The Identification Power of Equilibrium in Simple Games

Andres Aradillas-Lopez
Department of Economics, Princeton University, Princeton, NJ 08544

Elie Tamer
Department of Economics, Northwestern University, Evanston, IL 60208 (*tamer@northwestern.edu*)

We examine the identification power that (Nash) equilibrium assumptions play in conducting inference about parameters in some simple games. We focus on three static games in which we drop the Nash equilibrium assumption and instead use rationalizability as the basis for strategic play. The first example examines a bivariate discrete game with complete information of the kind studied in entry models. The second example considers the incomplete-information version of the discrete bivariate game. Finally, the third example considers a first-price auction with independent private values. In each example, we study the inferential question of what can be learned about the parameter of interest using a random sample of observations, under level-k rationality, where k is an integer ≥ 1. As k increases, our identified set shrinks, limiting to the identified set under full rationality or rationalizability (as $k \to \infty$). This is related to the concepts of iterated dominance and higher-order beliefs, which are incorporated into the econometric analysis in our framework. We are then able to categorize what can be learned about the parameters in a model under various maintained levels of rationality, highlighting the roles of different assumptions. We provide constructive identification results that lead naturally to consistent estimators.

KEY WORDS: Equilibrium vs rationality; Identification; Partial identification.

1. INTRODUCTION

In this article we examine the identification power of equilibrium in some simple games. In particular, we relax the assumption of Nash equilibrium (NE) behavior and assume that players are rational. Rationality posits that agents play strategies that are consistent with a set of proper beliefs. The object of interest in these games is a parameter vector that parameterizes payoff functions. We study the identified features of the model using a random sample of data under a set of rationality assumptions, culminating with rationalizability, a concept introduced jointly in the literature by Bernheim (1984) and Pearce (1984), and compare those to what we can learn under Nash. We find that in static discrete games with *complete information*, the identified features of the games with more than one level of rationality are similar to those obtained with Nash behavior assumption but allowing for multiple equilibria (including equilibria in mixed strategies). In a bivariate game with *incomplete information*, if the game has a unique (Bayesian) NE, then there is convergence between the identified features with and without equilibrium only when the level of rationality tends to infinity. When there are multiple equilibria, the identified features of the game under rationality and equilibrium are different: smaller identified sets (hence more information about the parameter of interest) when equilibrium is imposed, but computationally easier to construct identification regions under rationality (i.e., no need to solve for fixed points). In the auction game that we study, the situation is different. We follow the work of Battigalli and Siniscalchi (2003) where, under some assumptions given the valuations, rationalizability predicts only upper bounds on the bids. We show how these bounds can be used to learn about learn about the latent distribution of valuation. Another strategic assumptions in auctions resulting in tighter bounds is the concept of $\mathcal{P}$-dominance studied by Dekel and Wolinsky (2003).

Economists have observed that equilibrium play in noncooperative strategic environment is not necessary for rational behavior. Some can easily construct games in which NE strategy profiles are unreasonable, whereas others can find reasonable strategy profiles that are not Nash. Restrictions once Nash behavior is dropped typically are based on a set of "rationality" criteria, as has been enumerated in numerous works under different strategic scenarios. In this article we study the effect of adopting a particular rationality criterion on learning about parameters of interests. We do not advocate one type of strategic assumption over another, but simply explore one alternative to Nash and evaluate its effect on parameter inference. Thus, depending on the application, identification of parameters of interest certainly can be studied under strategic assumptions other than rationalizability. We provide such an example.

Because every Nash profile is rational under our definition, dropping equilibrium play complicates the identification problem, because under rationality only, the set of predictions is enlarged. As Pearce noted, "this indeterminacy is an accurate reflection of the difficult situation faced by players in a game, because logical guidelines and the rules of the game are not sufficient for uniqueness of predicted behavior." Thus it is interesting from the econometric perspective to examine how the identified features of a particular game changes as weaker assumptions on behavior are made.

We maintain that players in the game are rational, where heuristically, we define rationality as behavior consistent with an optimizing agent equipped with a proper set of beliefs or probability distributions about the unknown actions of others. Rationality comes in different levels or orders, where a profile is first-order rational if it is a best response to some profile for the other players. This intersection of layers of rationality constitutes rationalizable strategies. We study the identification question for level-k rationality for $k \geq 1$. When we study the identifying power of a game under a certain set of assumptions on the

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000105

strategic environment, we implicitly assume that all players in that game are abiding exactly by these assumptions and playing exactly that game. This is important, because theoretical work has challenged the multiplicity issues that arise under rationalizability. For example, Weinstein and Yildiz (2007) showed that for any rationalizable set of strategies in a given game, there is a local disturbance of that game in which these are the unique rationalizable strategies. This ambiguity about what is the exact game being played is why it is important to study the identified features of a model in the presence of multiplicity.

Using equilibrium as a restriction to gain identifying power is a well-known strategy in economics. The model of demand and supply uses equilibrium to equate the quantity demanded with quantity supplied, thus obtaining the classic simultaneous equation model. Other literature in econometrics, such as job search models and hedonic equilibrium models, explicitly use equilibrium as a "moment condition." In this article we study the identification question in simple game-theoretic models without the assumption of equilibrium by focusing on the weaker concept of rationality, k-level rationality and its limit rationalizability. This approach has two important advantages. First, it leads naturally to a well-defined concept of *levels of rationality*, which is attractive practically. Second, it can be adapted to a very wide class of models without the need to introduce ad hoc assumptions. Ultimately, interim rationalizability allows us to do inference (to varying degrees) both on the structural parameters of a model (e.g., the payoff parameters in a reduced-form game or the distribution of valuations in an auction) and on the properties of *higher-order beliefs* by the agents, which are incorporated into the econometric analysis. The features of this hierarchy of beliefs characterize what we call the *rationality level* of agents. In addition, it is possible to also provide testable restrictions that can be used to find an upper bound on the rationality level in a given data set.

Level-k thinking as an alternative to Nash equilibrium behavior also has been studied by Stahl and Wilson (1995), Nagel (1995), Ho, Camerer, and Weigelt (1998), Costa-Gomes, Crawford, and Broseta (2001), Costa-Gomes and Crawford (2006), and Crawford and Iriberri (2007). These models depart from equilibrium behavior by dropping the assumption that each player has a perfect model of others' decisions and replacing it with the assumption that such subjective models survive k rounds of iterated elimination of dominated decisions. Thus each player's subjective model about others' behavior is consistent with level-k interim rationalizability in the sense of Bernheim (1984). For identification, the aforementioned articles assume the existence of a small number of prespecified types, each of which is associated with a very specific behavior. For example, a particular type of player could perform two mental rounds of deletion of dominated strategies and best response to a uniform distribution over the surviving actions. Using carefully designed experiments, previous researchers sought to explain which type best fits the observed choices. This article differs from the aforementioned works by focusing on bounds for conditional choice probabilities that can be rationalized by beliefs that survive k steps of iterated thinking. We look at the largest possible set of level-k rationalizable beliefs but assume nothing about how players choose their actual (unobserved) beliefs from within this set. In addition, we focus on situations in which the researcher ignores how "rational" players are and in which other primitives of the game also are the object of interest: payoff parameters in discrete games or the distribution of valuations in an auction. In an experimental data set, the last set of objects are entirely under the control of the researcher, and strong parametric assumptions typically are made about behavioral types.

In Section 2 we review and define rational play in a noncooperative strategic game. Here we mainly adapt the definition provided by Pearce. We then examine the identification power of dropping Nash behavior in some commonly studied games in empirical economics. In Section 3 we consider discrete static games of complete information. This type of game is widely used in the empirical literature on (static) entry games with complete information and under NE (see, e.g., Bjorn and Vuong 1985; Bresnahan and Reiss 1991; Berry 1994; Tamer 2003; Andrews, Berry, and Jia 2003; Ciliberto and Tamer 2003; Bajari, Hong, and Ryan 2005). Here we find that in the 2×2 game with level-2 rationality, the outcomes of the game coincide with Nash, and thus econometric restrictions are the same. In Section 4 we consider static games with incomplete information. Empirical frameworks for these games have been studied by Aradillas-Lopez (2005), Aguiregabiria and Mira (2007), Seim (2002), Pakes, Porter, Ho, and Ishii (2005), Berry and Tamer (1996), and others. Characterization of rationalizability in the incomplete information game is closely related to the higher-order belief analysis in the global games literature (see Morris and Shin 2003) and to other recently developed concepts, such as those of Dekel, Fudenberg, and Morris (2007) and Dekel, Fudenberg, and Levine (2004). Here we show that level-k rationality implies restrictions on player beliefs in the 2×2 game that lead to simple restrictions that can be exploited in identification. As k increases, an iterative elimination procedure restricts the size of the allowable beliefs that map into stronger restrictions that can be used for identification. If the game admits a unique equilibrium, then the restrictions of the model converge toward Nash restrictions as the level of rationality k increases. With multiple equilibria, the iterative procedure converges to sets of beliefs that contain both the "large" and "small" equilibria. In particular, studying identification in these settings is simple, because we do not need to solve for fixed points, but simply iterate the beliefs toward the predetermined level of rationality k. In Section 5 we examine a first-price independent auction game, where we follow the work of Battigalli and Sinischalchi (2003). Here for any order k, we are only able to bound the unobserved valuation from above. Finally, in Section 6 we conclude.

2. NASH EQUILIBRIUM AND RATIONALITY

In noncooperative strategic environments, optimizing agents maximize a utility function that depends on what their opponents do. In simultaneous games, agents attempt to predict what their opponents will play, and then play accordingly. Nash behavior posits that players' expectations of what others are doing are mutually consistent, and so a strategy profile is Nash if no player has an incentive to change strategy given what the other agents are playing. This Nash behavior makes an implicit assumption on players' expectations. But, players "are not compelled by deductive logic" (Bernheim) to play Nash. In this article we examine the effect of assuming Nash behavior on identification by comparing restrictions under Nash with those obtained under rationality in the sense of Bernheim and Pearce. Here we follow Pearce's framework and first maintain the following assumptions on behavior:

- Players use proper subjective probability distribution, or use the axioms of Savage, when analyzing uncertain events.
- Players are expected utility maximizers.
- Rules and structure of the game are common knowledge.

We next describe heuristically what is meant by rationalizable strategies; precise definitions have been given by Pearce (1984), for example:

- We say that a strategy profile for player i (which can be a mixed strategy) is *dominated* if there exists another strategy for that player that does better no matter what other agents are playing.
- Given a profile of strategies for all players, a strategy for player i is a *best response* if that strategy does better for that player than any other strategy given that profile.

To define rationality, we use the following notation. Let $\mathcal{R}^i(0)$ be the set of all (possibly mixed) strategies that player i can play and $\mathcal{R}^{-i}(0)$ be the set of all strategies for players other than i. Then, heuristically, we have the following:

- *Level-1 rational strategies* for player i are strategy profiles $s^i \in \mathcal{R}^i(0)$ such that there exists a strategy profile for other players in $\mathcal{R}^{-i}(0)$ for which s^i is a best response. The set of level-1 strategies for player i is $\mathcal{R}^i(1)$.
- *Level-2 rational strategies* for player i are strategy profiles $s^i \in \mathcal{R}^i(0)$ such that there exists a strategy profile for other players in $\mathcal{R}^{-i}(1)$ for which s^i is a best response.
- *Level-t rational strategies* are defined recursively from level 1.

Note that by construction, $\mathcal{R}^i(t) \subseteq \mathcal{R}^i(t-1)$. Finally, *rationalizable strategies* are ones that lie in the intersection of the $\mathcal{R}$'s as t increases to infinity. In the complete information game of Section 3, we show that there exists a finite k such that for $\mathcal{R}^i(t) = \mathcal{R}^i(k)$ for all $t \geq k$. In the incomplete information models of Sections 4 and 5, we show that we can have $\mathcal{R}^i(t) \subset \mathcal{R}^i(k)$ for all $t > k$. In all of these settings, a strategy is level-k rational for a player if it is a best response to some strategy profile in $\mathcal{R}^i(k-1)$ by his opponents. Iterating this further, we arrive at the set of rationalizable strategies. Pearce provided properties of the rationalizable set; for example, NE profiles are always included in this set, and the set contains at least one profile in pure strategies.

3. BIVARIATE DISCRETE GAME WITH COMPLETE INFORMATION

Consider the following bivariate discrete 0/1 game where t_p is the payoff that player p obtains by playing 1 when player $-p$ is playing 0. Parameters α_1 and α_2 are of interest. The econometrician does not observe t_1 or t_2 and is interested in learning about the α's and the joint distribution of (t_1, t_2). (See Table 1.) Assume also, as in entry games, that the α's are negative. In this example and the next, we assume that we have access to a random sample of observations $(y_{1i}, y_{2i})_{i=1}^N$, which represent, for example, market structures in a set of N independent markets. To learn about the parameters, we map the observed distribution of the data (the choice probabilities) to the distribution predicted by the model. Because this is a game of complete information, players observe all of the payoff-relevant information. In particular, in the first round of rationality, player 1 will play 1 if $t_1 + \alpha_1 \geq 0$, because this will be a dominant strategy. In addition, if t_1 is negative, then player 1 will play 0. But when $t_1 + \alpha_1 \leq 0 \leq t_1$, both actions 1 and 0 are level-1 rational; action 1 is rational because it can be a best response to player 2 playing 0, whereas action 0 is a best response to player 2 playing 1. The set $\mathcal{R}(1)$ is summarized in Figure 1. Consider, for example the upper right corner. For values of t_1 and t_2 lying there, playing 0 is not a best response for either player. Thus $(1, 1)$ is the unique level-1 rationalizable strategy (which is also the unique NE). Consider now the middle region on the right side, that is, $(t_1, t_2) \in [-\alpha_1, \infty) \times [0, -\alpha_2]$. In level-1 rationality, 0 is not a best reply for player 1, but player 2 can play either 1 or 0; 1 is a best reply when player 1 plays 0, and 0 is a best reply for player 2 when player 1 plays 1. However, in the next round of rational play, given that player 2 now believes that player 1 will play 1 with probability 1, then player 2's response is to play 0. Thus $\mathcal{R}(1) = \{\{1\}, \{0, 1\}\}$ while the rationalizable set reduces to the outcome $(1, 0)$. Here $\mathcal{R}(k) = \mathcal{R}(2) = \{\{1\}, \{0\}\}$ for all $k \geq 2$. In the middle square, we see that the game provides no observable restrictions; any outcome can be *potentially observable*, because both strategies are rational at any level of rationality. Note also that in this game, the set of rationalizable strategies is the set of profiles that are undominated. This is a property of bivariate binary games.

