathbf{y}=(1,1)\big)=\lambda\Delta^2 \quad \text{and} \quad \mathbb{P}\big(\mathbf{y}=(0,0)\big)=1-\mathbb{P}\big(\mathbf{y}=(1,1)\big).$$

It is not possible to pin down λ and Δ from the distribution of outcomes. One of the issues here is that there are more unknowns than there are equations. To reduce the degrees of freedom in the problem, Bajari et al. impose additional structure. For example, the number of equations can be increased if the support of covariates $\mathbf{x}$ is relatively large, generating additional conditional moments. To keep the number of parameters under control, Bajari et al. assume that selection probabilities depend only on utility indices through a relatively low-dimensional sufficient statistic. In essence this allows them to parameterize the equilibrium selection mechanism as $\lambda(\cdot)=\lambda(\cdot;\gamma)$ for some parameter γ of relatively low dimension compared to the cardinality of covariate support. Once this is guaranteed and exclusion restrictions such as those in Tamer's (2003) theorem above are imposed, one can rely on there being at least as many conditional moments as there are parameters (i.e., θ and γ), which is necessary for identification (see Bajari et al. 2010b, theorem 3).

Of course, one should be careful not to introduce misspecifications while incorporating the equilibrium selection mechanism into the econometric model (especially a parametric one). Because the number of equilibria and the selection probability of an equilibrium may depend on all variables observed by the players and the parameter θ, the selection mechanism should typically allow for this possibility. Any misspecification will likely contaminate the estimation of θ. Furthermore, Bajari et al.'s (2010b) results provide necessary but not sufficient conditions for the point identification of the selection mechanism. Finally, as highlighted above, one ought to be careful not to extrapolate the equilibrium selection mechanism in counterfactual experiments mimicking policies that might affect the way in which equilibria are selected.

If point identification can be established, Bajari et al. (2010b) suggest estimating the parameters of interest using the conditional probability in Equation 2 by the method of simulated moments to handle the integration over the latent variables $\mathbf{u}$. (As the authors point out, the simulated maximum likelihood method could also be applied here.) Here the estimation procedure requires the determination of $\mathcal{E}(\mathbf{x},\mathbf{u},\theta)$ for given realizations of $\mathbf{x}$ and $\mathbf{u}$ and values of θ. Computation of the whole equilibrium set becomes prohibitive already with a relatively small number of players and actions. Bajari et al. (2010b) also suggest strategies to partly accommodate the computational issues.

With a random sample of $g = 1, \ldots, G$ games, one can then estimate the parameters of interest using moments such as

$$\mathbb{E}\big[\big(1_{\mathbf{y}_t=a}-\mathbb{P}(\mathbf{y}=a|\,\mathbf{x};\theta,\gamma,F)\big)\,h(\mathbf{x})\big]=0,$$

where $h(\cdot)$ are appropriately chosen weight functions of the observable covariates and a are admissible action profiles. The sample analog of the above moment for the method of simulated moments estimator is given by

$$\sum_{g=1}^{G}\left(1_{\mathbf{y}_g=a}-\widehat{\mathbb{P}}\left(\mathbf{y}=a|\mathbf{x};\theta,\gamma,F\right)\right)h\left(\mathbf{x}\right),$$

where $\widehat{\mathbb{P}}$ is a computer-simulated estimate of P.

3.3. Bounds

Typically the covariates' support is not rich enough, and the previous point-identification strategy will not suffice. Another avenue for inference in games with possibly multiple equilibria is to rely on partial identification and use bounds for estimation. Take again the example in Jovanovic (1989) and note that $\mathbb{P}(\mathbf{y}=(1,1))$ is at most Δ^2 [if $(1,1)$ is always selected when $(u_1,u_2)\in[-\Delta,0]^2$]. Hence because $\Delta<1$, we have $\sqrt{\mathbb{P}(\mathbf{y}=(1,1))}\leq\Delta<1$. This is the set of all parameters Δ that are consistent with the observable distribution of outcomes: For every Δ within $\left[\sqrt{\mathbb{P}(\mathbf{y}=(1,1))},1\right)$, there is an equilibrium selection probability that delivers the same distribution of observables. Because $\mathbb{P}(\mathbf{y}=(1,1))$ can be consistently estimated, one can estimate the identified set by $\left[\sqrt{\widehat{\mathbb{P}}(\mathbf{y}=(1,1))},1\right)$.

This insight appears in Jovanovic's discussion of this particular example and is explored in more detail by Tamer (2003). Consider the entry game in Section 2 for $\Delta_i<0, i,=1,2$. We handled this case above using the fact that, whenever multiple equilibria occur, the number of entrants is the same. In more general models (e.g., when there are more players and payoffs are heterogeneous), this is not always the case, even if we restrict ourselves to pure-strategy equilibria, as noted by Ciliberto & Tamer (2009). We use this simple example to illustrate the construction of bounds for the parameters of interest.

Remember that in this case (0, 0) and (1, 1) always occur as unique (pure) strategy equilibria. Note also that

$$\mathbb{P}\left(\mathbf{y}=(0,1)|\mathbf{x}\right)\leq1-\mathbb{P}\left(\mathbf{y}=(1,1)|\mathbf{x}\right)-\mathbb{P}\left(\mathbf{y}=(0,0)|\mathbf{x}\right)$$

and

$$\begin{aligned}\mathbb{P}\left(\mathbf{y}=(0,1)|\mathbf{x}\right)\geq1&-\mathbb{P}\left(\mathbf{y}=(1,1)|\mathbf{x}\right)-\mathbb{P}\left(\mathbf{y}=(0,0)|\mathbf{x}\right)\\&-\mathbb{P}\left((u_1,u_2)\in\times_{i=1,2}\left[-\mathbf{x}_i^\top\beta_i,\ -\mathbf{x}_i^\top\beta_i-\Delta\right]|\mathbf{x}\right),\end{aligned}$$

where, as above,

$$\begin{aligned}\mathbb{P}\left(\mathbf{y}=(0,0)|\mathbf{x}\right)&=F_{u_1,u_2}\left(-\mathbf{x}_1^\top\beta_1,\ -\mathbf{x}_2^\top\beta_2\right)\ \text{ and}\\ \mathbb{P}\left(\mathbf{y}=(1,1)|\mathbf{x}\right)&=1-F_{u_1}\left(-\mathbf{x}_1^\top\beta_1-\Delta_1\right)-F_{u_1}\left(-\mathbf{x}_2^\top\beta_2-\Delta_2\right)\\&\quad+F_{u_1,u_2}\left(-\mathbf{x}_1^\top\beta_1-\Delta_1,\ -\mathbf{x}_2^\top\beta_2-\Delta_2\right).\end{aligned}$$

The upper bound on $\mathbb{P}(\mathbf{y}=(0,1)|\mathbf{x})$ is attained if (0, 1) is always selected when (u_1,u_2) falls in the region of multiplicity. Conversely, the lower bound is attained when (0, 1) is never selected in that case. Hence we subtract the probability that (u_1, u_2) falls in that region from the previous probability. Similar bounds hold for $\mathbb{P}(\mathbf{y}=(1,0)|\mathbf{x})$. The bounds can be modified to allow for mixed-strategy equilibria, as done by Berry & Tamer (2006).