Table 1. Bivariate discrete game

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	$0, 0$	$0, t_2$
$a_1 = 1$	$t_1, 0$	$t_1 + \alpha_1, t_2 + \alpha_2$

3.1 Implications of Level-k Rationality

A random sample of observations allows us to obtain a consistent estimator of the choice probabilities (or the data). The object of interest here is $\theta = (\alpha_1, \alpha_2, F(\cdot, \cdot))$, where $F(\cdot, \cdot)$ is the joint distribution of (t_1, t_2). One interesting approach to conduct inference on the identified set, Θ_I, is to assume that both t_1 and t_2 are discrete random variables with identical support on $s_1, \ldots, s_K$ such that $P(t_1 = s_i; t_2 = s_j) = p_{ij} \geq 0$ for $i, j \in \{1, \ldots, k\}$ with $\sum_{i,j} p_{ij} = 1$. Thus we make inference on the set of probabilities $(p_{ij}, i, j \leq k)$ and (α_1, α_2). We highlight this for level-2 rationality. In particular, we say that

$$\theta = ((p_{ij}), \alpha_1, \alpha_2) \in \Theta_I$$

if and only if

$$P_{11} = \sum_{i,j} p_{ij}\big(1[s_i \geq -\alpha_1; s_j \geq -\alpha_2] + l_{ij}^{(1,1)} 1[0 \leq s_i \leq -\alpha_1; 0 \leq s_j \leq -\alpha_2]\big),$$

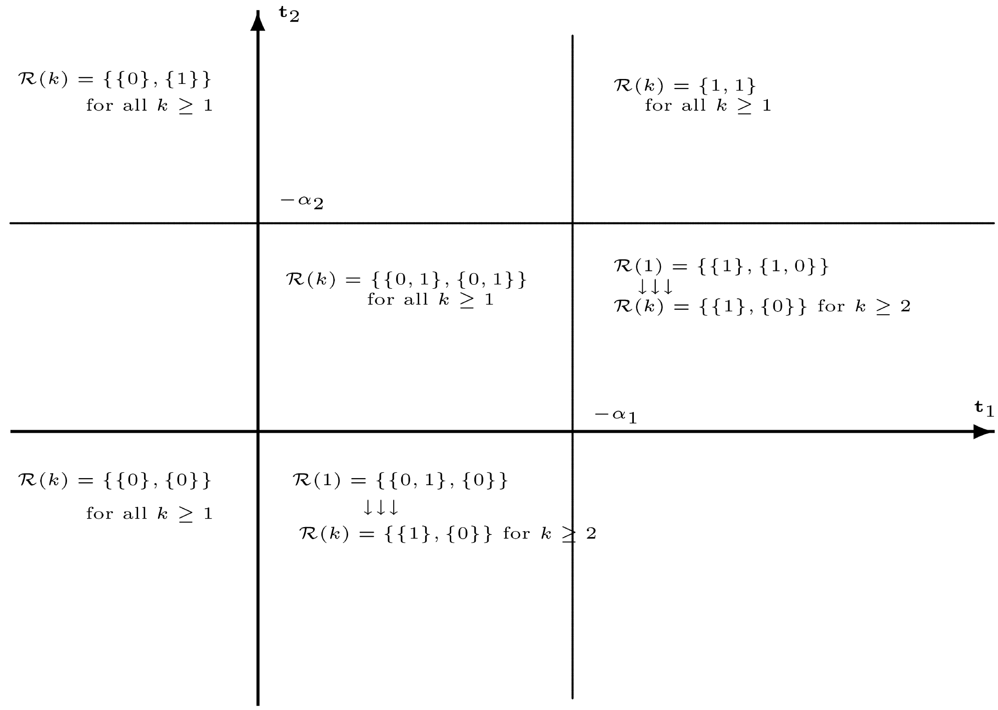

Figure 1. Rationalizable profiles in a bivariate game with complete information.

$$P_{00} = \sum_{i,j} p_{ij}\big(1[s_i \le 0; s_j \le 0] + l_{ij}^{(0,0)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big),$$

$$P_{10} = \sum_{i,j} p_{ij}\big(1[s_i \ge 0; s_j \le 0] + 1[s_i \ge -\alpha_1; 0 \le s_j \le -\alpha_2] + l_{ij}^{(1,0)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big),$$

and

$$P_{01} = \sum_{i,j} p_{ij}\big(1[s_i \le 0; s_j \ge 0] + 1[0 \le s_i \le -\alpha_1; s_j \ge -\alpha_2] + l_{ij}^{(0,1)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big)$$

for some $(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) \ge 0$ and $l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1$ for all $i, j \le k$. The l's can be thought of as the "selection mechanisms" that choose an outcome in the region where the model predicts multiple outcomes. We treat the support points as known, but this is without loss of generality, because those also can be made part of θ. The foregoing equalities (and inequalities) for a given θ are similar to first-order conditions from a linear programming problem and thus can be solved rapidly using linear programming algorithms. In particular, consider the objective function in (1). Note first that $Q(\theta) \le 0$ for all θ's in the parameter space. Moreover,

$$\theta \in \Theta_I$$

if and only if $Q(\theta) = 0$.

$$\begin{aligned}
Q(\theta) = &\max_{v_i,\dots,v_8,(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)})} -(v_1 + \cdots + v_8) \quad \text{s.t.} \\
&P_{11} - \sum_{i,j} p_{ij}\big(1[s_i \ge -\alpha_1; s_j \ge -\alpha_2] + l_{ij}^{(1,1)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big) = v_1 - v_2, \\
&P_{00} - \sum_{i,j} p_{ij}\big(1[s_i \le 0; s_j \le 0] + l_{ij}^{(0,0)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big) = v_3 - v_4, \\
&P_{10} - \sum_{i,j} p_{ij}\big(1[s_i \ge 0; s_j \le 0] + 1[s_i \ge -\alpha_1; 0 \le s_j \le -\alpha_2] + l_{ij}^{(1,0)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big) = v_5 - v_6, \\
&P_{01} - \sum_{i,j} p_{ij}\big(1[s_i \le 0; s_j \ge 0] + 1[0 \le s_i \le -\alpha_1; s_j \ge -\alpha_2] + l_{ij}^{(0,1)} 1[0 \le s_i \le -\alpha_1; 0 \le s_j \le -\alpha_2]\big) = v_7 - v_8, \\
&v_i \ge 0; \qquad \big(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}\big) \ge 0; \\
&l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1 \quad \text{for all } 1 \le i, j \le k.
\end{aligned} \tag{1}$$

First, note that for any θ, the program is feasible; for example, set $(l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) = 0$ and then set $v_1 = P_{11} - \sum_{i,j} p_{ij} 1[s_i \ge -\alpha_1; s_j \ge -\alpha_2]$ and $v_2 = 0$ if $P_{11} - \sum_{i,j} p_{ij} 1[s_i \ge -\alpha_1; s_j \ge -\alpha_2] \ge 0$, or set $v_2 = -(P_{11} - \sum_{i,j} p_{ij} 1[s_i \ge -\alpha_1; s_j \ge -\alpha_2])$ and $v_1 = 0$ and similarly for the rest. Moreover,

$\theta \in \Theta_I$ if and only if $Q(\theta) = 0$. We can collect all of the parameter values for which the foregoing objective function is equal to 0 (or approximately equal to 0). A similar linear programming procedure was used by Honoré and Tamer (2006). The sampling variation comes from having to replace the choice probabilities $(P_{11}, P_{12}, P_{21}, P_{22})$ with their sample analogs, which results in a sample objective function $Q_n(\cdot)$ that can be used to conduct inference.

More generally, and without making support assumptions, a practical way to conduct inference with one level of rationality, say, is to use an implication of the model. In particular, under $k = 1$ rationality, the statistical structure of the model is one of moment inequalities,

$$\Pr(t_1 \geq -\alpha_1; t_2 \geq -\alpha_2) \leq P(1,1) \leq \Pr(t_1 \geq 0; t_2 \geq 0),$$
$$\Pr(t_1 \leq 0; t_2 \leq 0) \leq P(0,0) \leq \Pr(t_1 \leq -\alpha_1; t_2 \leq \alpha_2),$$
$$\Pr(t_1 \geq -\alpha_1; t_2 \leq 0) \leq P(1,0) \leq \Pr(t_1 \geq 0; t_2 \leq -\alpha_2),$$
$$\Pr(t_1 \leq 0; t_2 \geq -\alpha_2) \leq P(0,1) \leq \Pr(t_1 \leq -\alpha_1; t_2 \geq 0).$$

The foregoing inequalities do not exploit all of the information, and thus the identified set based on these inequalities is not sharp. But these inequality-based moment conditions are simple to use and can be generalized to large games. Heuristically, then, by definition the model identifies the set of parameters Θ_I such that the above inequalities are satisfied. Moreover, we say that the model *point identifies* a unique θ if the set Θ_I is a singleton. Figure 2 shows the mapping between the predictions of the game and the observed data under Nash and level-k rationality. The observable implication of Nash is different depending on whether or not we allow for mixed strategies. In particular, without allowing for mixed strategies, in the middle square of Figure 2(a), the only observable implication is $(1, 0)$ and $(0, 1)$; however, it reverts to all outcomes once the mixed strategy equilibrium is considered. To get an idea of the identification gains when we assume rationality versus equilibrium, we simulated a stylized version of the foregoing game in the case where t_p is standard normal for $p = 1, 2$ and the only object of interest is the vector (α_1, α_2). We compare the identified set of the foregoing game under $k = 1$ rationality and NE when we consider only pure strategies. Figure 3 shows that there is identifying power in assuming Nash equilibrium. In particular, under Nash, the identified set is a somehow tight "circle" around the simulated truth, whereas under rationality, the model provides only upper bounds on the alphas. But if we add exogenous variations in the profits (X's), then the identified region under rationality will shrink. In the next section we examine the identifying power of the same game under incomplete information.

4. DISCRETE GAME WITH INCOMPLETE INFORMATION

Consider now the discrete game presented in Table 1 but under the assumptions that player 1 (2) does not observe t_2 (t_1) or that the signals are private information. We denote player $p \in \{1, 2\}$'s opponent by $-p$. We let $\mathcal{I}_p$ denote the signals used by player p to obtain information about t_{-p}, where $t_p \in \mathcal{I}_p$ could be a special case. Player p holds beliefs about his opponent's type conditional on $\mathcal{I}_p$, and those beliefs can be summarized by a subjective distribution function. Let $\pi_2(\mathcal{I}_1)$ denote player 1's *subjective* probability of entry for player 2, and define $\pi_1(\mathcal{I}_2)$ analogously for player 2. Given his beliefs, the expected utility function of player 1 is

$$U(a_1, t_1, \mathcal{I}_1) = \begin{cases} t_1 + \alpha_1 \pi_2(\mathcal{I}_1) & \text{if } a_1 = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, for player 2, we have

$$U(a_2, t_2, \mathcal{I}_2) = \begin{cases} t_2 + \alpha_1 \pi_1(\mathcal{I}_2) & \text{if } a_2 = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Both players are assumed to be expected-utility maximizers who make choices simultaneously and independently. (This includes NE behavior as a special case.) This yields threshold-crossing decision rules

$$Y_1 = \mathbb{1}\{U(1, t_1, \mathcal{I}_1) \geq 0\}, \qquad Y_2 = \mathbb{1}\{U(1, t_2, \mathcal{I}_2) \geq 0\}. \quad (2)$$

Incomplete information makes it impossible for player p to randomize in a way that makes his opponent exactly indifferent between his two actions. In addition, because we focus on the case where t_p is continuously distributed, the event $U(1, t_p, \mathcal{I}_p) = 0$ occurs with probability 0. Our assumptions differ from NE be-

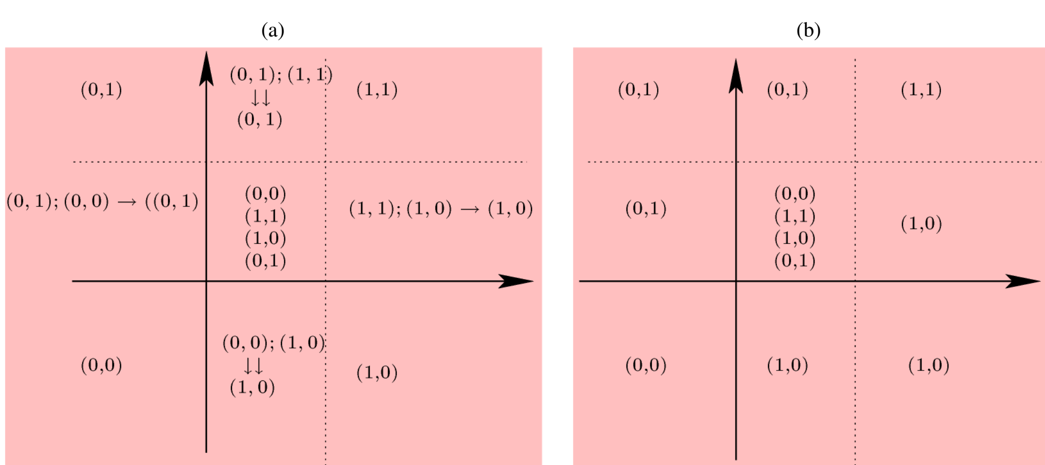

Figure 2. Observable implications of equilibrium (b) versus rationality (a).

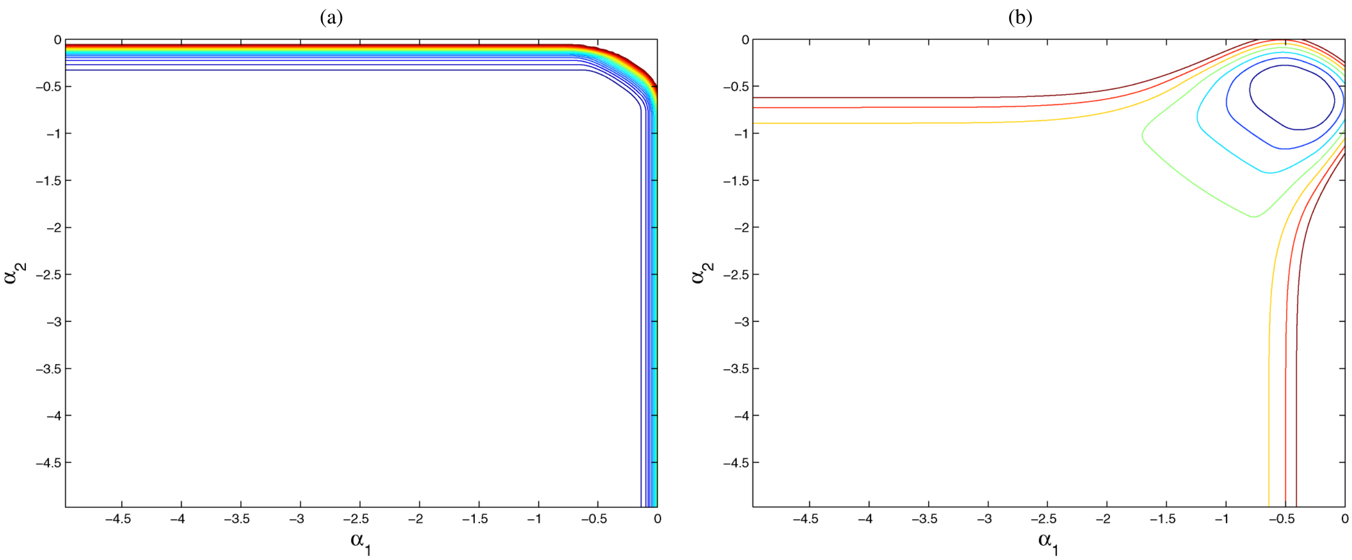

Figure 3. Identification set under Nash and 1-level rationality. Shown the identified regions for (α_1, α_2) under $k = 1$ rationality (a) and Nash (b). We set in the underlying model $(\alpha_1, \alpha_2) = (-.5, -.5)$. The model was simulated assuming Nash with (0, 1) selected with probability one in regions of multiplicity. Note that in (a), the model only places upper bounds on the alphas, whereas in (b) (α_1, α_2) are constrained to lie a much smaller set (the inner "circle").

cause we do not impose the restriction that subjective beliefs are consistent with players' actual behavior. Again, here we assume that both α_1 and α_2 are negative.