These inequalities define a region of the parameter space in which the true parameter vector resides. When the covariates' support is not as rich as outlined in the previous section to guarantee point identification, this set may (and typically will) be larger than a singleton. Inference may nonetheless be carried out using methods developed in the recent literature on the estimation of partially identified models. Papers on the topic that pay special attention to games with possibly multiple equilibria include Andrews et al. (2004), Ciliberto & Tamer (2009), Beresteanu et al. (2009), Pakes et al. (2011), Galichon & Henry (2011), Moon & Schorfheide (2012), and Chesher & Rosen (2012). Other papers, more general in treatment, could also be cited, and many of the ideas can be applied to games under different information structures. Pakes et al. (2011), for instance, handle games in which there is uncertainty that is symmetric across players and information is complete, as well as games of incomplete information (i.e., some information about payoffs is private information to players). A thorough treatment of the techniques in that literature is beyond the scope of this review. For a recent introduction to that literature, readers are referred to Tamer (2010), and references therein.

3.4. Additional Topics

I close this section with a short discussion. First, much of the work dealing with multiplicity has focused on discrete games. Many issues encountered in discrete games will nevertheless also appear when the action space is continuous. These models are particularly relevant in the study of pricing games among firms, for example. If more than one solution for a given realization of the payoff structure is possible, observed outcomes will be a mixture over equilibria, having the equilibrium selection mechanism as the mixing distribution. Most of the ideas above in principle can be applied in those settings (see, e.g., Pakes et al. 2011), but future work may reveal further advantages or disadvantages of environments with continuous action spaces.

One notable reference in econometrics focusing on continuous action spaces and multiple equilibria is Echenique & Komunjer (2009). These authors analyze a coordination game with continuous actions and use equilibrium results typically found in the literature of supermodular games. In this case, Tarski's fixed-point theorem can be employed to establish maximal and minimal equilibria. Certain features of the structure under study (e.g., monotone comparative statics) can then be econometrically tested using these extremal equilibria (see also de Paula 2009 and Lazzati 2012 for econometric multiagent models using Tarski's fixed-point theorem and Molinari & Rosen 2008 for a discussion of identification in supermodular games with discrete actions). A similar environment (with payoff complementarities and discrete action space) also appears for the empirical analyses of Ackerberg & Gowrinsankaran (2006), in which an equilibrium selection mechanism with support on the best and worst equilibria is incorporated into the econometric model, and of Jia (2008), in which one of the extremal equilibria is assumed to be selected. Jia (2008) also exploits a constructive version of Tarski's theorem for computation.

4. GAMES OF INCOMPLETE INFORMATION

The previous section presents ideas for the analysis of games of complete information. There, individuals are assumed to know the payoff realizations of every other player. This is not always an adequate assumption. In the case of firms contemplating entry into a particular market, at least some part of the cost structure is private information to each firm. It turns out that games in which information is incomplete and agents possess at least partially private information about payoffs provide different avenues for identification and estimation of the structure of interest when there are multiple equilibria.

I start with the same payoff structure as in **Table 1** but assume that u_i is known only to person i, where $i = 1, 2$. As in the previous section, I assume that roles or labels can be assigned to players. Given realizations of $\mathbf{x}$ and $\mathbf{u}$, agents now decide on their actions based on the payoffs they expect, given the distribution of types for the other individual involved in the game (i.e., the private-information components of their payoffs, u_i). As before, I consider only pure-strategy equilibria for simplicity. Given i's opponent's strategy, the best response dictates that

$$y_i = 1 \quad \text{if} \quad \mathbf{x}_i^\top \beta_i + \mathbb{P}(y_j = 1|\mathbf{x}, u_i)\Delta_i + u_i \geq 0, \quad j \neq i, \tag{3}$$

and 0 otherwise. The term $\mathbb{P}(y_j = 1|\mathbf{x}, u_i)$ accommodates i's beliefs about the other person's type given i's information set (comprising $\mathbf{x}$ and u_i) and the other person's strategy about choosing 0 or 1. A Bayes-Nash equilibrium (in pure strategies) is characterized by mutual best responses of players 1 and 2 and consistent beliefs.

When information is incomplete, the expression above highlights that the joint distribution of the latent variables (u_1, u_2) plays an important role. Unless these variables are independent, $\mathbb{P}(y_j = 1|\mathbf{x}, u_i)$ will be a nontrivial function of u_i, $i = 1, 2$. When u_1 and u_2 are not independent, player i's type (i.e., u_i) is informative about the opponents type (i.e., u_j) and, in equilibrium, will likely affect expectations of the opponent's play. If, alternatively, $u_1 \perp\!\!\!\perp u_2|\mathbf{x}$, i's expectations about j's action will not depend on u_i as i's type brings no hint about u_j. Because the equilibrium y_i depends solely on u_i, conditional on $\mathbf{x}$, this assumption also implies that y_1 and y_2 are conditionally independent, given $\mathbf{x}$ for a particular equilibrium. Unless otherwise noted, and except for the last segment in this section, I assume that the private shocks are (conditionally) independent within a game. In this case, we can write $\mathbb{P}(y_j = 1|\mathbf{x}, u_\mathrm{i}) = p_j(\mathbf{x})$, and equilibrium choice probabilities should solve the following system of equations:

$$p_i(\mathbf{x}) = 1 - F_{u_i|\mathbf{x}}\left(-\mathbf{x}_i^\top \beta_i - p_j(\mathbf{x})\Delta_i|\mathbf{x}\right), \quad i = 1, 2, \; i \neq j. \tag{4}$$

One can easily see that this system might admit multiple solutions. Assume, for example, a symmetric payoff structure such that $\Delta_1 = \Delta_2 = 10.5$, u_1 and u_2 follow independent logistic distributions, and $\mathbf{x}_1^\top \beta_1$ and $\mathbf{x}_2^\top \beta_2$ are drawn such that $\mathbf{x}_1^\top \beta_1 = \mathbf{x}_2^\top \beta_2 = -3.5$. Then $p_1(\mathbf{x}) = p_2(\mathbf{x}) = 0.047$ and $p_1(\mathbf{x}) = p_2(\mathbf{x}) = 0.204$ are symmetric solutions to the system above. (A third solution in which both players choose 1 with probability close to 1 also exists.)

In general the constituents of this model are not point identified. To see this, denote by $\mathcal{E}(\mathbf{x}, \theta, F)$ the set of equilibria for a given realization of $\mathbf{x}$ and primitives (θ, F). Note that the equilibrium set now depends not on $\mathbf{u}$ (which are only privately observed) but on its distribution. Because for now I retain the assumption that $\mathrm{u}_1 \perp\!\!\!\perp u_2|\mathbf{x}$, recall that in equilibrium y_1 and y_2 are conditionally independent, and an equilibrium is hence characterized by the marginal distributions of y_i given $\mathbf{x}$, $i = 1, 2$. Let $(p_1^k(\mathbf{x}), p_2^k(\mathbf{x}))_{k=1}^{|\mathcal{E}(\mathbf{x},\theta,F)|}$ be an enumeration of the equilibrium set $\mathcal{E}(\mathbf{x}, \theta, F)$. Then the distribution of outcomes conditional on covariates $\mathbf{x}$ is given by

$$\mathbb{P}(y_1, y_2|\mathbf{x}) = \sum_{k=1}^{|\mathcal{E}(\mathbf{x},\theta,F)|} \lambda\left(k|\mathcal{E}(\mathbf{x}, \theta, F), \mathbf{x}, \theta, F\right) p_1^k(\mathbf{x})^{y_1}\left(1 - p_1^k(\mathbf{x})\right)^{1-y_1} p_2^k(\mathbf{x})^{y_2}\left(1 - p_2^k(\mathbf{x})\right)^{1-y_2},$$

where $\lambda(\cdot)$ is the equilibrium selection mechanism and places probability $\lambda(k|\mathcal{E}(\mathbf{x}, \theta, F), \mathbf{x}, \theta, F)$ on the selection of equilibrium k. This is a mixture of independent random variables, and results such as those by Hall & Zhou (2003) can be used to demonstrate that the component distribution $(p_1^k(\mathbf{x}), p_2^k(\mathbf{x}))_{k=1}^{|\mathcal{E}(\mathbf{x},\theta,F)|}$ and mixing probabilities are not point identified by the model if $|\mathcal{E}(\mathbf{x}, \theta, F)|$ is strictly greater than 2. Similar results can be obtained for more players (e.g., using Hall et al. 2005): If the number of equilibria on the support of the selection mechanism is large relative to the number of players, the structure is not point identified by the model (see, e.g., the supplementary appendix in de Paula & Tang 2012).

As in the previous section, I now discuss different approaches to inference in games of incomplete information like the one just outlined above.

4.1. Degenerate Equilibrium Selection Mechanism

The previous discussion underscores the benefits of further restrictions on the equilibrium selection mechanism for the econometric analysis of incomplete-information games with possibly many equilibria. One common strategy is to assume that

$$\lambda(k|\mathcal{E}(\mathbf{x},\theta,F),\mathbf{x},\theta,F)=1_{k=\mathcal{K}}$$

for some $\mathcal{K}\in\mathcal{E}(\mathbf{x},\theta,F)$. In words, whenever primitives and covariates coincide for two games, thus inducing an identical equilibrium set, the same equilibrium is played in these two games. One can nevertheless be agnostic about which equilibrium is selected [i.e., which element of $\mathcal{E}(\mathbf{x}, \theta, F)$ is selected].

When is it realistic to assume that the same equilibrium is played across games? As Mailath (1998) points out, "the evolution of conventions and social norms is an instance of players learning to play an equilibrium." If an equilibrium is established as a mode of behavior by past play, custom, or culture, this equilibrium becomes a focal point for those involved. When observed games are drawn from a population that is culturally or geographically close, sharing similar norms and conventions, one would expect this assumption to be adequate.

If a single equilibrium occurs in the data for every given payoff configuration, this restriction essentially allows one to treat the model in standard ways for identification and inference. Once identification is guaranteed, one possible strategy for estimation is the two-step estimator proposed by Bajari et al. (2010a). The strategy builds on the earlier work by Hotz & Miller (1993) for the estimation of empirical (individual) dynamic discrete choice models and later work by Aguirregabiria & Mira (2007) and Bajari et al. (2007) for dynamic games (in which information is incomplete). It also resembles estimators proposed in the social interactions literature and covered below. I describe the procedure using the example in **Table 1** (still with conditionally independent latent variables u_1 and u_2). Bajari et al. (2010a) present the method in more generality.

In the first step, the (reduced-form) conditional choice probabilities for the various outcomes in the action space are estimated nonparametrically: $\widehat{\mathbb{P}}(y_i|\mathbf{x}), i=1, 2$. This can be done using sieves or kernel methods.

In the second step, estimates for the structural parameters are obtained from the conditional choice probability estimates. If a unique equilibrium is played in the data for the same realizations of the payoff matrix, p_i $(\mathbf{x})$ and the expected payoff for player i of choosing 1 using equilibrium beliefs [i.e., $\mathbf{x}_i^\top\beta_i+p_j(\mathbf{x})\Delta_i$] are in one-to-one correspondence. In the example, this correspondence is simply obtained as

$$\mathbf{x}_i^\top\beta_i+p_j(\mathbf{x})\Delta_i=-F^{-1}_{u_i|\mathbf{x}}(1-p_i(\mathbf{x})|\mathbf{x}).$$

Given a sample of G i.i.d. games, the estimation then proceeds to minimize the following least-squares function with respect to θ:

$$\sum_{g=1}^{G}\left[-F^{-1}_{u_i|\mathbf{x}}\left(1-\widehat{p}_i(\mathbf{x}_g)|\mathbf{x}_g\right)-\mathbf{x}_{ig}^\top\beta_i-\widehat{p}_j(\mathbf{x}_g)\Delta\right]^2 h(\mathbf{x}_g),$$

where $h(\cdot)$ is again an appropriately chosen weight function. Under adequate conditions, the estimator for θ converges at parametric rates as pointed out by Bajari et al. (2010a).

The assumption of a unique equilibrium in the data is crucial to travel from the conditional choice probabilities to the second-step estimator as it guarantees (together with other additional restrictions) that $p_i(\mathbf{x})$ and the expected payoff for player i of choosing 1 using equilibrium beliefs $\left[\text{i.e., } \mathbf{x}_i^\top \beta_i + p_j(\mathbf{x})\Delta_i\right]$ are in a one-to-one relationship. Because of this, the above procedure bypasses the calculation of all equilibria for each possible value of θ, much as Hotz & Miller (1993) avoid the computation of a dynamic program in the estimation of individual dynamic discrete choice models. Of course, as long as there is a unique selected equilibrium in the data, other estimators that rely on uniqueness, such as those in Aradillas-Lopez (2010), can also be used. The assumption is also employed, for example, by Kasy (2012), who proposes a procedure to perform inference on the number of feasible equilibria in a game: Even if the equilibrium selection mechanism is assumed to be degenerate, once the primitives are estimated [i.e., $(\beta_1, \beta_2, \Delta_1, \Delta_2)$ in my example, but more general nonparametric functions in Kasy 2012], one can infer the number of equilibria for the estimated particular payoff structure at a given realization of covariates.