4.1 Implications of Level-1 Rationality

We maintain the expected utility maximization assumption and the resulting decision rules (2). In the first round of rationality, we know that for any belief function, or without making any common prior assumptions, the following hold:

$$\begin{aligned} t_1 + \alpha_1 \geq 0 \quad &\Longrightarrow \quad U(1, t_1, \mathcal{I}_1) = t_1 + \alpha_1 \pi_2(\mathcal{I}_1) \geq 0 \\ &\qquad \forall \pi_2(\mathcal{I}_1) \in [0, 1], \\ t_1 \leq 0 \quad &\Longrightarrow \quad U(1, t_1, \mathcal{I}_1) = t_1 + \alpha_1 \pi_2(\mathcal{I}_1) \leq 0 \\ &\qquad \forall \pi_2(\mathcal{I}_1) \in [0, 1]. \end{aligned} \tag{3}$$

Well-defined beliefs satisfy $\pi_2(\cdot) \in [0, 1]$. This implies that if player 1 is an expected-utility maximizer and holds well-defined beliefs, then he must satisfy

$$t_1 + \alpha_1 \geq 0 \quad \Longrightarrow \quad a_1 = 1$$

and

$$t_1 \leq 0 \quad \Longrightarrow \quad a_1 = 0.$$

Now, let $0 \leq t_1 \leq -\alpha_1$. For a player that is rational of order 1, there exists well defined beliefs that rationalizes either 1 or 0. Thus when $0 \leq t_1 \leq -\alpha_1$, both $a_1 = 1$ and $a_1 = 0$ are rationalizable. So, the implication of the game are summarized in Figure 4. Note here that the (t_1, t_2) space is divided into nine regions: four regions where the outcome is unique, four regions with two potentially observable outcomes, and the middle square where any outcome is potentially observed. To make inference based on this model, we need to map these regions into predicted choice probabilities. To obtain the sharp set of parameters that is identified by the model, we can supplement this model with consistent "selection rules" that specify, in regions of multiplicity, the probability of observing the various outcomes, which would be a function of both $(t_1, \mathcal{I}_1)$ and $(t_2, \mathcal{I}_2)$. The probabilities can be given a "structural" interpretation in which they would be interpreted as proper selection mechanisms. Given the level-1 behavioral assumptions, the only valid selection mechanisms are those that can be produced (rationalized) by the choice rules (2) for some well-defined beliefs. Expected utility maximization explains [through (2)] how players' choices are produced in an incomplete information environment given beliefs. Finally, let the joint distribution of (t_1, t_2) be noted by $F(\cdot)$.

Result 1. For the game with incomplete information, let the players be rational with order 1 (level-1 rational) and write $W_p \equiv t_p \cup \mathcal{I}_p$. Then the choice probabilities predicted by the model are

$$\begin{aligned} P(1,1) &= \int_{III} dF + \int_{II} S^{II}_{(1,1)}(W_1, W_2)\, dF \\ &\quad + \int_{VI} S^{VI}_{(1,1)}(W_1, W_2)\, dF + \int_{V} S^{V}_{(1,1)}(W_1, W_2)\, dF, \\ P(0,0) &= \int_{VII} dF + \int_{VIII} S^{VIII}_{(0,0)}(W_1, W_2)\, dF \\ &\quad + \int_{IV} S^{IV}_{(0,0)}(W_1, W_2)\, dF + \int_{V} S^{V}_{(0,0)}(W_1, W_2)\, dF, \\ P(0,1) &= \int_{I} dF + \int_{II} S^{II}_{(0,1)}(W_1, W_2)\, dF \\ &\quad + \int_{IV} S^{VI}_{(0,1)}(W_1, W_2)\, dF + \int_{V} S^{V}_{(0,1)}(W_1, W_2)\, dF, \end{aligned} \tag{4}$$

where $S^i_j \geq 0$ are such that for example, $S^{II}_{(1,0)} + S^{II}_{(1,1)} = 1$, and so on, and I, II, III, IV, V, VI, and VIII are regions for (t_1, t_2) shown in Figure 4.

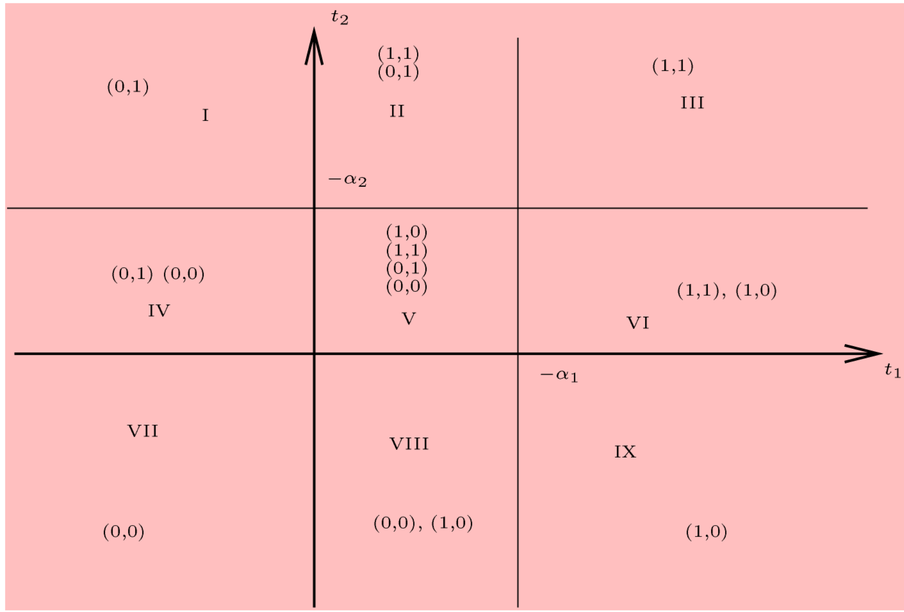

Figure 4. Observable implications of level-1 rationality.

The functions S are unknown and represent "selection" functions that represent the probabilities of selecting a particular outcome in a region of multiplicity. Suppose, for simplicity, that $\mathcal{I}_p = t_p$, so that players condition their beliefs only on the realization of their own type. The (sharp) identified set Θ_I, is the set of parameters where there exists proper selection functions S such that the predicted choice probabilities in (4) are equal to the ones obtained from data. The restrictions in (4) can be exploited by, for example, discretizing the joint distribution of (t_1, t_2), such as discussed for the complete-information case, to construct the identified set. The latter, Θ_I, is the set of parameters for which the equalities in (4) are satisfied for well-defined selection functions. Implications of the foregoing set of equalities is a set of moment *inequalities* constructed by exploiting the fact that the S functions are probabilities and thus are positive. So, for example, an implication of Result 1 is that $\int_{III} dF \leq P(1,1) \leq \int_{III} dF + \int_{II \cup III \cup V \cup VI} dF$, where the bounds of this inequalities do not involve the unknown functions S.

Next we analyze the behavior of players who assume that their opponents are (at least) level-k rational for $k \geq 1$. Level-2 rational players are those whose second-order beliefs for their opponent are compatible with the bounds implied by (3). As we show, by eliminating beliefs that violate (3), we are able to reduce the set of level-2 rational beliefs from the entire [0, 1] interval to a segment of it. Further rounds of iterated thinking will refine these bounds even more. Unlike in the Bayesian Nash equilibrium (BNE) case, we do not impose the requirement that beliefs are correct; we will rule only out those that are not compatible with the assumption that opponents are level-k rational.

4.2 Implications of Level-*k* Rationality

Level-1 rationality is characterized simply by expected utility maximization and any arbitrary system of well-defined beliefs. We now generalize the results of the previous section by characterizing bounds for beliefs that are compatible with assuming that opponents are level-k rational. This means that, for example, level-2 rational players are all of those whose beliefs are consistent with the bounds implied by (3). As we show later, by eliminating beliefs that violate (3), we will be able to reduce the set of level-2 rational beliefs from the entire [0, 1] interval to a segment of it. Level-3 rational players are those whose beliefs are compatible with the bounds for level-2 rational beliefs. This iterative construction can then be used to characterize bounds for level-k rational beliefs. Each "round of rationality" refines these bounds by deleting all beliefs that assign positive probability to opponents' dominated strategies. As a reminder, the realization of t_p is privately observed by player p, who conditions his beliefs about the expected action of his opponent on the realization of signals $\mathcal{I}_p$, with $t_p \in \mathcal{I}_p$ being a special case. The true distribution of $(t_1 \cup t_2 \cup \mathcal{I}_1 \cup \mathcal{I}_2)$ is common knowledge to both players. This is the common prior assumption. Even though it plays no role in the analysis of level-1 rational behavior, the common prior assumption is important for higher levels of rationality. We consider strategies (decision rules) for player p that are threshold functions of t_p,

$$Y_p = \mathbb{1}\{t_p \geq \mu_p\} \quad \text{for } p = 1, 2. \tag{5}$$

It follows from the normal-form payoffs in Table 1 that this family of decision rules includes those of all expected utility-maximizing players in this simple binary choice game with incomplete information. Level-1 rational players and those who play a BNE are two special cases. In the construction of his expected utility, player p forms *subjective* beliefs about μ_{-p} that can be summarized by a probability distribution for μ_{-p} given $\mathcal{I}_p$. These beliefs are derived as part of a solution concept. They may include BNE beliefs as a special case (in which case all players know those equilibrium beliefs to be correct). Here let $\widehat{G}_1(\mu_2|\mathcal{I}_1)$ denote player 1's subjective distribution function for μ_2 given $\mathcal{I}_1$, and define $\widehat{G}_2(\mu_1|\mathcal{I}_2)$ analogously for player 2. A strategy by player p is *rationalizable* if it is the best response (in the expected-utility sense) given some beliefs $\widehat{G}_p(\mu_{-p}|\mathcal{I}_p)$

that assign zero probability mass to strictly dominated strategies by player $-p$. A rationalizable strategy by player p is described by

$$Y_p = \mathbb{1}\left\{t_p + \alpha_p \int_{\mathbb{S}(\widehat{G}_p)} E[\mathbb{1}\{t_{-p} \geq \mu\}|\mathcal{I}_p, \mu]\, d\widehat{G}_p(\mu|I_p) \geq 0\right\}, \tag{6}$$

where the support $\mathbb{S}(\widehat{G}_p)$ excludes values of μ that result in dominated strategies within the class (5). Throughout, we focus on the case where μ_{-p} is continuously distributed conditional on $\mathcal{I}_p$, and ignore the distinction between strictly and weakly dominated strategies. Note that the subset of rationalizable strategies within the class (5) is of the form $\mu_p = -\alpha_p \int_{\mathbb{S}(\widehat{G}_p)} E[\mathbb{1}\{t_{-p} \geq \mu\}|\mathcal{I}_p, \mu]\, d\widehat{G}_p(\mu|I_p)$. In this setting, rationalizability requires expected utility maximization for a given set of beliefs but does not require those beliefs to be correct. It only imposes the condition that $\mathbb{S}(\widehat{G}_p)$ exclude values of μ_{-p} that are dominated. We eliminate such values by iterated deletion of dominated strategies.

Now we describe the iterative procedure that restricts $\mathbb{S}(\widehat{G}_p)$ by iterated dominance. As before, we maintain that the signs of the strategic interaction parameters (α_1, α_2) are known. Specifically, suppose that $\alpha_p \leq 0$. Then, repeating arguments from the previous section on $k = 1$–rationalizable outcomes, looking at (6), we see that we must have eventwise comparisons

$$\mathbb{1}\{t_p + \alpha_p \geq 0\} \leq \mathbb{1}\{Y_p = 1\} \qquad \text{and} \\ \mathbb{1}\{t_p < 0\} \leq \mathbb{1}\{Y_p = 0\}. \tag{7}$$

Decision rules that do not satisfy these conditions are strictly dominated for all possible beliefs. Therefore, the subset of strategies within the class (5) that are not strictly dominated must satisfy $\Pr(t_p + \alpha_p \geq 0) \leq \Pr(t_p \geq \mu_p) \leq \Pr(t_p \geq 0)$ or, equivalently, $\mu_p \in [0, -\alpha_p]$. All other values of μ_p correspond to dominated strategies. In this setup, we refer that to the subset of strategies that satisfy $\mu_p \in [0, -\alpha_p]$ as *level-1 rationalizable strategies*. Note that, as before, these μ's do not involve the common prior distributions.

Level-2 rational players are those whose beliefs are consistent with assuming that their opponents are level-1 rational. Without any further assumptions, level-2 rational players are those whose beliefs about others satisfy (7). Consequently, a level-2 rational player must have beliefs that assign zero probability mass to values $\mu_{-p} \notin [0, -\alpha_p]$. As before, we impose no further requirements (such as having unbiased beliefs). A strategy is level-2 rationalizable if it can be justified by level-2 rationalizable beliefs, that is,

$$\mu_p = -\alpha_p \int_0^{-\alpha_{-p}} E[\mathbb{1}\{t_{-p} \geq \mu\}|\mathcal{I}_p, \mu]\, d\widehat{G}_p(\mu|\mathcal{I}_p),$$

where player p's beliefs $\widehat{G}_p(\cdot|\mathcal{I}_p)$ satisfy $\widehat{G}_p(0|\mathcal{I}_p) = 0$ and $\widehat{G}_p(-\alpha_{-p}|\mathcal{I}_p) = 1$; that is, those beliefs give zero weight to level-1–dominated strategies. Moreover, the expectation within the integral is taken with respect to the common prior conditional on $\mathcal{I}_p$, which includes player p's type. Thus, exploiting this monotonicity, it is easy to see that for an outside observer, the subset of level-2 rationalizable strategies must satisfy

$$\mu_1 \in \left[-\alpha_1 E[\mathbb{1}\{t_2 \geq -\alpha_2\}|\mathcal{I}_1], -\alpha_1 E[\mathbb{1}\{t_2 \geq 0\}|\mathcal{I}_1]\right]$$

and

$$\mu_2 \in \left[-\alpha_2 E[\mathbb{1}\{t_1 \geq -\alpha_1\}|\mathcal{I}_2], -\alpha_2 E[\mathbb{1}\{t_1 \geq 0\}|\mathcal{I}_2]\right].$$

Level-k rational players are those whose beliefs are consistent with assuming that their opponents are level-$(k-1)$ rational. Note that this definition is a statement about a player's higher-order beliefs up to order $k-1$; specifically, any player who believes that his opponent undertakes (at least) $k-1$ rounds of iterated deletion of dominated strategies in the construction of his expected utility will be a level-k rational player. By induction, it is easy to prove the following claim.

Claim 1. If $\alpha_p \leq 0$, then a strategy of the type $Y_p = \mathbb{1}\{t_p \geq \mu_p\}$ is level-k rationalizable if and only if μ_1 and μ_2 satisfy

$$\begin{aligned}
&\mu_p \in [0, -\alpha_{-p}] \equiv [\mu_{p,1}^L, \mu_{p,1}^U], \quad \text{for } k = 1 \text{ and } p \in \{1,2\}; \\
&\mu_1 \in \left[-\alpha_1 E[\mathbb{1}\{t_2 \geq \mu_{2,k-1}^U\}|\mathcal{I}_1], -\alpha_1 E[\mathbb{1}\{t_2 \geq \mu_{2,k-1}^L\}|\mathcal{I}_1]\right] \\
&\qquad \equiv [\mu_{1,k}^L, \mu_{1,k}^U], \quad \text{for } k > 1; \\
&\mu_2 \in \left[-\alpha_2 E[\mathbb{1}\{t_1 \geq \mu_{1,k-1}^U\}|\mathcal{I}_2], -\alpha_2 E[\mathbb{1}\{t_1 \geq \mu_{1,k-1}^L\}|\mathcal{I}_2]\right] \\
&\qquad \equiv [\mu_{2,k}^L, \mu_{2,k}^U], \quad \text{for } k > 1.
\end{aligned} \tag{8}$$

The bounds described in (8) contain any set of beliefs that can be rationalized after $k-1$ rounds of iterated deletion of dominated strategies. We present identification results based on this entire range with no additional restrictions on how level-k players actually choose their beliefs from within this space of rationalizable beliefs.

Remark 1. Any level-k rational player also is level-k' rational for any $1 \leq k' \leq k-1$. Also, for $p \in \{1,2\}$, with probability 1, we have that $[\mu_{p,k}^L, \mu_{p,k}^U] \subseteq [\mu_{p,k-1}^L, \mu_{p,k-1}^U]$ for any $k > 1$, with strict inclusion if $\alpha_p \neq 0$ and t_{-p} has unbounded support conditional on $\mathcal{I}_p$. This monotonic feature of bounds (as k increases) is a consequence of the payoff parameterization in the game. Note also that these bounds are a function of $\mathcal{I}_p$, the information on which player p conditions his beliefs.

The two statements in Remark 1 follow because conditional on $\mathcal{I}_p$, the support $\mathbb{S}(\widehat{G}_p)$ of a k-level rational player is contained in that of a $k-1$–level rational player. In fact, if there is a unique BNE (conditional on $\mathcal{I}_p$), then $\mathbb{S}(\widehat{G}_p)$ will collapse to the singleton given by BNE beliefs as $k \to \infty$. Whenever warranted, we clarify whether a k-level rational player is "at most k-level rational" or "at least k-level rational." For inference based on level-2 rationality, we can use inequalities similar to (4) to map the observed choice probabilities to the predicted ones; in particular, we can use the thresholds from Claim 1 to construct a map between the model and the observable outcomes using (5). This is illustrated in Figure 5 for the case where $\mathcal{I}_p = t_p$ (i.e., players condition their beliefs exclusively on the realization of their own type). There $\mathbb{P}_{t_1}(\cdot)$ denotes the conditional distribution of $t_2|t_1$, with $\mathbb{P}_{t_2}(\cdot)$ defined analogously. We see that, moving from level 1 to level 2, the middle square shrinks. As we show later, higher rationality levels (properly speaking, further rounds of deletion of dominated strategies) will shrink it further. The set of choice probabilities predicted by the model with level-k rational players can be characterized by generalizing Result 1.

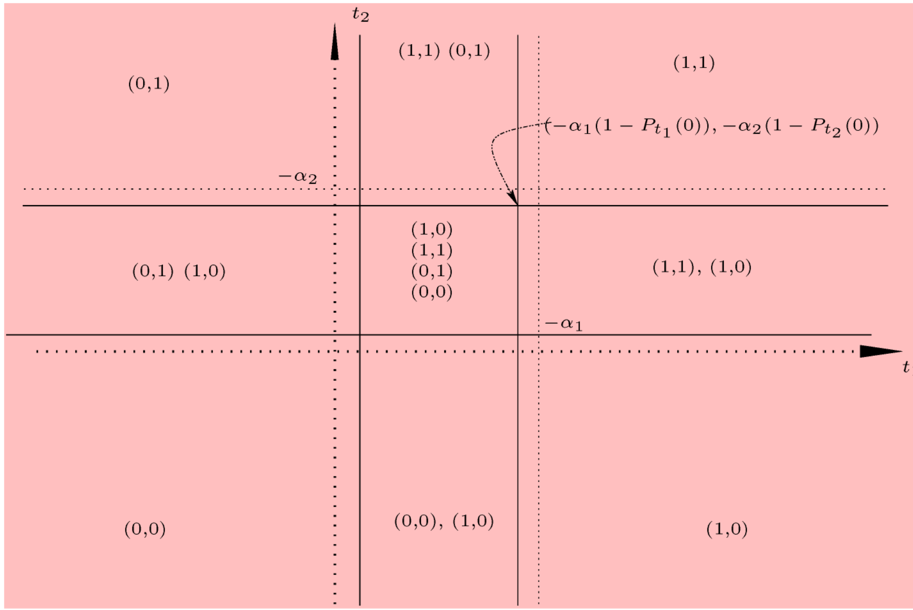

Figure 5. Observable implications of level-2 rationality.

Result 2. Let

$$\pi_p^L(1;\mathcal{I}_p)=0 \qquad \text{and} \qquad \pi_p^U(1;\mathcal{I}_p)=1$$

for $p=1,2$, and let for $k>1$,

$$\pi_1^L(k;\mathcal{I}_1)=E\big[\mathbb{1}\{t_2+\alpha_2\pi_2^U(k-1;\mathcal{I}_2)\geq 0\}|\mathcal{I}_1\big],$$
$$\pi_1^U(k;\mathcal{I}_1)=E\big[\mathbb{1}\{t_2+\alpha_2\pi_2^L(k-1;\mathcal{I}_2)\geq 0\}|\mathcal{I}_1\big],$$
$$\pi_2^L(k;\mathcal{I}_2)=E\big[\mathbb{1}\{t_1+\alpha_1\pi_1^U(k-1;\mathcal{I}_1)\geq 0\}|\mathcal{I}_2\big],$$
$$\pi_2^U(k;\mathcal{I}_2)=E\big[\mathbb{1}\{t_1+\alpha_1\pi_1^L(k-1;\mathcal{I}_1)\geq 0\}|\mathcal{I}_2\big].$$

Using the notation in Section 2, the space of strategies for level-k rational players is

$$\mathcal{R}^p(k)=\big\{Y_p=\mathbb{1}\{t_p+\alpha_p\pi_{-p}(\mathcal{I}_p)\geq 0\}: \pi_{-p}\in[\pi_{-p}^L(k;\cdot),\pi_{-p}^U(k;\cdot)]\big\} \quad \text{for } p=1,2.$$

In the next section we parameterize the types t_p to allow for observable heterogeneity and provide sufficient point identification conditions based exclusively on the restrictions implied by level-k rationality.

4.3 Identification With Level-k Rationality in a Parametric Model

From here on, we express $t_p=X_p'\beta_p-\varepsilon_p$, where X_p is observable to the econometrician, ε_p is not, and β_p must be estimated (along with α_p, the strategic interaction parameter for player p). Player p observes the realization of his own X_p and ε_p, where the latter is *only* privately observed. We also allow the possibility that some elements in X_p are private information to player p and, as before, denote the vector of signals used by player p to condition his beliefs by $\mathcal{I}_p$. Throughout, we assume $(\varepsilon_1,\varepsilon_2)$ to be continuously distributed, with scale normalized to 1 and unbounded support. The results that follow only require that for each player p, the support of ε_p be larger than that of $X_p'\beta_p$ for all possible realizations of $\mathcal{I}_{-p}$. For simplicity, we assume that ε_1 is independent of ε_2 and denote their cumulative distribution function (cdf) as $H_p(\cdot)$ for $p=1,2$. Conceptually, we can extend the result that follow and obtain constructive identification results for the case where ε_1 and ε_2 are correlated, but we do not deal with that case here. For simplicity, we limit ourselves to the case where beliefs are conditioned on observables to the researcher; that is, $\mathcal{I}_p$ is observable. We define the identified set of parameters and then provide an objective function that can be used to construct the identified set. This function depends on the level k of rationality that the econometrician assumes ex ante. We discuss the identification of k, then we provide a set of sufficient conditions to guarantee point identification under some assumptions. These point identification results provide insight into the kind of "variation" needed to shrink the identified set to a point. Our results can be extended to cases where beliefs are conditioned on unobservables to the researcher, as long as the joint distribution of all unobservables in the model is assumed known, possibly up to a finite-dimensional parameter.

As in the previous section, we make a common prior assumption. This assumption is only needed to compute bounds on beliefs for levels of rationality k that are strictly larger than 1. Specifically, we assume that H_1 and H_2 are common knowledge among the players, and also assume that the econometrician knows these common prior distributions. We assume that players use the true distributions as their priors for payoff covariates X_p and signals $\mathcal{I}_p$, both of which are observed by the econometrician. Implicitly, we also assume that the true values of β_p and α_p are common knowledge to both players. Given this setup, we can construct bounds on beliefs iteratively. For any parameter value and any "rationality level," these bounds are identified, and they constitute the foundation for our identification results.

Iterated Dominance and Bounds for Beliefs. For ease of exposition, we assume that both players condition on the same vector of signals, which we denote by $\mathcal{I}$. This would include the case where the only source of private information for player p is ε_p and $\mathcal{I} = X_1 \cup X_2$. We will return to the more general case and allow for $\mathcal{I}_1 \neq \mathcal{I}_2$ later. As in Section 4.2, we derive bounds for the range of rationalizable beliefs iteratively by deleting those that assign positive probability to opponents' dominated strategies. For each player p, let

$$\pi^L_{-p}(\theta|k=1,\mathcal{I}) = 0, \qquad \text{and} \qquad \pi^U_{-p}(\theta|k=1,\mathcal{I}) = 1,$$

and for $k \geq 2$, let

$$\begin{aligned}
\pi^L_1(\theta|k,\mathcal{I}) &= E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi^U_2(\theta|k-1,\mathcal{I})\big)|\mathcal{I}\big];\\
\pi^U_1(\theta|k,\mathcal{I}) &= E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi^L_2(\theta|k-1,\mathcal{I})\big)|\mathcal{I}\big];\\
\pi^L_2(\theta|k,\mathcal{I}) &= E\big[H_2\big(X_2'\beta_2 + \alpha_2\pi^U_1(\theta|k-1,\mathcal{I})\big)|\mathcal{I}\big];\\
\pi^U_2(\theta|k,\mathcal{I}) &= E\big[H_2\big(X_2'\beta_2 + \alpha_2\pi^L_1(\theta|k-1,\mathcal{I})\big)|\mathcal{I}\big],
\end{aligned} \tag{9}$$

where $\pi^L_{-p}(\theta|k,\mathcal{I})$ and $\pi^U_{-p}(\theta|k,\mathcal{I})$ are the lower and upper bounds for level-k rationalizable beliefs by player p for $\Pr(Y_{-p}|\mathcal{I})$. Given our foregoing assumptions, these bounds are identified for any θ and k. In the case where we want to allow for correlation in ε_1 and ε_2, the belief function for player p will depend on ε_p, which would be part of a player-specific information set, and H_p would be the conditional cdf of $\varepsilon_p|\varepsilon_{-p}$. By induction, it is easy to show that

$$\begin{aligned}
&[\pi^L_{-p}(\theta|k;\mathcal{I}), \pi^U_{-p}(\theta|k;\mathcal{I})]\\
&\quad\subseteq [\pi^L_{-p}(\theta|k-1;\mathcal{I}), \pi^U_{-p}(\theta|k-1;\mathcal{I})]\\
&\quad\quad \text{with probability 1 in } \mathbb{S}(\mathcal{I}).
\end{aligned} \tag{10}$$

This monotonic feature holds even if players condition on different information sets. Moreover, the inclusion in (10) is strict if the strategic interaction coefficients are nonzero and if ε_p has unbounded support conditional on $X_p'\beta_p$ and $\mathcal{I}$. Figure 6 depicts this case for a fixed realization $\mathcal{I}$, a given parameter vector θ, and $k \in \{2,3,4,5\}$.

Identified Set for θ Based on Level-k Rationality. Let $W_p = X_p \cup \mathcal{I}$. It follows from the discussion in Section 4.2 (see Result 2) that the identified set for θ under the assumption that players are level-k rational is given by

$$\begin{aligned}
\Theta_I(k) = \Big\{\theta \in \Theta : \exists \pi_1(\cdot), \pi_2(\cdot) \in [\pi^L_1(\theta|k;\cdot), \pi^U_1(\theta|k;\cdot)]&\\
\times [\pi^L_2(\theta|k;\cdot), \pi^U_2(\theta|k;\cdot)]&\\
\text{such that } E[Y_p|W_p] = H_p(X_p'\beta_p + \alpha_p\pi_{-p}(\mathcal{I}))&\\
\text{with probability 1 for } p = 1,2\Big\}.&
\end{aligned} \tag{11}$$

We exploit the fact that under our assumptions, the bounds for level-k rational beliefs are identified to characterize a set $\Theta(k)$ that includes $\Theta_I(k)$. Our characterization constructive and based on conditional moment inequalities. To proceed, note that player p is level-k rational if and only if

$$\begin{aligned}
&\mathbb{1}\{X_p'\beta_p + \alpha_p\pi^U_{-p}(\theta|k;\mathcal{I}) \geq \varepsilon_p\}\\
&\quad \leq \mathbb{1}\{Y_p = 1\}\\
&\quad \leq \mathbb{1}\{X_p'\beta_p + \alpha_p\pi^L_{-p}(\theta|k;\mathcal{I}) \geq \varepsilon_p\} \quad \text{with probability 1.}
\end{aligned}$$

Recall that we are studying the case where $\alpha_p \leq 0$ for $p = 1,2$. These inequalities must hold with probability 1 for all realizations of $(X_p, \varepsilon_p, \mathcal{I})$. It follows that level-$k$ rational players must satisfy

$$\begin{aligned}
&H_p\big(X_p'\beta_p + \alpha_p\pi^U_{-p}(\theta|k;\mathcal{I})\big)\\
&\quad \leq E[Y_p|W_p]\\
&\quad \leq H_p\big(X_p'\beta_p + \alpha_p\pi^L_{-p}(\theta|k;\mathcal{I})\big) \quad \text{with probability 1,}
\end{aligned}$$

where $W_p = X_p \cup \mathcal{I}$. Define the set

$$\begin{aligned}
\Theta(k) = \Big\{\theta \in \Theta : H_p\big(X_p'\beta_p + \alpha_p\pi^U_{-p}(\theta|k;\mathcal{I})\big) \leq E[Y_p|W_p]&\\
\leq H_p(X_p'\beta_p + \alpha_p\pi^L_{-p}(\theta|k;\mathcal{I}))&\\
\text{with probability } 1, p = 1,2\Big\}.&
\end{aligned} \tag{12}$$

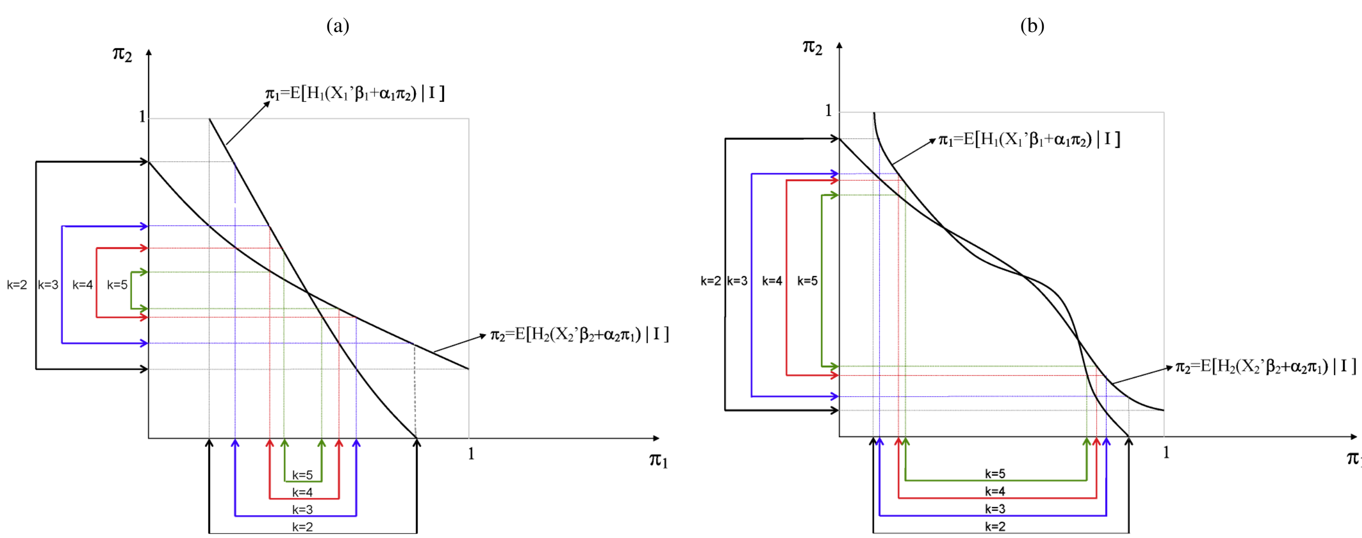

Figure 6. Rationalizable beliefs for $k = 2, 3, 4$, and 5. (a) Belief iterations with a unique BNE. (b) Belief iterations with a multiple BNE. Bounds for level-k rationalizable beliefs when $\mathcal{I}_1 = \mathcal{I}_2 \equiv \mathcal{I}$ (players condition on the same set of signals). The vertical axis shows level-k rationalizable bounds for player 1's beliefs about $\Pr(Y_2 = 1|\mathcal{I})$. The horizontal axis shows the equivalent objects for player 2. The graphs correspond to a particular realization $\mathcal{I}$ and a given parameter value θ.