As discussed by Bajari et al. (2010a), when covariates $\mathbf{x}$ have continuous support, the requirement that the data come from a unique equilibrium imposes subtle additional restrictions, as the implementation of the above estimation strategy typically demands smoothness of the first-stage estimator with respect to the covariates. **Figure 3** sketches the possible ways in which this prerequisite might fail. The horizontal axis represents a generic continuous covariate $x \in \mathbb{R}$, and the vertical axis schematically represents the equilibrium conditional choice probabilities $p(x)$. The graph sketches the potential equilibrium probabilities corresponding to each x. For $x > x'$, there are three potential equilibria, whereas for $x < x'$, there is only one. Because the number of equilibria $|\mathcal{E}(x, \theta, F)|$ may change with x, as shown in **Figure 3**, small changes in the covariates may provoke abrupt changes in the cardinality of the equilibrium set. In the figure, this happens around x'. At this bifurcation point, the selected equilibrium conditional choice probability selected $p(x)$ may vary nonsmoothly as is the case for the equilibrium selection mechanism. An implicit condition then is that such bifurcations happen in ways that do not affect the estimation. It is also possible that minute changes in the covariates $\mathbf{x}$ may tip the selection of the equilibrium observed in the data

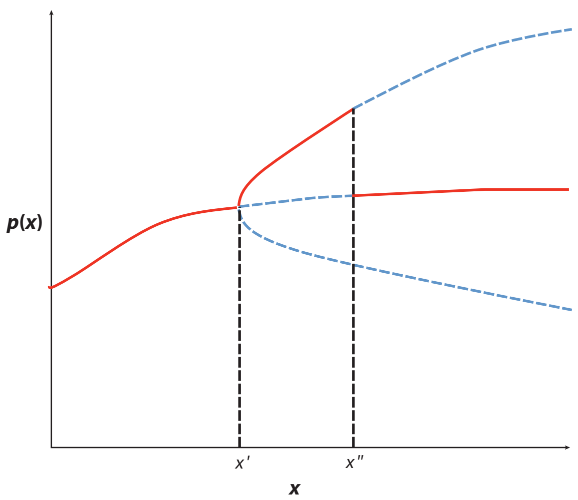

Figure 3

Equilibrium choice probability. The red line marks the equilibrium that is selected in the data and stands for a possible (degenerate) equilibrium selection mechanism.

in a discontinuous manner. In the figure, this happens at x''. Both points (x' and x'') present possible challenges to the smoothness requirement.

In the presence of nonsmooth equilibrium selection mechanisms, the two-step estimator suggested above will perform poorly. Aguirregabiria & Mira (2012) produce some Monte Carlo evidence of the deficiencies of this estimator when there are discontinuities in the equilibrium selection mechanism and suggest an alternative strategy, combining a pseudo-maximum likelihood estimator with a genetic algorithm.

The assumption of a degenerate equilibrium selection mechanism is usually invoked as well in the estimation of dynamic incomplete-information games. In these environments, an action profile affects not only the players' payoffs, but also the evolution of the state variables. The prescribed equilibrium strategies in this case generalize Equation 3 to account for future repercussions. Common estimation strategies for dynamic games such as those presented by Aguirregabiria & Mira (2007), Bajari et al. (2007), Pakes et al. (2007), and Pesendorfer & Schmidt-Dengler (2008) adapt well-known procedures for the estimation of single-agent Markov decision problems, such as that by Hotz & Miller (1993). In essence, these estimators will rely on the existence of a one-to-one mapping between value functions and conditional choice probabilities. For this to hold, they rely on there being a single equilibrium in the data for games with identical payoff realizations as well as other assumptions usually invoked to guarantee the existence of such mapping (e.g., the additive separability of u_i).

4.2. Identifying Power of Multiplicity

Sweeting (2009) follows a different route and incorporates the equilibrium selection mechanism in the estimation (see references in the previous section for similar strategies in complete-information models). He studies coordination (or not) on the timing of commercial breaks among radio stations (i.e., players) within a geographical market (i.e., game). In terms of Sweeting's (2009, p. 713) application, the intuition is as follows:

> If stations want to coordinate then there may be an equilibrium where stations cluster their commercials at time 1 and another equilibrium where they cluster their commercials at time 2. . . . If stations want to play commercials at different times then we would expect to observe excess dispersion within markets (market distributions less concentrated than the aggregate) rather than clustering.

In his discussion, he focuses on symmetric (stable) pure-strategy Bayesian-Nash equilibria in a coordination game. To keep matters simple, I consider again the two-player game with payoff structure from **Table 1**. Assume that $\mathbf{x}_i^\top \beta_i = \alpha$, $i = 1, 2$, and $\Delta = \Delta_1 = \Delta_2 \geq 0$. As is the case above, private information is assumed to be independent across agents and from $\mathbf{x}$. Given the parameterization in his model, multiplicity always produces three symmetric equilibria, two of which are stable. Denote the choice probabilities in these two stable equilibria p^l and p^h, where $p_l < p_h$.

To complete the model, Sweeting assumes that the stable equilibria are selected with probability λ_k, $k = l, h$, whereas the unstable equilibrium is never selected. The outcome probability is then

$$\mathbb{P}(y_1, y_2|\mathbf{x}) = \sum_{k=l,h} \lambda_k \binom{2}{y_1 + y_2} \left(p^k\right)^{y_1+y_2} \left(1 - p^k\right)^{2-y_1-y_2}, \quad \lambda_l + \lambda_h = 1. \tag{5}$$

Notice that with a unique equilibrium in the data, Equation 5 corresponds to the distribution of a binomial random variable (with parameters 2 and p). Using simulations, Sweeting (2009, p. 723) notes that when $\Delta > 0$, "a mixture generates greater variance in the number of stations choosing a particular outcome than can be generated by a single binomial component." As he points out, this suggests that multiplicity provides additional information about the payoff structure of the game under analysis.

De Paula & Tang (2012) formalize and generalize this idea in many directions. For the basic insight, take the expression in Equation 4 and compare two equilibria where $p_j^h(\mathbf{x}) > p_j^l(\mathbf{x})$. If $\Delta_i > 0$, it has to be the case that $p_i^h(\mathbf{x}) > p_i^l(\mathbf{x})$. (This is because $F_{u_i|\mathbf{x}}$ is increasing as it is a CDF.) Conversely, if $\Delta_i < 0$, one must have $p_i^h(\mathbf{x}) < p_i^l(\mathbf{x})$. With two players, multiple equilibria are possible only if both Δ_1 and Δ_2 have the same sign (in contrast with the complete-information case). Consequently, if Δ_1 and Δ_2 are positive, $p_1^h(\mathbf{x}) > p_1^l(\mathbf{x})$ if and only if $p_2^h(\mathbf{x}) > p_2^l(\mathbf{x})$. In other words, the equilibrium choice probabilities by player 2 are an increasing function of the equilibrium choice probabilities by player 1. Conversely, when both Δ_1 and Δ_2 are negative, the equilibrium choice probabilities by player 2 are a decreasing function of the equilibrium choice probabilities by player 1. Because of this monotonicity, the correlation of players' actions across games will be positive when both Δ_1 and Δ_2 are positive and negative otherwise. Interestingly, because private types $u_i, i = 1, 2$, are assumed to be independent (conditionally on $\mathbf{x}$), the actions will be uncorrelated and uninformative about $\Delta_i, i = 1, 2$, if there is only one equilibrium in the data. This gives the following specialization of proposition 1 in de Paula & Tang (2012):

> **Proposition (de Paula & Tang 2012):** Suppose $u_1 \perp\!\!\!\perp u_2|\mathbf{x}$. (*a*) For any given $\mathbf{x}$, multiple equilibria exist in the data-generating process if and only if $\mathbb{E}[y_1 y_2|\mathbf{x}] \neq \mathbb{E}[y_1|\mathbf{x}]\mathbb{E}[y_2|\mathbf{x}]$ and (*b*) if that is the case for some $\mathbf{x}$,

$$sign\big(\mathbb{E}[y_1 y_2|\mathbf{x}] - \mathbb{E}[y_1|\mathbf{x}]\mathbb{E}[y_2|\mathbf{x}]\big) = sign(\Delta_i),\ i = 1,\ 2.$$

In de Paula & Tang (2012), the baseline utility $\mathbf{x}_i^\top \beta_i$ can be replaced by a generic function of $\mathbf{x}_i$, and the parameter Δ_i is also allowed to depend on $\mathbf{x}_i$, but I keep with the payoffs in **Table 1** for simplicity and consistency. The result also holds for an arbitrary number of players and allows for asymmetric equilibria and situations in which the number of equilibria is unknown. [Kline (2012) extends the idea of this result to environments with complete information.]

The subtle implication of the proposition is that multiplicity is informative about aspects of the model in ways that uniqueness is not. Of course, for the set of covariate values in which there is only one equilibrium in the data, the estimation can then be performed under various restrictions on utilities, $\Delta_i, i = 1, 2$, and $F_{\mathbf{u}|\mathbf{x}}$ (for more details, see Berry & Tamer 2006, Aradillas-Lopez 2010, Bajari et al. 2010a).

For games with more than two players, de Paula & Tang (2012) rely on this result to suggest a test for the hypothesis that there are multiple solutions in the data and, conditional on there being more than one equilibrium, on the sign of $\Delta_i, i = 1, 2$. To implement this test, they build on recent developments in the statistical literature on multiple comparisons (Romano & Wolf 2005; see the recent survey in Romano et al. 2010). Based on his model, Sweeting (2009) also suggests tests for multiple symmetric equilibria when the cardinality of the equilibrium set is known. The first is based on calculating the percentage of pairs of players whose actions are correlated. The other is a test of the null of a unique equilibrium against the alternative of exactly two equilibria using a maximum likelihood estimator.

Finally, I note that the essential assumption of the conditional independence of the latent variables **u** is also commonly found in dynamic games of incomplete information. Optimal decision rules in those settings involve not only equilibrium beliefs but also continuation value functions that may change across equilibria. Nevertheless, the characterization of optimal policy rules in that context suggests that the existence of a unique equilibrium in the data can still be detected by the lack of correlation in actions across players of a given game, as presented here in the case of the static game. [The identification of $sign(\Delta_i)$ would require additional restrictions, however.] Because most of the known methods for the semiparametric estimation of incomplete information (static or dynamic) rely on the existence of a single equilibrium in the data (see above), a formal test for the assumption of a unique equilibrium in the data-generating process can be quite useful.

4.3. Game-Level Heterogeneity and Correlated Private Signals

To establish proposition 1 in de Paula & Tang (2012), it is paramount that the latent variables be conditionally independent. Any association between u_1 and u_2 will lead to correlation in actions even under a unique equilibrium, but it will also change the nature of equilibrium decision rules in important ways [i.e., $\mathbb{P}(y_j = 1|\mathbf{x}, u_i)$ in Equation 3 is now a nontrivial function of u_i]. Aradillas-Lopez (2010) suggests an estimation procedure to handle cases allowing for correlated private values but relies on the assumption that a single equilibrium is played in the data. Another example is provided by Wan & Xu (2010), who also require that a unique (monotone) Bayesian-Nash equilibrium be played in the data. As long as one is comfortable with the assumption of a single equilibrium in the data described above, these methods can be used for estimation. To my knowledge, general results along the lines of de Paula & Tang (2012) that allow for both correlation of private values and multiplicity have not been proposed [although see the working-paper version of that article (de Paula & Tang 2011) for a discussion of some possible characterizations].

One empirically important form of association between latent variables that has received some attention amounts to decomposing u_i into a publicly observed component ε_i and a privately observed component ν_i. Whereas we retain the assumption that ν_1 and ν_2 are (conditionally) independent, I assume for simplicity that the publicly observed errors take the form of a game-level shock, $\varepsilon_1 = \varepsilon_2 = \varepsilon$. In the empirical games literature, the presence of this market-level shock is typically referred to as unobserved heterogeneity (even though ν_1 and ν_2 are themselves heterogeneous and unobserved within and across games).

The presence of ε prevents one from employing the results in de Paula & Tang (2012), much as the correlation in fully private u_1 and u_2 would. The practical solution is to account for it by modeling the distribution of publicly observed latent shocks when a cross section is employed or using a market-invariant fixed or (possibly correlated) random effect, again under the assumption that a unique equilibrium occurs in the data for each particular market when panel data are available. In a dynamic context, for instance, Aguirregabiria & Mira (2007) introduce game-level shocks as a correlated random effect with finite support [in the tradition of Heckman & Singer (1984) for duration models]. Grieco (2012) also models the distribution of publicly observed shocks in a static game estimated on a cross section.

Alternatively, Bajari et al. (2010a, p. 475) note that "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." The requirement of a long panel helps circumvent the incidental parameters problem in this fairly nonlinear context. I also note that, provided the equilibrium is the same for the different editions of the game within a particular market, different equilibria are allowed across different markets. Similar ideas also appear in other papers in the literature, such as Bajari et al. (2007) and Pesendorfer & Schmidt-Dengler (2008). Even in the absence of long panels, Bajari et al. (2010a) suggest a few interesting strategies, such as the use of conditional likelihood methods when the ν errors are logistic or the use of Manski's (1987) panel data rank estimator. These proposals are not further developed in that paper, and some caution may be warranted given the simultaneous-equation nature of the problem [i.e., the presence of $p_j(\mathbf{x})$ among the regressors in Equation 4 might have repercussions for the two-step procedure as $p_j(\mathbf{x})$ is also affected by the game-level shock ε].