Clearly, if players are level-k rational, then we have $\Theta_I(k) \subseteq \Theta(k)$. If $X_p \in \mathcal{I}$ for both players, then it is easy to show that $\Theta_I(k) = \Theta(k)$. This follows because the set $[\pi_1^L(\theta|k;\cdot), \pi_1^U(\theta|k;\cdot)] \times [\pi_2^L(\theta|k;\cdot), \pi_2^U(\theta|k;\cdot)]$ is connected and the distributions H_1 and H_2 are continuous. The characterization $\Theta(k)$ is constructive and is the one that we use even though it might be a strict superset of the sharp set $\Theta_2(k)$. This is because dealing with the set $\Theta(k)$ is simple and computationally attractive. To allow for the case where W_p has continuous support, we reexpress $\Theta(k)$ as the set of minimizers of an objective function (see Dominguez and Lobato 2004). For two vectors $a, b \in \mathbb{R}^{\dim(W_p)}$, let

$$\begin{aligned}\Lambda_p(\theta|a,b;k) &= E\big[\big(1 - \mathbb{1}\big\{H_p\big(X_p'\beta_p + \alpha_p\pi_{-p}^U(\theta|k;\mathcal{I})\big) \le \Pr(Y_p = 1|W_p) \\ &\qquad \le H_p\big(X_p'\beta_p + \alpha_p\pi_{-p}^L(\theta|k;\mathcal{I})\big)\big\}\big) \\ &\qquad \times \mathbb{1}\{a \le W_p \le b\}\big];\end{aligned} \tag{13}$$

$$\Gamma_p(\theta|k) = \int\int \Lambda_p(\theta|a,b;k)\,dF_{W_p}(a)\,dF_{W_p}(b);$$

$$\Gamma(\theta|k) = (\Gamma_1(\theta|k), \Gamma_2(\theta|k))',$$

where the inequality $a \le W_p \le b$ is elementwise and $W_p \sim F_{W_p}(\cdot)$. Take any 2×2 positive definite matrix Ω. The set in (12) can be expressed as

$$\Theta(k) = \{\theta \in \Theta : \Gamma(\theta|k)'\Omega\Gamma(\theta|k) = 0\}. \tag{14}$$

This is the definition of identified set for θ that we use under the assumption that all players in the game are level-k rational. More precisely, returning to Remark 1, $\Theta(k)$ is the identified set if we assume that all players in the population are at least level-k rational. By construction, $\Theta(k+1) \subseteq \Theta(k)$ for all k. Methods meant for set inference can be adapted to construct a sample estimator of $\Theta(k)$ based on a random sample of games where all players are level-k rational for a given k. Note also that, compared with the Bayesian Nash solution, here we do not need to solve a fixed-point map to obtain the equilibrium; rather, rationalizability requires restrictions on player beliefs, which can be implemented iteratively. We formally show that $\Theta(k)$ contains the set of BNE for any $k > 0$. Having $\Theta(k) = \emptyset$ would reject the hypothesis that all players are at least level-k rational.

Remark 2. Note that when $k = 1$, we do not need to specify the common prior assumption, because here beliefs play no role. Thus results will be robust to this assumption. However, depending on the magnitude of the α_p's, the bounds on choice probabilities predicted by such a model (where $k = 1$) can be wide.

Under certain conditions, the identified set in (14) would consist only of θ_0, the true parameter value. An example of this is a case in which there exist realizations of the vector of signals $\mathcal{I}$ where the players are "forced" to take one of their actions with probability 1 regardless of their beliefs. To be concrete, suppose that the linear index $X_p'\beta_p$ has unbounded support for both players, and suppose that both of them are at least level-2 rational (i.e., they both perform at least one round of deletion of dominated strategies). Then, if the vector of signals $\mathcal{I}$ is such that there exist regions of $\mathbb{S}(\mathcal{I})$ such that $\mathbb{S}(X_p'\beta_p|\mathcal{I})$ is concentrated around arbitrarily large positive or arbitrarily large negative values, the identified set $\Theta(k=2)$ defined in (14) would collapse to a singleton θ_0, the true parameter value. We refer to this as a case of "informative signals" and formalize this point identification result in the next section.

4.4 Sufficient Point Identification Conditions

In this section we study the problem of point identification of the parameter of interests in the foregoing game. In particular, we provide sufficient point identification conditions for level-1 rational play and for levels $k > 1$. These conditions can provide insight into what is required to shrink the identified set to a point (or a vector). Here we allow for the information sets to be different; that is, player p conditions on $\mathcal{I}_p$ when making decisions and allow for exclusion restrictions where $\mathcal{I}_1 \neq \mathcal{I}_2$. We start with sufficient conditions for level-1 rationalizability.

4.4.1 Identification Results With Level-1 Rationality. Let $\theta_p = (\beta_p, \alpha_p)$ and $\theta = (\theta_1, \theta_2)$; then we have the following identification result.

Theorem 1. Suppose that X_p has full rank for $p = 1, 2$, and let $X \equiv (X_1, X_2)$; assume that $\alpha_p < 0$ for $p = 1, 2$, and let Θ denote the parameter space. Let there be a random sample of size N from the foregoing game. Consider the following conditions:

A1.1 For each player p, there exists a continuously distributed $X_{\ell,p} \in X_p$ with nonzero coefficient $\beta_{\ell,p}$ and unbounded support conditional on $X \setminus X_{\ell,p}$ such that for any $c \in (0,1)$, $b \neq 0$, and $q \in \mathbb{R}^{\dim(X_{-\ell,p})}$, there exists $C_{b,q,m} > 0$ such that

$$\Pr(\varepsilon_p \le bX_{\ell,p} + q'X_{-\ell,p}|X) > m \quad \forall X_{\ell,p} : \operatorname{sign}(b) \cdot X_{\ell,p} > C_{b,q,m}. \tag{15}$$

A1.2 For $p = 1, 2$, let $X_{d,p}$ denote the regressors that have bounded support but are not constant. Suppose that Θ is such that for any $\beta_{d,p}, \widetilde{\beta}_{d,p} \in \Theta$ with $\widetilde{\beta}_{d,p} \neq \beta_{d,p}$ and for any $\alpha_p \in \Theta$,

$$\Pr\big(|X_{d,p}'(\beta_{d,p} - \widetilde{\beta}_{d,p})| > |\alpha_p|\big|X \setminus X_{d,p}\big) > 0. \tag{16}$$

If all we know is that players are level-1 rational, then the following hold:

a. If (A1.1) holds, then the coefficients $\beta_{\ell,p}$ are identified.
b. If (A1.2) holds, then the coefficients $\beta_{d,p}$ are identified.
c. We say that player p is pessimistic with positive probability if for any $\Delta > 0$, there exists $\mathcal{X}_\Delta \in \mathbb{S}(X_p)$ such that $\Pr(Y_p = 1|X) < \Pr(\varepsilon_p \le X_p'\beta_{b_0} + \alpha_{p_0}|X) + \Delta$ whenever $X_p \in \mathcal{X}_\Delta$. If (A1.1) and (A1.2) hold and player p is pessimistic with positive probability, then the identified set for α_p is $\{\alpha_p \in \Theta : \alpha_p \le \alpha_{p_0}\}$. (Here we refer to the identified set as the set of values of α_p that are observationally equivalent, conditional on observables, to the true value α_{p_0}.)

The results in Theorem 1 imposed no restrictions on $\mathcal{I}_p$. In particular, players can condition their beliefs on unobservables to the econometrician. A special case of condition (A1.1) is when ε_p is independent of X. The condition in (A1.2) says how rich the support of the bounded shifters must be in relation to the parameter space. Covariates with unbounded support satisfy this condition immediately given the full-rank assumption. Finally, similar identification results to Proposition 1 hold for the cases where $\alpha_p \geq 0$ and $\alpha_1\alpha_2 \leq 0$. The proof of Theorem 1 is given in the Appendix.

4.4.2 Identification With Level-k Rationality. We now move on to the case of rationalizable beliefs of higher order. Our goal is to investigate whether a higher degree of rationality will the task of point-identifying α_p. To simplify the analysis, we assume from here on that ε_p is independent of X and of $\mathcal{I} \equiv (\mathcal{I}_1, \mathcal{I}_2)$. This assumption could be replaced with one along the lines of (A1) in Theorem 1. We make the assumption that $\mathcal{I}$ is observed by the econometrician; we relax it later. Again, denote the common prior assumption by $H_p(\cdot)$. The beliefs of the players for any level-k rationality can be constructed as done in the previous section. Our point identification–sufficient conditions are summarized in Theorem 2.

Theorem 2. Suppose that there exists a subset $\mathcal{X}_1^* \subseteq \mathbb{S}(X_1)$ where X_1 has full-column rank such that for any $X_1 \in \mathcal{X}_1^*$, $\varepsilon > 0$, and $\theta_2 \in \Theta$, there exist $\Im^*_{1_\varepsilon} \subset \mathbb{S}(\mathcal{I}_1|X_1)$ and $\Im^{**}_{1_\varepsilon} \subset \mathbb{S}(\mathcal{I}_1|X_1)$ such that

$$\begin{aligned}
&\text{for all } \mathcal{I}_1 \in \Im^*_{1_\varepsilon},\\
&\max\{1 - E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1],\\
&\qquad E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1] - E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1]\}\\
&\quad < \varepsilon,\\
&\text{for all } \mathcal{I}_1 \in \Im^{**}_{1_\varepsilon},\\
&\max\{E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1],\\
&\qquad E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1] - E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1]\}\\
&\quad < \varepsilon.
\end{aligned} \tag{17}$$

A special case in which (17) holds is when there exists $X_{2_\ell} \in (X_2 \cap W_1)$ with nonzero coefficient in Θ such that X_{2_ℓ} has unbounded support conditional on $(X_2 \cup W_1) \setminus X_{2_\ell}$. We can call (17) an "informative signal" condition. Note that implicit in (17) is an exclusion restriction in the parameter space that precludes $\beta_2 = 0$ for any $\theta_2 \in \Theta$. If (17) holds, then for any $\theta \in \Theta$ such that $\theta_1 \neq \theta_{1_0}$, there exists either $\mathcal{W}_1^* \subset \mathbb{S}(W_1)$ or $\mathcal{W}_1^{**} \subset \mathbb{S}(W_1)$ such that

$$\begin{aligned}
&H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k;\mathcal{I}_1)\big)\\
&\quad < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^U(\theta_0|k;\mathcal{I}_1)\big)\\
&\quad\quad \forall W_1 \in \mathcal{W}_1^*,\ k \geq 2;\\
&H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k;\mathcal{I}_1)\big)\\
&\quad > H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k;\mathcal{I}_1)\big)\\
&\quad\quad \forall W_1 \in \mathcal{W}_1^{**},\ k \geq 2.
\end{aligned} \tag{18}$$

Therefore, for any $k \geq 2$, the level-k rationalizable bounds for player 1's conditional choice probability of $Y_1 = 1|W_1$ that correspond to θ will be disjoint with those of θ_0 with positive probability. Consequently, if (17) holds and the population of player 1's are at least level-2 rational, θ_{1_0} is identified. By symmetry, θ_{2_0} will be point-identified if the foregoing conditions hold, with the subscripts "1" and "2" interchanged.

For the case in which $\mathcal{I}_1 = \mathcal{I}_2 = X$ and the only source of private information in payoffs is ε_p, Figures 7 and 8 illustrate four graphical examples of how the "informative signals" condition (17) in Theorem 2 yields disjoint level-2 bounds.

The ability to shift the upper and lower bounds for level-2 rationalizable beliefs arbitrarily close to 1 or 0 is essential for the point-identification result in Theorem 2. For simplicity, the intercept Δ_1 is subsumed in $X_1'\beta_1$ in the labels of these figures.

4.5 On Identification of Players' Rationality Level

Without further structure, our setup is not capable of identifying each individual player's rationality level (measured by k). Furthermore, without strong assumptions about the support of ε_p relative to that of $X_p'\beta_p$, it is not possible to reject a value of

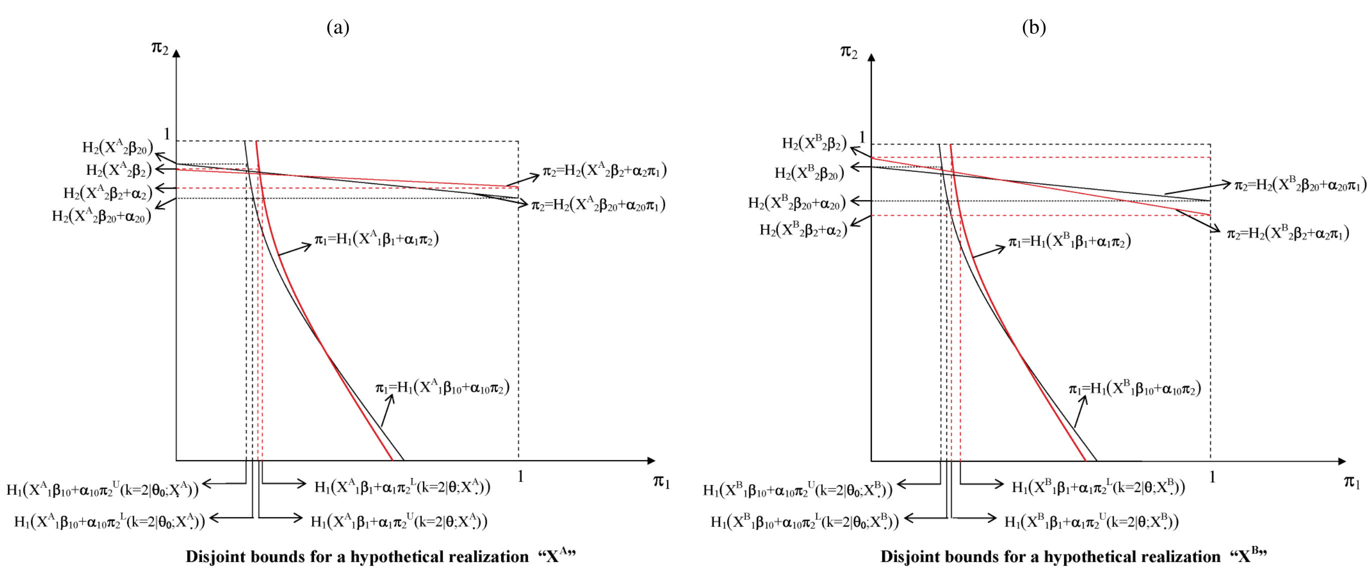

Figure 7. Graphical examples of informative signals, I.