5. MODELS OF SOCIAL INTERACTIONS

Social interactions models have gained widespread attention since Manski (1993). Whereas Manski (1993) focuses on linear social interactions systems in which equilibrium is unique, Brock & Durlauf (2001, 2007) consider a model of social interactions with discrete choices in which multiple equilibria are possible. In the model, the value of a particular choice depends on the distribution of actions among other individuals in the community. In what follows, I adapt the expressions of Brock & Durlauf (2001, 2007) to keep with the previous notation above. Normalizing the value of choosing zero to 0, the utility obtained from choosing 1 can be written as (Brock & Durlauf 2001, p. 238)

$$g_i\big(\mathbb{P}(\mathbf{y}_{-i}|\mathbf{x}),\, \mathbf{x},\, \theta\big) + u_i.$$

The term $g_i(\mathbb{P}(\mathbf{y}_{-i}|\mathbf{x}),\, \mathbf{x},\, \theta)$ encompasses what Brock & Durlauf call the private and the social utilities attached to this particular choice (or, given our normalization, the difference in those utilities of choosing 1 over 0). The social utility depends on the conditional probability that i places on the choices of others when making a choice. As before, u_i is a random utility component that is known to the individual but unknown to others and is i.i.d. across agents. Typically, u_i is taken to be logistically distributed.

The probability distribution that i puts on the choices of others is denoted by $\mathbb{P}(\mathbf{y}_{-i}|\mathbf{x})$ and is determined in equilibrium. Person i's choice is then given by

$$y_i = 1 \text{ if } g_i\big(\mathbb{P}(\mathbf{y}_{-i}|\mathbf{x}),\, \mathbf{x},\, \theta\big) + u_i \geq 0. \tag{6}$$

As with incomplete-information games, an equilibrium will consist of mutual best responses and self-consistent beliefs. Given the nonlinearities in the model, multiplicity is again a possibility. Nevertheless, instead of comparing (equilibrium) expected utilities as in Equation 3, the expression above compares utility functions that depend on the (equilibrium) distributions of actions. These two will differ for general parameterizations (possibly nonlinear in $\mathbb{P}$). This is a case, for example, in which there is preference for conformity, as discussed by Brock & Durlauf (2001, p. 239). This is an important distinction. In the main (linear) parameterization considered by Brock & Durlauf, they nevertheless coincide. This specification postulates that

$$g_i\big(\mathbb{P}(\mathbf{y}_{-i}|\mathbf{x}),\, \mathbf{x},\, \theta\big) = \mathbf{x}_i^{\top}\beta + \Delta\mathbb{E}\left[\frac{\sum_{j\neq i} y_j}{N-1}|\mathbf{x}\right],$$

where N is the group size as before, and β and Δ are parameters to be estimated. Because

$$\mathbf{x}_i^\top \beta + \Delta \mathbb{E}\left[\frac{\sum_{j \neq i} y_j}{N-1} | \mathbf{x}\right] = \mathbb{E}\left[\mathbf{x}_i^\top \beta + \Delta \frac{\sum_{j \neq i} y_j}{N-1} | \mathbf{x}\right],$$

the (equilibrium) expected utility agrees with the utility at the expected (equilibrium) profile of actions. This is, in particular, the best response predicament in my example with two players and two actions under incomplete information when $\beta_1 = \beta_2$ and $\Delta_1 = \Delta_2$. A noteworthy difference between this setup and that in the previous section is that I do not assume that players' roles or labels (e.g., firm identity, husband or wife) can be assigned. This anonymity assumption is more natural in the social interactions systems in which a large number of people is typically observed per group.

Because the equilibrium moment $\mathbb{E}\left[\sum_{j \neq i} y_j/(N-1)|\mathbf{x}\right]$ is observable, it can be estimated by the average choice in this particular game and equilibrium (e.g., Brock & Durlauf 2007, p. 58). When there is a small number of players (e.g., as is the case in industrial organization applications), the choice probabilities will not be reliably estimated by averaging choices within a game. In this case, some combination of observations across games is unavoidable, and I refer the reader to the discussion in the previous section. For most applications of social interactions models, however, the number of agents within a game is much larger. Consequently, the average peer choice in a game will consistently estimate the average action within the game (see Brock & Durlauf 2001, proposition 6). Because average peer choice will display little variability within a group, variation across games can then be exploited to estimate Δ and β.

I now briefly discuss estimation and additional topics in this class of models in two different subsections.

5.1. Estimation

In this context, Bisin et al. (2011), building on Moro (2003), suggest two alternative estimators. The first procedure estimates the game via maximum likelihood, imposing the self-consistency equilibrium conditions. Because here a parameter vector might induce multiple equilibria for a given data set, the estimation proceeds by selecting that equilibrium that maximizes the likelihood (for a given parameter vector) and subsequently maximizing the likelihood function over the parameter space. [This resembles the procedure suggested by Chen et al. (2011), who suggest profiling the equilibrium selection mechanism in a sieve–maximum likelihood procedure.] The second procedure is a plug-in estimator in which $\mathbb{E}\left[\sum_{j \neq i} y_j/(N-1)|\mathbf{x}\right]$ is replaced by the average peer choice within the game, and parameters are then estimated via maximum likelihood. The log-likelihood function in the case of my example is given by

$$\sum_{g=1}^{G} \sum_{i \in g} \ln\left[1 - F_u\left(-\mathbf{x}_i^\top \beta - \Delta \frac{\sum_{j \neq i, j \in g} y_j}{N-1}\right)\right],$$

where $i \in g$ indicates that person i belongs to neighborhood g. This estimator can be seen as a two-step estimator akin to the procedure outlined in the previous section for incomplete-information games. As is the case in the previous section, this estimator avoids the calculation of all the equilibria for a given parameter value. In response to the computational difficulties that this entails, Bisin et al. advise using the two-step estimator as an initial guess in the direct estimator. Similar two-step procedures with (possibly) group-level unobservables are also suggested by Shang & Lee

(2011). Bisin et al. present Monte Carlo evidence in a model with possibly many equilibria for certain parameter configurations that highlights the computational costs and statistical properties of the two estimators. Because the asymptotic approximations rely on $N \to \infty$, I must point out that the econometric estimators in such large population games might present some delicate issues given the dependence in equilibrium outcomes within a game as the number of players grows. This is a topic of ongoing research (see, e.g., Menzel 2010, Song 2012).

5.2. Additional Topics

As is the case in the previous section, group-level unobserved heterogeneity is potentially important in many applications. Ignoring it essentially rules out an important channel of unobserved contextual effects (or correlated effects) in the terminology coined by Manski (1993). Brock & Durlauf (2007, section 4) also discuss a series of potential scenarios that would allow the model to identify (at least partially) the parameters of interest. These include the use of panels, restrictions on the distribution of unobserved group shocks (i.e., large support, stochastic monotonicity, unimodality), and other features of the model (i.e., linearity in terms of group shocks). Some of these restrictions may also be useful in the incomplete-information games described above.

6. DISCUSSION

I finish this review with a brief discussion on postestimation analysis and potential areas of development and applications.

6.1. Counterfactuals and Postestimation

Once in possession of point or set estimates for the parameters of interest, one may be interested in the effect of counterfactual policies. After all, this goal is behind the very development of the Cowles Commission agenda delineated in Section 2. If the equilibrium selection mechanism is included in the structure as an estimation object, one has a complete model that can be simulated to generate counterfactual distributions after the introduction of alternative policies (see, e.g., Bajari et al. 2010b, p. 1537).

As pointed out above, one must be cautious about the policy invariance of the estimated equilibrium selection mechanism. Quoting Pakes et al. (2007, p. 375),

> though our assumptions are sufficient to use the data to pick out the equilibrium that was played in the past, they do not allow us to pick out the equilibrium that would follow the introduction of a new policy. On the other hand, the . . . estimates should give the researcher the ability to examine what could happen after a policy change (say, by examining all possible post-policy-change equilibria).

This is done, for example, by Ciliberto & Tamer (2009), who provide a range of possible counterfactual outcomes for an intervention repealing a particular piece of legislation on the entry behavior of airlines in their empirical analysis. Of course, any counterfactual analysis (with or without an estimated equilibrium selection policy) would require the computation of all equilibria, although only for the range of parameters estimated.

An additional motivation for including the equilibrium selection mechanism as an estimation object is to retrospectively learn about the process and covariates determining which equilibria come to be played in the data. Even when the game is estimated under the assumption that a unique equilibrium is played in the data, the possession of estimated parameters allows one to go back and calculate all the potential equilibria for a particular parametric configuration. In their study of stock analysts' recommendations, for instance, Bajari et al. (2010a) notice the existence of multiple equilibria before New York Attorney General Eliot Spitzer launched a series of investigations on conflict of interest, with one equilibrium yielding much more optimistic ratings than those granted in the equilibrium post-Spitzer.

6.2. Potential Avenues for Future Research

Above I try to present many tools used in the econometric analysis of games with multiple equilibria. There is nevertheless still much to be understood in these settings. One interesting avenue that appears in some papers cited here is the connection with panel data methods. Just as the distribution of outcomes in game theoretic models is a mixture over equilibrium-specific outcome distributions under multiplicity, the observable distribution of outcomes in panel data models is a mixture over the distribution of individuals effects. Important idiosyncrasies such as the (typical) finiteness of the equilibrium set (which would correspond to a finite support for the individual effects) may help bring in interesting technical results in the panel literature to shed light on some properties of econometric game theoretic models. Examples of such studies include Hahn & Moon (2010) and Bajari et al. (2011). Here a important caveat, mentioned above, is that the cardinality of the equilibrium set, $|\mathcal{E}(\theta, F, \mathbf{x})|$, will depend on the covariates and parametric configurations, whereas the support of the individual effects in the usual panel data model suffers no such restrictions. This might introduce important complications.

In a similar fashion, Grieco (2012) and Chen et al. (2011) suggest treating the equilibrium selection mechanisms as a (possibly infinite dimensional) nuisance parameter that is concentrated out in a profile sieve–maximum likelihood estimator procedure aimed at estimating semi-parametric partially identified models. Again, in this case, the dependence of the cardinality of the equilibrium set, $|\mathcal{E}(\theta, F, \mathbf{x})|$, on the covariates and parametric configurations might introduce subtle complications, as the class of functions that contain the equilibrium set might have to vary with θ, F, and $\mathbf{x}$ to accommodate this dependence.

6.3. Applications

Above I discuss some of the commonly analyzed models in which multiplicity shows up prominently in the econometric analysis of games and the usual procedures to handle this feature. I left out other models in which nonuniqueness is rampant but so far has been given less attention. One example includes static models of network formation à la Jackson & Wolinsky (1996). Although the solution concepts used in that literature differ from those used in the noncooperative games studied above, it is not uncommon to encounter many possible stable networks for a given payoff configuration. Those difficulties are perhaps compounded in econometric analogs of network-formation models because equilibrium existence results for heterogeneous payoffs are less readily available in the theory literature, and computational challenges in their analysis are notorious.

For space considerations, above I only briefly mention empirical applications and focus on more methodological aspects of the analysis. Many papers cited here provide (or actually even focus on) empirical questions (e.g., Ciliberto & Tamer 2009, Sweeting 2009), and many other examples in the empirical industrial organization or peer effects literature can be enumerated.

Recent applications in other areas of economics are available (e.g., Card & Giuliano 2011, Todd & Wolpin 2012) and are likely to become more common.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I thank Tim Halliday, Chris Julliard, Elie Tamer, and Matthew Weinberg for discussions and suggestions. I am grateful to the Reviewing Editor, Chuck Manski, who provided detailed comments that markedly improved this review. Financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001 and the National Science Foundation through grant award number SES 1123990 is gratefully acknowledged.

LITERATURE CITED

Ackerberg D, Gowrinsankaran G. 2006. Quantifying equilibrium network externalities in the ACH banking industry. *Rand J. Econ.* 37:738–61

Aguirregabiria V, Mira P. 2007. Sequential estimation of dynamic discrete games. *Econometrica* 75:1–53

Aguirregabiria V, Mira P. 2012. *Structural estimation of games with multiple equilibria in the data*. Work. Pap., Univ. Toronto

Andrews D, Berry S, Jia P. 2004. *Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location*. Work. Pap., Yale Univ., New Haven, CT

Aradillas-Lopez A. 2010. Semi-parametric estimation of simultaneous games with incomplete information. *J. Econom.* 157:409–31

Aradillas-Lopez A, Tamer E. 2008. The identification power of equilibrium in simple games. *J. Bus. Econ. Stat.* 26:261–310

Bajari P, Benkard L, Levin J. 2007. Estimating dynamic models of imperfect competition. *Econometrica* 75:1331–70

Bajari P, Hahn J, Hong H, Ridder G. 2011. A note on semiparametric estimation of finite mixtures of discrete choice models with application to game theoretic models. *Int. Econ. Rev.* 52:807–24

Bajari P, Hong H, Krainer J, Nekipelov D. 2010a. Estimating static games of incomplete information. *J. Bus. Econ. Stat.* 28:469–82

Bajari P, Hong H, Ryan S. 2010b. Identification and estimation of a discrete game of complete information. *Econometrica* 78:1529–68

Beresteanu A, Molchanov I, Molinari F. 2009. *Sharp identification regions in models with convex predictions: games, individual choice, and incomplete data*. CeMMAP Work. Pap. CWP27/09, Inst. Fisc. Stud., London

Berry S, Tamer E. 2006. Empirical models of oligopoly entry. In *Advances in Economics and Econometrics: Theory and Applications; Ninth World Congress of the Econometric Society*, ed. R Blundell, W Newey, T Persson, pp. 46–85. Cambridge, UK: Cambridge Univ. Press

Berry ST. 1992. Estimation of a model of entry in the airline industry. *Econometrica* 60:889–917

Bisin A, Moro A, Topa G. 2011. *Empirical content of models with multiple equilibria in economies with social interactions*. NBER Work. Pap. 17196