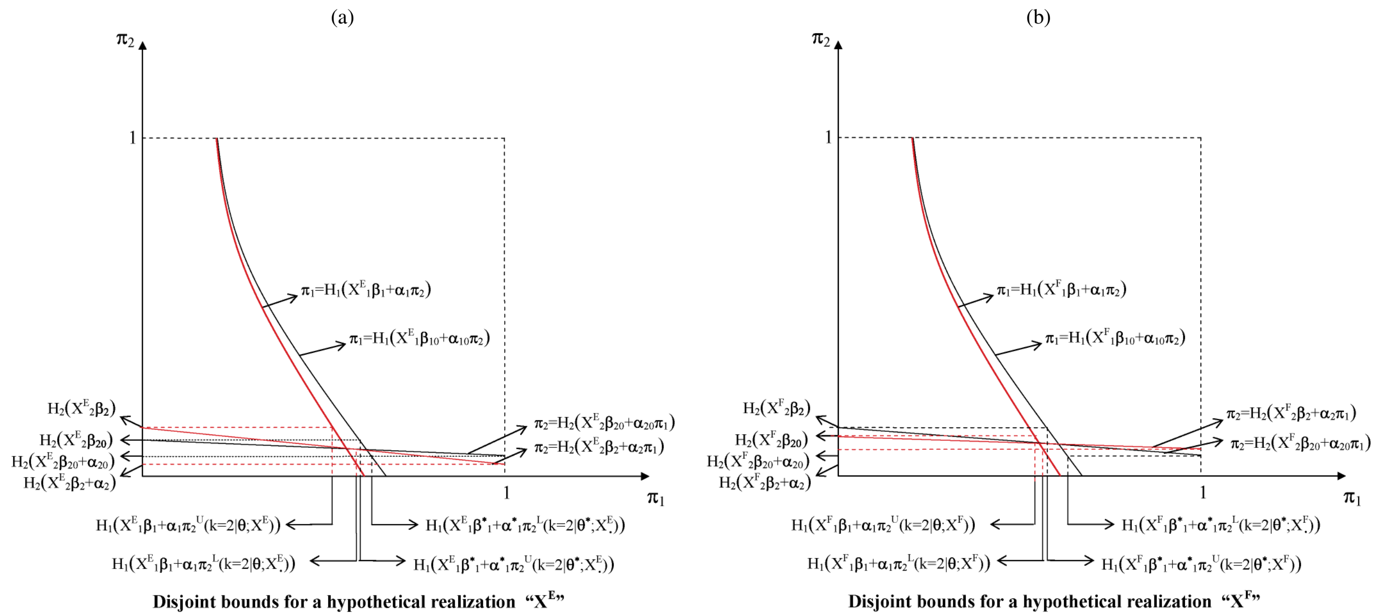

Figure 8. Graphical examples of informative signals, II.

k on the basis of observed choices. But our setup is capable of producing identification results for the value k_0 such that players in the population are at most level-k_0 rational. This refers to the value such that the level-k_0 bounds hold with probability 1, but the level-(k_0+1) bounds are violated with positive probability in the population. In other words, our setup has the potential to identify the rationality level k_0 such that a portion of players in the population have beliefs that violate the level-(k_0+1) rationalizable bounds. Whether or not we can identify k_0 depends on how much we can identify about θ. If all players are at least level-2 rational and the conditions for point identification of θ described in Theorem 2 hold, then k_0 would be point-identified because $Q(\theta_0|k)=0$ if and only if $k \le k_0$, where $Q(\theta|k)$ is as defined in (14). To see why this is not true when θ is set-identified, refer to parts (a) and (b) following (19). Otherwise, if the conditions for Theorem 2 do not hold, then suppose that we maintain the assumption $k_0 \ge 1$ (the only interesting case). We can start with $k=1$ and construct $\Theta(1)$, as defined in (14). Next, for any $k \ge 2$, define

$$\underline{Q}(k) = \min_{\theta \in \Theta(1)} Q(\theta|k), \tag{19}$$

where $Q(\theta|k)$ is as defined in (14). Then the following hold:

a. $\underline{Q}(k)=0$ for all $k \le k_0$; however, $\underline{Q}(k)=0$ does not imply, $k \le k_0$.
b. $\underline{Q}(k)>0$ implies that $k>k_0$.

Suppose that different observations in the data set correspond to a game with a different level of rationality; then if $\underline{Q}(k)>0$ and $\underline{Q}(k-1)=0$, we would reject the hypothesis (strictly speaking, this would be a joint test of the rationality hypothesis and all other maintained assumptions) that all of the population is at least level-k rational. If we assumed ex ante that $k_0 \ge \underline{k} > 1$, then we could simply replace $\Theta(1)$ with $\Theta(\underline{k})$ in the definition of $\underline{Q}(k)$ in (19). Alternatively, in settings where at least a subset of the structural parameter θ is known (e.g., experiments), we could evaluate whether players are at least level-k_0 rational by testing whether or not $\theta_0 \in \Theta(k_0)$ (the identified set for level-k_0 rationality). Otherwise, a test that would fail to reject $\Theta(k_0+1)=\emptyset$ would indicate that players are at most level-k_0 rational.

4.6 Bayesian Nash Equilibria and Rationalizable Beliefs

As before, let $\mathcal{I}_p$ be the signal that player p uses to condition his beliefs about his opponent's expected choice, and let $\mathcal{I} \equiv (\mathcal{I}_1, \mathcal{I}_2)$. The set of BNE is defined as any pair $(\pi_1^*(\mathcal{I}_2), \pi_2^*(\mathcal{I}_1)) \equiv \pi^*(\mathcal{I})$ that satisfies

$$\begin{aligned} \pi_1^*(\mathcal{I}_2) &= E\big[H_1(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1))|\mathcal{I}_2\big], \\ \pi_2^*(\mathcal{I}_1) &= E\big[H_2(X_2'\beta_2 + \alpha_2\pi_1^*(\mathcal{I}_2))|\mathcal{I}_1\big]. \end{aligned} \tag{20}$$

By construction, the set of rationalizable beliefs for $\mathcal{I}$ must include the BNE set for any rational level k. The following result formalizes this claim.

Proposition 1. Let

$$\mathcal{R}(\mathcal{I};k) = [\pi_1^L(\theta|k;\mathcal{I}_2), \pi_1^U(\theta|k;\mathcal{I}_2)] \times [\pi_2^L(\theta|k;\mathcal{I}_1), \pi_2^U(\theta|k;\mathcal{I}_1)]$$

denote the set of level-k rationalizable beliefs. Then, with probability 1, the BNE set described in (20) is contained in $\mathcal{R}(\mathcal{I};k)$ for *any* $k \ge 1$.

We present the proof for the case where $\alpha_p \le 0$ for $p=1,2$, on which we have focused. The proof can be adapted to all other cases. We proceed by induction by first proving the following claim.

Claim 2. Let $\pi^*(\mathcal{I}) \equiv (\pi_1^*(\mathcal{I}_2), \pi_2^*(\mathcal{I}_1))$ be any BNE. Then, for any $k \ge 1$, with probability 1, we have that $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k)$ implies that $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I};k+1)$ with probability 1.

Proof. If $\alpha_p = 0$ for $p=1$ or $p=2$, then the result follows trivially. Suppose that $\alpha_1 = 0$; then $\pi_1^L(\theta|k;\mathcal{I}_2) = \pi_1^U(\theta|k;\mathcal{I}_2) = \pi_1^*(\mathcal{I}_2) = E[H_1(X_1'\beta_1)|\mathcal{I}_2]$ and $\pi_2^L(\theta|k;\mathcal{I}_1) = \pi_2^U(\theta|k;\mathcal{I}_1) = \pi_2^*(\mathcal{I}_1) = E[H_2(X_2'\beta_2 + \alpha_2\pi_1^*(\mathcal{I}_2))|\mathcal{I}_1]$ for all

$k \geq 1$. We focus on the case where $\alpha_p < 0$ for $p = 1, 2$. Now suppose that $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k)$ but $\pi^*(\mathcal{I}) \notin \mathcal{R}(\mathcal{I}; k)$. Suppose, for example, that $\pi_1^L(\theta|k+1; \mathcal{I}_2) > \pi_1^*(\mathcal{I}_2)$. Because $\alpha_1 < 0$, this can be true if and only if

$$\underbrace{E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^U(\theta|k; \mathcal{I}_1)\big)\big|\mathcal{I}_2\big]}_{=\pi_1^L(\theta|k+1;\mathcal{I}_2)} > \underbrace{E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1)\big)\big|\mathcal{I}_2\big]}_{=\pi_1^*(\mathcal{I}_2)}.$$

For this inequality to be satisfied, it cannot be the case that $\pi_2^*(\mathcal{I}_1) \leq \pi_2^U(\theta|k; \mathcal{I}_1)$. But this violates the assumption that $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k)$; therefore, we must have $\pi_1^L(\theta|k+1; \mathcal{I}_2) \leq \pi_1^*(\mathcal{I}_2)$. Suppose now that $\pi_1^U(\theta|k; \mathcal{I}_2) < \pi_1^*(\mathcal{I}_2)$. This can be true if and only if

$$\underbrace{E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^L(\theta|k; \mathcal{I}_1)\big)\big|\mathcal{I}_2\big]}_{=\pi_1^U(\theta|k+1;\mathcal{I}_2)} < E\big[H_1\big(X_1'\beta_1 + \alpha_1\pi_2^*(\mathcal{I}_1)\big)\big|\mathcal{I}_2\big].$$

For this inequality to be satisfied, it cannot be the case that $\pi_2^*(\mathcal{I}_1) \geq \pi_2^L(\theta|k; \mathcal{I}_1)$. Once again, this violates the assumption $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k)$; therefore, we must have $\pi_1^U(\theta|k+1; \mathcal{I}_2) \geq \pi_1^*(\mathcal{I}_2)$. These results imply that we must have $\pi_1^L(\theta|k+1; \mathcal{I}_2) \leq \pi_1^*(\mathcal{I}_2) \leq \pi_1^U(\theta|k+1; \mathcal{I}_2)$. Following the same steps, we can establish that we must have $\pi_2^L(\theta|k+1; \mathcal{I}_1) \leq \pi_2^*(\mathcal{I}_1) \leq \pi_2^U(\theta|k+1; \mathcal{I}_1)$. Combined, these yield $\pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k+1)$, as claimed.

Proof of Proposition 1. Follows from Claim 2 and the fact that level-1 rational players satisfy $H_p(X_p'\beta_p + \alpha_p) \leq E[Y_p|X_p] \leq H_p(X_p'\beta_p)$, which yields $\mathcal{R}(\mathcal{I}; k = 1) = [0, 1] \times [0, 1]$. Consequently, $\mathcal{R}(\mathcal{I}; k = 1)$ contains all BNE. It follows from Claim 2 that $\mathcal{R}(\mathcal{I}; k = 1)$ contains all BNE for all $k \geq 1$.

4.7 BNE versus Rationalizability: Identification

Naturally, it is always guaranteed that one gets a weakly *smaller* identified set with BNE assumptions, because the predicted outcomes based on equilibrium use stronger assumptions on player beliefs. The size of the rationalizable outcome set depends on the distance between the smallest and largest equilibria [more precisely, the distance (in the unit square) between the smallest and the largest equilibrium beliefs]. In the case of a unique equilibrium, we can see that in the foregoing game and as $k \to \infty$, the predicted outcomes under both solution concepts converge. This convergence feature is not a general property of rationalizability, but rather is a consequence of the normal-form payoff parameterization of the game. In addition, in the foregoing simple example, predicted outcomes based on rationality of order k for any k are much easier to solve for, because they do not require solutions to fixed-point problems, especially in cases of multiple equilibria.

5. IDENTIFICATION IN FIRST–PRICE INDEPENDENT PRIVATE VALUE AUCTIONS WITH RATIONALIZABLE BIDS

This section considers a situation in which a population of symmetric, risk-neutral potential buyers must bid simultaneously for a single good. We focus on a first-price auction with independent private values, although our results can be adapted to the case of interdependent private values and affiliated signals. As is usually the case in the econometric analysis of auctions, the object of interest is the distribution of private values. Under the assumption that observed bids conform to a BNE, nonparametric point identification for this distribution has been established by, for example, Guerre, Perrigne, and Vuong (1999). Thus equilibrium assumptions (and other conditions) deliver point identification of the valuation distribution. Here we relax the BNE requirement and assume only that buyers are strategically sophisticated in the sense of Battigalli and Siniscalchi (2003, henceforth BS). Other strategic assumptions that can be used and that deliver qualitatively different results than BS's interim rationalizability is the $\mathcal{P}$-dominance concept introduced for auction setups by Dekel and Wolinsky (2003) and used more recently by Crawford and Iriberri (2007). Here we highlight just what can be learned with the BS setup and compare that with BNE. BNE requires rational, expected utility maximizing buyers with correct beliefs. Strategically sophisticated buyers are rational and expected-utility maximizers, but their beliefs may or may not be correct. This characterization includes BNE as a special case. The degree of sophistication will be characterized using the concept of interim rationalizability. As we show, this will lead to the notion of *level-k rationalizable bids* for $k \in \mathbb{N}$. We describe these concepts next.

Let $F_0(\cdot)$ denote the distribution of v_i, the private valuation of bidder i. We assume $F_0(\cdot)$ to be common knowledge among the bidders, and focus on the case where $F_0(\cdot)$ is log-concave and absolutely continuous with respect to Lebesgue measure. We assume its support to be of the form $[0, \omega)$ (i.e., normalize its lower bound by 0) and allow, in principle, the case where $\omega = +\infty$. Assume for the moment that the seller's reservation price p_0 is equal to 0. We explicitly introduce a strictly positive reservation price later.

5.1 Assumptions About Bidders' Beliefs

Following BS, we assume that bidders expect all positive bids to win with strictly positive probability and that this is common knowledge. This condition will ensure that it is common knowledge that no bidder will bid beyond his or her valuation irrespective of his or her beliefs. It also implies that every bidder with nonzero private value will submit a strictly positive bid. Interim rationalizability will naturally produce only upper bounds for rationalizable bids. Additional, ad hoc assumptions can be made to characterize a lower bound. Therefore, with probability 1, the number of potential bidders $\mathcal{N}$ is equal to the number of actual bidders. (Only a bidder with valuation equal to 0 is indifferent between entering the bid or not.) We restrict further attention to beliefs that assign positive probability only to increasing bidding functions. Formally, let $\mathcal{B}$ denote the space of all functions of the form

$$\mathcal{B} = \big\{b : [0, \omega) \to \mathbb{R}_+ : b(v) \leq v, \text{ and } v > v' \Rightarrow b(v) > b(v')\big\}. \tag{21}$$

We let $\mathcal{N}$ denote the number of potential bidders in the population and write $\mathcal{B}_{-i} = \mathcal{B}^{\mathcal{N}-1}$. Beliefs for bidder i are probability distributions defined over a sigma-algebra $\Delta_{\mathcal{B}_{-i}}$, where this sigma-algebra is such that singletons in $\mathcal{B}_{-i}$ are measurable. The results that we analyze here do not depend on the specific choice of the sigma-algebra, as long as this choice satisfies the singleton-measurability mentioned here (see footnote 10 in BS). A conjecture by bidder i is a degenerate belief that assigns probability mass 1 to a singleton $\{b_j\}_{j\neq i} \in \mathcal{B}_{-i}$. The distribution of valuations $F_0(\cdot)$ as well as $\mathcal{N}$ are common knowledge among potential bidders. This is similar to the common prior assumption made in the previous section.

As we show, restricting attention to beliefs in $\mathcal{B}$ will yield rationalizable upper bounds for bids that also belong in $\mathcal{B}$. It also simplifies the analysis by, for example, ruling out ties in the characterization of players' expected utility. Finally, as we argue below (and formally shown in BS), restricting attention to beliefs in $\mathcal{B}$ implies that BNE-optimal bids are always rationalizable.

5.2 Implications of Level-k Rationality in Bids

Here we follow the setup in BS, with our notation differing from that of previous sections. We have a population of $\mathcal{N}$ risk-neutral potential buyers bidding simultaneously for a single object. With a 0 reservation price of 0, we can interpret $\mathcal{N}$ as the number of observed bids that is common knowledge among the bidders. Each bidder i observes his valuation v_i, independent of those of other bidders with an identical log-concave, continuous distribution $F_0(\cdot)$. The highest bid wins the object, ties are broken at random and only the winner pays his bid. The space of beliefs on which we focus assigns probability 0 to ties. Therefore, the decision problem for bidder i can be expressed as

$$\max_{b\geq 0}(v_i - b)\widehat{\Pr}_i\Big[\max_{j\neq i} b(v_j) \leq b\Big], \tag{22}$$

where $\widehat{\Pr}_i(\cdot)$ denotes bidder i's subjective probability, derived from his beliefs and knowledge of $F_0(\cdot)$. (Strictly speaking, what matters is that ties have probability 0 for the most pessimistic conjecture.) For level-1 rational bidding, any bidder i whose bids satisfy

$$b \leq v_i \equiv \overline{B}_1(v_i; \mathcal{N}) \quad \text{with probability 1} \tag{23}$$

are called *level-1 rational bidders*. Any expected-utility maximizer bidder i must be level-1 rational regardless of whether or not his beliefs live in $\mathcal{B}_{-i}$. Thus we have the following:

Result (BS). Any bid with $b_i \leq v_i$ is level-1 rational.

This was proved by BS, who also showed that the bound is sharp; that is, that for any bid in the bound, there exists a consistent and valid level-1 belief function for which that bid is a best response. This result is interesting because in this setup, the bids cannot be bound from below. This situation is in marked contrast to the BNE prediction. Note that the bound above depends on the continuity of the valuation and the assumption that any positive bid has a positive chance of winning. In another case where the valuations are assumed to take countable values, Dekel and Wolinsky (2003) showed that a form of rationalizability implies tight bounds on the bidding function in the limit as the number of bidders increases. Here we derive strategies for identification of $F(\cdot)$ based on the results of BS, but these strategies can be easily adapted to other strategic setups, like those suggested by Dekel and Wolinsky.

Higher-Order Rationality. We now characterize the identified features in an auction with higher rationality levels. Focus on bidders with beliefs in $\mathcal{B}_{-i}$. The most pessimistic assessment in $\mathcal{B}_{-i}$ is given by the conjecture $b(v_j) = \overline{B}_1(v_j; \mathcal{N}) = v_j$ for all $j \neq i$ (the upper bound for bids for level-1 rational bidders). Because bidder i knows F_0, his optimal expected utility for this assessment is

$$\begin{aligned}\max_{b\geq 0}&(v_i - b)\Pr\Big[\max_{j\neq i}\overline{B}_1(v_j;\mathcal{N}) \leq b\Big]\\ &= \max_{b\geq 0}(v_i - b)\Pr\Big[\max_{j\neq i} v_j \leq b\Big]\\ &= \max_{b\geq 0}(v_i - b)F_0(b)^{\mathcal{N}-1}\\ &\equiv \underline{\pi}_2^*(v_i;\mathcal{N}),\end{aligned} \tag{24}$$

where $\underline{\pi}_2^*(v_i;\mathcal{N})$ is the lower bound for optimal expected utility (22) for all beliefs in $\mathcal{B}_{-i}$. The upper expected utility bound for an arbitrary bid b is trivially given by $(v_i - b)$ for any possible beliefs. (No bidder would ever expect to win the good with probability higher than 1.) Any bid submitted by a rational (i.e., expected-utility maximizer) bidder with beliefs in $\mathcal{B}_{-i}$ must satisfy

$$v_i - b \geq \underline{\pi}_2^*(v_i;\mathcal{N}) \quad \Rightarrow \quad b \leq v_i - \underline{\pi}_2^*(v_i;\mathcal{N}) \equiv \overline{B}_2(v_i;\mathcal{N}) \quad \text{with probability 1.} \tag{25}$$

We refer to bidders who satisfy (25) as *level-2 rational bidders*. Given our assumptions, $\overline{B}_2(v_i;\mathcal{N})$ is increasing and concave and satisfies $\overline{B}_2(v_i;\mathcal{N}) \leq \overline{B}_1(v_i;\mathcal{N}) = v_i$, with strict inequality for all $v_i > 0$. These and more properties were enumerated by BS, who focused on a more general case that allows for interdependent values. Therefore, $\overline{B}_2 \in \mathcal{B}$. Let $\overline{S}_2(\cdot;\mathcal{N})$ denote the inverse of $\overline{B}_2(\cdot;\mathcal{N})$. We call *level-3 rational bidders* those whose beliefs incorporate the level-2 upper bound (25). The most pessimistic assessment for level-3 rational bidders is the conjecture $b(v_j) = \overline{B}_2(v_j;\mathcal{N})$ for all $j \neq i$. The optimal expected utility for this pessimistic assessment is

$$\begin{aligned}\max_{b\geq 0}&(v_i - b)\Pr\Big[\max_{j\neq i}\overline{B}_2(v_j;\mathcal{N}) \leq b\Big]\\ &= \max_{b\geq 0}(v_i - b)F_0(\overline{S}_2(b;\mathcal{N}))^{\mathcal{N}-1}\\ &\equiv \underline{\pi}_3^*(v_i;\mathcal{N}).\end{aligned} \tag{26}$$

Using the same logic that led to (25), the set of rationalizable bids for level-3 rational bidders must satisfy

$$v_i - b \geq \underline{\pi}_3^*(v_i;\mathcal{N}) \quad \Rightarrow \quad b \leq v_i - \underline{\pi}_3^*(v_i;\mathcal{N}) \equiv \overline{B}_3(v_i;\mathcal{N}) \quad \text{with probability 1.} \tag{27}$$

The level-3 upper bound for rationalizable bids, $\overline{B}_3(\cdot;\mathcal{N})$ is increasing and concave and satisfies $\overline{B}_3(\cdot;\mathcal{N}) \leq \overline{B}_2(\cdot;\mathcal{N})$, with strict inequality for nonzero valuations. To see why the last result holds, recall that $\overline{B}_2(v_i;\mathcal{N}) = v_i - \underline{\pi}_2^*(v_i;\mathcal{N}) \equiv \overline{B}_1(v_i;\mathcal{N}) - \underline{\pi}_2^*(v_i;\mathcal{N})$. Therefore, for any b, we have that

$\Pr[\max_{j\neq i} \overline{B}_2(v_j;\mathcal{N}) \leq b] \geq \Pr[\max_{j\neq i} \overline{B}_1(v_j;\mathcal{N}) \leq b]$. Immediately, this implies that $\underline{\pi}_3^*(\cdot;\mathcal{N}) \geq \underline{\pi}_2^*(\cdot;\mathcal{N})$ and thus $\overline{B}_3(\cdot;\mathcal{N}) \leq \overline{B}_2(\cdot;\mathcal{N})$. Because $F_0(\cdot)$ is not assumed to have point masses, all of the foregoing inequalities are strict for any $v_i > 0$. Proceeding iteratively, the level-k bound for rationalizable bids is given by

$$b_i \leq v_i - \underline{\pi}_k^*(v_i;\mathcal{N}) \equiv \overline{B}_k(v_i;\mathcal{N})$$

with probability 1, where

$$\underline{\pi}_k^*(v_i;\mathcal{N}) = \max_{b\geq 0}(v_i - b)F_0(\overline{S}_{k-1}(b;\mathcal{N}))^{\mathcal{N}-1} \quad (28)$$

and $\overline{S}_{k-1}(\cdot;\mathcal{N})$ is the inverse function of $\overline{B}_{k-1}(\cdot;\mathcal{N})$. The level-$k$ upper bounds for rationalizable bids, $\overline{B}_k(\cdot;\mathcal{N})$, are increasing and concave and satisfy $\overline{B}_{k+1}(v;\mathcal{N}) \leq \overline{B}_k(v;\mathcal{N})$ for all k, with strict inequality for all $v > 0$. Let $b^{\text{BNE}}(v;\mathcal{N})$ denote the optimal BNE bidding function, produced by self-consistent, correct beliefs. BS showed that $\overline{B}_k(\cdot;\mathcal{N}) \geq b^{\text{BNE}}(\cdot;\mathcal{N})$ for all $k \in \mathbb{N}$. In particular, this is true for $\lim_{k\to\infty} \overline{B}_k(\cdot;\mathcal{N})$, which is well defined by the aforementioned monotonicity property of the sequence $\{\overline{B}_k(\cdot;\mathcal{N})\}_{k\in\mathbb{N}}$. Bidding below $b^{\text{BNE}}(\cdot;\mathcal{N})$ is always rationalizable for any rationality level k. All results presented here are consistent with this type of behavior.

Example. Suppose that private values are exponentially distributed, with $F_0(v) = 1 - \exp\{-\theta v\}$ and $\theta > 0$. We have that $F_0(v)/f_0(v) = \frac{1-\exp\{-\theta v\}}{\theta \exp\{-\theta v\}} = \frac{1}{\theta}\exp\{\theta v\} - \frac{1}{\theta}$, which is an increasing function of v for all $\theta > 0$, establishing log-concavity of F_0. Figure 9 depicts $\overline{B}_k(\cdot;\mathcal{N})$, the level-$k$ rationalizable bounds for bids for the case where $\theta = -.25$, $\mathcal{N} = 2$ (two bidders), and $k = 1, 2, 3, 4$. This graphical example illustrates the features described earlier for these bounds, namely $\overline{B}_k(\cdot;\mathcal{N})$, is continuous, increasing, concave, and invertible and satisfies $\overline{B}_{k+1}(v;\mathcal{N}) \leq \overline{B}_k(v;\mathcal{N})$ for all k, with strict inequality for all $v > 0$. For this particular example, the bounds corresponding to $k \geq 5$ are graphically indistinguishable from $\overline{B}_4(v;\mathcal{N})$.

5.3 Identification With Level-*k* Rationality in a Parametric Model

In this section we exploit the foregoing bounds to learn about the distribution of valuation given a random sample of bids. We first focus on the hypothetical case where there is no reserve price set by the seller and for each auction we observe all bids submitted and also know $\mathcal{N}$, the number of potential entrants. In the next section we deal with the more general case where there is a nonzero reserve price and only winning bids are observed.

We assume a semiparametric setting where F_0 belongs to a space of log-concave, absolutely continuous distribution functions with support $[0,\omega)$ of the form

$$\mathcal{F}_v^{\Theta} = \big\{F(\cdot;\theta) : \theta \in \Theta, \text{ and } F_0(\cdot) = F(\cdot;\theta_0) \text{ for some } \theta_0 \in \Theta\big\}. \quad (29)$$

Here we also can think of Θ as a set of functions, in which case the foregoing definition accommodates nonparametric analysis. Denote the level-k upper bound that corresponds to $F(\cdot;\theta)$ by $\overline{B}_k(\cdot;\mathcal{N}|\theta)$.

Level-1 Rationality. For rationality of level 1, the game predicts that

$$0 \leq b_i^l \leq v_i^l \quad \text{for all } i = 1,\ldots,\mathcal{N},\ l = 1,\ldots,L.$$

This is a problem of inference with interval data. The b's are observed and the v's are not, but we observe a bound on every observation. The object of interest is the distribution function F of the valuations v. (Here we can introduce auction heterogeneity that is observed.) This implies that

$$F_0(t;\theta) \equiv P(v \leq t) \leq P(b \leq t) \equiv G_b(t).$$

Thus, with the first level of rationality, we can bound the valuation distribution above by the observed distribution of the bids. Here inference is handled below and is based on replacing the observed bids distribution with its consistent empirical analog.

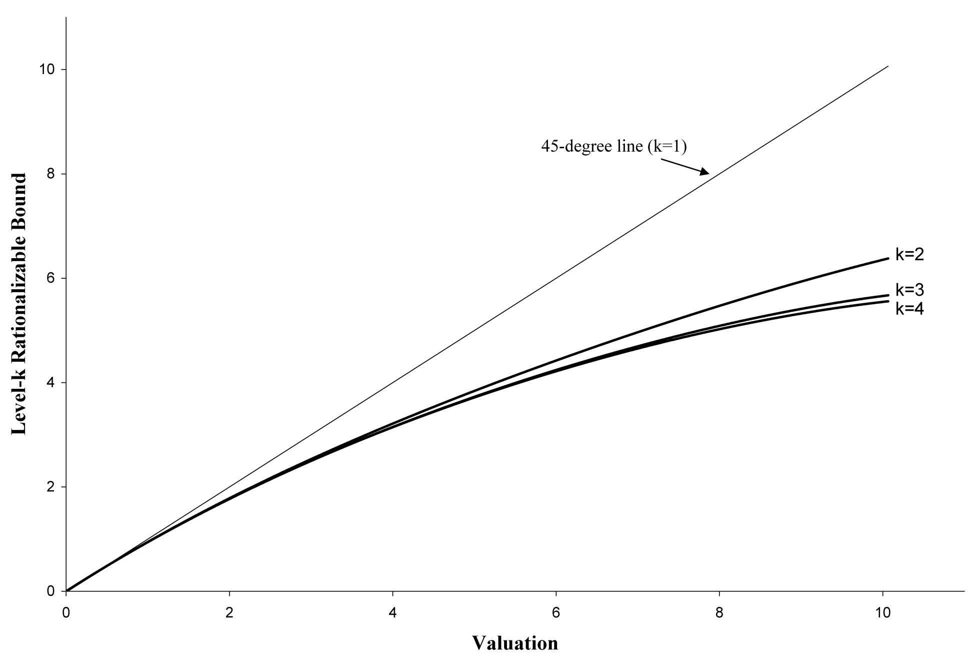

Figure 9. Level-k rationalizable bounds $\overline{B}_k(\cdot;\mathcal{N})$ for $F_0(v) = 1 - e^{-.25v}$, $\mathcal{N} = 2$, and $k = 1, 2, 3, 4$.

Level-k Rationality. Similar to the foregoing, for level-k and any $\theta \in \Theta$, we have that

$$0 \le b_i^l \le v_i^l - \underline{\pi}_k^*(v_i; \mathcal{N}|\theta) \equiv \overline{B}_k(v_i; \mathcal{N}) \quad \text{for all } i = 1, \dots, \mathcal{N},\ l = 1, \dots, L.$$

Thus this means that if bidders are level-k rational, then

$$F(\overline{S}_k(t; \mathcal{N}|\theta_0); \theta_0) \le P(b \le t) \equiv G_b(t),$$

where, as before, $\overline{S}_k$ denotes the inverse function of $\overline{B}_k$. Here the bound is a bit more complicated, because the function $\overline{S}$ also depends on F_0. Using the notation of Section 2, the space of strategies (bidding functions) for level-k rational players is

$$\mathcal{R}^i(k) = \{b \in \mathcal{B} : b(\cdot) \le \overline{B}_k(\cdot; \mathcal{N}|\theta)\}.$$

As we did in Section 4.3, we characterize the identified set for θ based on level-k rationality in terms of an objective function.