Bjorn P, Vuong Q. 1984. *Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation*. Caltech Work. Pap. 537, Calif. Inst. Technol, Pasadena

Bresnahan TF, Reiss PC. 1990. Entry in monopoly markets. *Rev. Econ. Stud.* 57:57–81

Bresnahan TF, Reiss PC. 1991. Entry and competition in concentrated markets. *J. Polit. Econ.* 99:977–1009

Brock W, Durlauf S. 2001. Discrete choice with social interactions. *Rev. Econ. Stud.* 62:235–60
Brock W, Durlauf S. 2007. Identification of binary choice models with social interactions. *J. Econom.* 140:52–75
Card D, Giuliano P. 2011. *Peer effects with multiple equilibria in the risky behavior of friends*. NBER Work. Pap. 17088
Chen X, Tamer E, Torgovitsky A. 2011. *Sensitivity analysis in semiparametric likelihood models*. Work. Pap., Yale Univ., New Haven, CT
Chesher A, Rosen A. 2012. *Simultaneous equations models for discrete outcomes: coherence, completeness, and identification*. CeMMAP Work. Pap., Inst. Fisc. Stud., London
Ciliberto F, Tamer E. 2009. Market structure and multiple equilibria in airline markets. *Econometrica* 77:1791–828
de Paula A. 2009. Inference in a synchronization game with social interactions. *J. Econom.* 148:56–71
de Paula A, Tang X. 2011. *Inference of signs of interaction effects in simultaneous games with incomplete information*. Work. Pap. 11-003, Penn Inst. Econ. Res., Univ. Penn., Philadelphia
de Paula A, Tang X. 2012. Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica* 80:143–72
Echenique F, Komunjer I. 2009. Testing models with multiple equilibria by quantile methods. *Econometrica* 77:1281–98
Galichon A, Henry M. 2011. Set identification in models with multiple equilibria. *Rev. Econ. Stud.* 78:1264–98
Grieco P. 2012. *Discrete choice games with flexible information structure: an application to local grocery markets*. Work. Pap., Penn. State Univ., University Park
Hahn J, Moon R. 2010. Panel data models with finite number of multiple equilibria. *Econom. Theory* 26:863–81
Hahn J, Tamer E. 2004. *Corrigendum on "Incomplete bivariate discrete response model with multiple equilibria."* Work. Pap., Northwestern Univ., Evanston, IL
Hall P, Neeman A, Pakyari R, Elmore R. 2005. Nonparametric inference in multivariate mixtures. *Biometrika* 92:667–78
Hall P, Zhou X-H. 2003. Nonparametric estimation of component distributions in a multivariate mixture. *Ann. Stat.* 31:201–24
Heckman JJ. 1990. Varieties of selection bias. *Am. Econ. Rev.* 80:313–18
Heckman JJ, Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52:271–320
Honoré B, de Paula A. 2010. Interdependent durations. *Rev. Econ. Stud.* 77:1138–63
Hotz J, Miller R. 1993. Conditional choice probabilities and the estimation of dynamic models. *Rev. Econ. Stud.* 60:397–429
Jackson M, Wolinsky A. 1996. A strategic model of social and economic networks. *J. Econ. Theory* 71:44–74
Jia P. 2008. What happens when Wal-Mart comes to town: an empirical analysis of the discount retailing industry. *Econometrica* 76:1263–316
Jovanovic B. 1989. Observable implications of models with multiple equilibria. *Econometrica* 57: 1431–37
Kasy M. 2012. *Nonparametric inference on the number of equilibria*. Work. Pap., Harvard Univ., Cambridge, MA
Khan S, Tamer E. 2010. Irregular identification, support conditions and inverse weight estimation. *Econometrica* 78:2021–42
Kline B. 2012. *Identification of complete information games*. Work. Pap., Univ. Texas, Austin
Koopmans T, Reiersol O. 1950. The identification of structural characteristics. *Ann. Math. Stat.* 21:165–81
Lazzati N. 2012. *Treatment response with social interactions: partial identification via monotone comparative statics*. Work. Pap., Univ. Michigan, Ann Arbor
Mailath G. 1998. Do people play Nash equilibrium? Lessons from evolutionary game theory. *J. Econ. Lit.* 36:1347–74
Manski C. 1987. Semiparametric analysis of random effects linear models from binary response data. *Econometrica* 55:357–62

Manski C. 1988. Identification of binary response models. *J. Am. Stat. Assoc.* 83:729–38

Manski C. 1993. Identification of endogenous social effects: the reflection problem. *Rev. Econ. Stud.* 60:531–42

Menzel K. 2010. *Inference for large games with exchangeable players*. Work. Pap., New York Univ.

Molinari F, Rosen A. 2008. The identification power of equilibrium in games: the supermodular case. *J. Bus. Econ. Stat.* 26:297–302

Moon R, Schorfheide F. 2012. Bayesian and frequentist inference in partially identified models. *Econometrica* 80:755–82

Moro A. 2003. The effect of statistical discrimination on black-white wage inequality: estimating a model with multiple equilibria. *Int. Econ. Rev.* 44:457–500

Myerson R. 1991. *Game Theory: The Analysis of Conflict*. Cambridge, MA: Harvard Univ. Press

Pakes A, Ostrovsky M, Berry S. 2007. Simple estimators for the parameters of dynamic games, with entry/exit examples. *Rand J. Econ.* 38:373–99

Pakes A, Porter J, Ho K, Ishii J. 2011. *Moment inequalities and their applications*. Work. Pap., Harvard Univ., Cambridge, MA

Pesendorfer M, Schmidt-Dengler P. 2008. Asymptotic least square estimator for dynamic games. *Rev. Econ. Stud.* 75:901–28

Romano JP, Shaikh AM, Wolf M. 2010. Hypothesis testing in econometrics. *Annu. Rev. Econ.* 2:75–104

Romano JP, Wolf M. 2005. Stepwise multiple testing as formalized data snooping. *Econometrica* 73:1237–82

Shang Q, Lee LF. 2011. Two-step estimation of endogenous and exogenous group effects. *Econom. Rev.* 30:173–207

Song K. 2012. *Econometric inference on a large Bayesian game*. Work. Pap., Univ. British Columbia, Vancouver

Sweeting A. 2009. The strategic timing of radio commercials: an empirical analysis using multiple equilibria. *Rand J. Econ.* 40:710–42

Tamer E. 2003. Incomplete simultaneous discrete response model with multiple equilibria. *Rev. Econ. Stud.* 70:147–67

Tamer E. 2010. Partial identification in econometrics. *Annu. Rev. Econ.* 2:167–95

Todd P, Wolpin K. 2012. *Estimating a coordination game within the classroom*. Work. Pap., Univ. Penn., Philadelphia

Wan Y, Xu H. 2010. *Semiparametric estimation of binary decision games of incomplete information with correlated private signals*. Work. Pap., Penn. State Univ., University Park

Annual Review of Economics

Volume 5, 2013

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://econ.annualreviews.org