Proposition 2. Suppose that F_0 belongs to a space of distribution functions as described in (29). Moreover, suppose that we have a random sample of size L of auctions, each of which has $\mathcal{N}$ bidders and where we observe all bids. Take $k \in \mathbb{N}^+$, and let

$$\Lambda(\theta|a, c; k) = \int \left(1 - \mathbb{1}\left\{F_b(b) \ge F(\overline{S}_k(b; \mathcal{N}|\theta); \theta)\right\}\right) \times \mathbb{1}\{a \le b \le c\}\, dF_b(b), \tag{30}$$

$$\Gamma(\theta|k) = \int \int \Lambda(\theta|a, c; k)\, dF_b(a)\, dF_b(c).$$

Then, under the sole assumption that all bidders are level-k rational, the identified set is

$$\Theta(k) = \{\theta \in \Theta : \Gamma(\theta|k)^2 = 0\}.$$

If the following condition holds for a known k_0, then a stronger identification result can be obtained.

Assumption B1. Suppose that there exists k_0 such that all bidders are level-k_0 rational and, with positive probability, bids are equal to the level-k_0 bounds; that is, suppose that $\Pr(b_i \le \overline{B}_{k_0}(v_i; \mathcal{N}|\theta_0)) = 1$ and $\Pr(b_i = \overline{B}_{k_0}(v_i; \mathcal{N}|\theta_0)) > 0$.

Proposition 3. Suppose that Assumption B1 holds, and let $\Theta(k_0)$ be as defined in Proposition 2. For $\theta \in \Theta$, let

$$\mathcal{F}^c(\theta) = \left\{\theta' \in \Theta : \overline{B}_{k_0}(v_i; \mathcal{N}|\theta') < \overline{B}_{k_0}(v_i; \mathcal{N}|\theta) \text{ with probability } 1\right\}. \tag{31}$$

Then the identified set is

$$\Theta_0^* = \left\{\theta \in \Theta(k_0) : \nexists \theta' \in \Theta \text{ such that } \theta' \in \mathcal{F}^c(\theta)\right\}. \tag{32}$$

Consequently, if there exist $\overline{\theta} \in \Theta(k_0)$ such that

$$F(\cdot; \theta) < F(\cdot; \overline{\theta}) \quad \text{for all } \theta \in \Theta(k_0), \tag{33}$$

then $\Theta_0^* = \{\overline{\theta}\}$ and, consequently, $\theta_0 = \overline{\theta}$.

Under Assumption B1, $\theta \notin \Theta(k_0)$ implies that $\theta \neq \theta_0$ and $\theta \in \Theta(k_0)$ holds only if $\theta \notin \mathcal{F}^c(\theta_0)$. Suppose that we have $\theta, \theta' \in \Theta(k_0)$ such that $\theta' \in \mathcal{F}^c(\theta)$; then it cannot be the case that $\theta = \theta_0$, because then $\theta' \in \mathcal{F}^c(\theta_0)$ would imply that $\theta' \notin \Theta(k_0)$. Thus any such θ can be discarded as the true θ_0. To see how this result is constructive, suppose that $\mathcal{F}_v^{\Theta}$ is a space of exponentially distributed valuations, Assumption B1 holds, and the largest value of $\Theta(k_0)$ is $\overline{\theta} < \infty$. This would immediately imply that $\theta_0 = \overline{\theta}$. Figure 10 illustrates this result for the exponential distribution. As shown, if Assumption B1 holds with $k_0 = 2$ and if we know that $\{.25, .50, .75, 1.00, 1.25\} \subset \Theta(k_0)$, then it will follow immediately that $\theta_0 \ge 1.25$. More generally, as in previous sections, the characterization of the identified set $\Theta(k)$ is amenable to recently developed set inference methods.

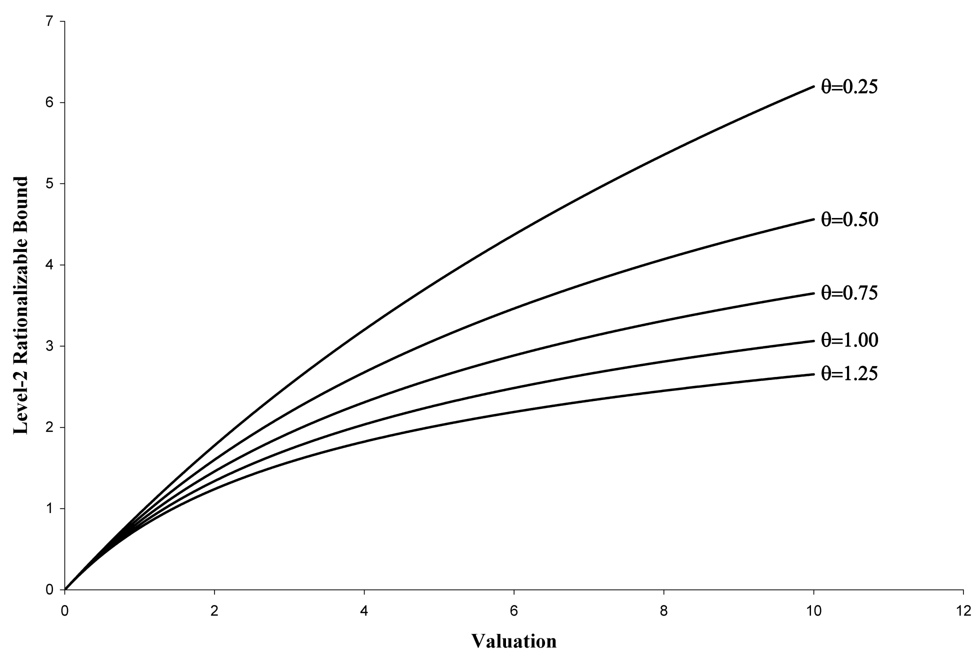

Figure 10. Level-2 upper bound for rationalizable bids for valuations with exponential distribution $F(v) = 1 - e^{-\theta v}$ and various levels of θ (number of bidders is 2). If Assumption B1 holds with $k_0 = 2$ and if we knew that $\{.25, .50, .75, 1.00, 1.25\} \subset \Theta(k_0)$, then it would follow immediately that $\theta_0 \ge 1.25$.

Remark 3. Let

$$\overline{B}_\infty(\cdot;\mathcal{N}|\theta)=\lim_{k\to\infty}\overline{B}_k(\cdot;\mathcal{N}|\theta). \quad (34)$$

Given our assumptions, the results of BS can be used to show that $\overline{B}_\infty(\cdot;\mathcal{N}|\theta)$ exists and is a continuous, increasing, concave, and invertible mapping that satisfies $\overline{B}_\infty(\cdot;\mathcal{N}|\theta)\geq b^{\text{BNE}}(\cdot;\mathcal{N}|\theta)$. Note that, unlike in the incomplete information game, here rationalizable behavior does not converge to BNE as $k\to\infty$; in particular, bidding below BNE is rationalizable for arbitrarily large k. If the rationality bound k_0 described in Assumption B1 does not exist, then we must have $b_i\leq\overline{B}_\infty(v_i;\mathcal{N}|\theta)$ with probability 1. The results in Proposition 3 will follow if the conditions stated there hold for the mapping $\overline{B}_\infty(\cdot;\mathcal{N}|\theta)$.

5.3.1 Identification When Only Winning Bids Are Observed. Suppose now that for each auction, we observe only the winning bid and the number of *actual* (as opposed to potential) entrants. We defer the introduction of nonzero reserve prices by the seller to the next section. In particular, we observe that

$$b^*=\max_{i=1,\dots,\mathcal{N}} b_i. \quad (35)$$

Under these conditions, it follows from the monotonic nature of rationalizable upper bounds that if bidders are level-k rational, with probability 1, then

$$\begin{aligned}b^*&\equiv\max_{i=1,\dots,\mathcal{N}} b_i\leq\max_{i=1,\dots,\mathcal{N}}\overline{B}_k(v_i;\mathcal{N}|\theta_0)\\&=\overline{B}_k\Big(\max_{i=1,\dots,\mathcal{N}} v_i;\mathcal{N}|\theta_0\Big)\equiv\overline{B}_k(v^*;\mathcal{N}|\theta_0).\end{aligned} \quad (36)$$

Then we must have

$$\Pr(b^*\leq b)\geq\Pr(\overline{B}_k(v^*;\mathcal{N}|\theta_0)\leq b)\qquad\forall b\in\mathbb{R}. \quad (37)$$

Because private values are iid, it follows that $v^*\sim F(\cdot;\theta_0)^{\mathcal{N}}$. Let $F_{b^*}(\cdot)$ denote the distribution function of b^*, the highest bid. Equation (37) then becomes

$$F_{b^*}(b)\geq F(\overline{S}_k(b;\mathcal{N}|\theta_0);\theta_0)^{\mathcal{N}}\qquad\forall b\in\mathbb{R}, \quad (38)$$

where, as before, $\overline{S}_k(\cdot;\mathcal{N}|\theta)$ denotes the inverse function of the upper bound $\overline{B}_k(\cdot;\mathcal{N}|\theta)$. Clearly, by the nondecreasing properties of distribution functions, (38) holds for all $b\in\mathbb{R}$ if and only if it holds for all $b\in\mathbb{S}(b^*)$ (the support of b^*). We conclude that this implies that

$$F_{b^*}(b)\geq F(\overline{S}_k(b;\mathcal{N}|\theta_0);\theta_0)^{\mathcal{N}}\qquad\forall b\in\mathbb{S}(b^*). \quad (39)$$

Equation (39) can be used as earlier to conduct inference on the set of consistent models. To do this, an objective function similar to that in Proposition 2 can be used. The results in Proposition 3 also would follow if Assumption B1 held for b^*. It appears that even if B1 were assumed to hold for all bids, then we would have to explicitly assume that it holds for b^*, because with heterogeneous beliefs, it is no longer true that the highest bid corresponds to the highest valuation among potential bidders.

5.4 Introducing a Binding Reserve Price

Suppose that there is a nonzero reserve price p_0 set by the seller and publicly observed by all potential buyers. We modify Assumption B0 accordingly as follows.

Assumption B0′. Assume now that all bidders expect any bid $b\geq p_0$ to win with strictly positive probability, and this is common knowledge. The implication of this for submitted bids is that $b_i\geq p_0$ *if and only if* $v_i\geq p_0$. We restrict attention to beliefs that assign positive probability only to bidding functions that are increasing for all $v\geq p_0$ and are equal to p_0 for $v=p_0$. Formally, let $\mathcal{B}(p_0)$ denote the space of all Borel-measurable functions of the form

$$\begin{aligned}\big\{b&:[0,\omega)\to\mathbb{R}_+:b(v)<p_0\,\forall v<p_0;\\&b(p_0)=p_0,\ \text{and for all } v>p_0:b(v)\leq v\\&\text{and } v>v'\Rightarrow b(v)>b(v')\big\}.\end{aligned} \quad (40)$$

We let $\mathcal{N}$ denote the number of potential bidders in the population and denote $\mathcal{B}_{-i}(p_0)=\mathcal{B}(p_0)^{\mathcal{N}-1}$. Beliefs for bidder i are probability distributions defined over a sigma-algebra $\Delta_{\mathcal{B}_{-i}}(p_0)$, where this sigma-algebra is such that singletons in $\mathcal{B}_{-i}$ are measurable. As before, conjectures are defined as degenerate beliefs that assign probability mass 1 to a singleton $\{b_j\}_{j\neq i}\in\mathcal{B}_{-i}$. We maintain the assumption that $F_0(\cdot)$ and $\mathcal{N}$ are common knowledge among potential bidders.

A consequence of a binding reserve price is that the number of potential bidders $\mathcal{N}$ no longer may be equal to the number of bidders who participate in the auction. Potential bidders with valuation $v_i<p_0$ will not submit a bid. Beliefs for valuations $v<p_0$ will be irrelevant for participating bidders, except for the fact that it is common knowledge that $v_j<p_0$ implies that $b_j<p_0$ with probability 1 for all potential bidders. As in the case of zero reservation price, restricting attention to beliefs in $\mathcal{B}(p_0)$ will yield rationalizable upper bounds that also belong in $\mathcal{B}(p_0)$. It also rules out ties in the characterization of expected utility for bidders with valuation $v\geq p_0$ (the only ones who participate in the auction). As in the case of zero reservation price, restricting attention to beliefs in $\mathcal{B}(p_0)$ will imply that BNE-optimal bids are always rationalizable.

5.4.1 Level-k Rationalizable Bids With a Nonzero Reserve Price. The construction of rationalizable upper bounds will follow the same interim-rationalizability steps as in Section 5.2. Any bidder i with $v_i\geq p_0$ whose bids satisfy

$$b\leq v_i\quad\text{with probability 1} \quad (41)$$

is called *level-1 rational*. Higher-rationality levels are characterized as before. The decision problem for any bidder i with $v_i\geq p_0$ now can be expressed as

$$\max_{b\geq p_0}(v_i-b)\widehat{\Pr}_i\Big[\max\Big\{p_0,\max_{j\neq i} b(v_j)\Big\}\leq b\Big], \quad (42)$$

where $\widehat{\Pr}_i(\cdot)$ denotes bidder i's subjective probability, derived from his beliefs and knowledge of $F_0(\cdot)$. The optimal bid for any assessment in $\mathcal{B}_{-i}(p_0)$ for any bidder with $v_i=p_0$ will always be $v_i=p_0$. Focusing on the case where $v_i>p_0$, the most pessimistic assessment in $\mathcal{B}_{-i}(p_0)$ is given by the conjecture that "$b(v_j) = v_j$ for all $j \neq i$ such that $v_j \geq p_0$." The optimal expected utility for this assessment is

$$\max_{b \geq p_0} (v_i - b) F_0(b)^{\mathcal{N}-1} \equiv \pi_2^*(v_i; \mathcal{N}, p_0), \tag{43}$$

which follows because $\widehat{\Pr}_i[\max\{p_0, \max_{j \neq i} v_j\} \leq b] = F_0(b)^{\mathcal{N}-1} \mathbb{1}\{b \geq p_0\}$. (Recall that F_0, $\mathcal{N}$, and p_0 are common knowledge among bidders.) Using the same arguments that followed (24), level-2 rational bidders with $v_i \geq p_0$ must satisfy

$$p_0 \leq b \leq v_i - \pi_2^*(v_i; \mathcal{N}, p_0) \equiv \overline{B}_2(v_i; \mathcal{N}, p_0). \tag{44}$$

$\overline{B}_2(v_i; \mathcal{N}, p_0)$ is the level-1 rationalizable upper bound for all bidders with $v_i \geq p_0$. It is continuous, increasing, and invertible for all $v_i \geq p_0$, with $\overline{B}_2(p_0; \mathcal{N}, p_0) = p_0$. In particular, the inverse function of $\overline{B}_2(\cdot; \mathcal{N}, p_0)$ is well defined for all values and bids $\geq p_0$. As before, we denote this inverse function by $\overline{S}_2(\cdot; \mathcal{N}, p_0)$. Note that in general, (43) has corner solutions; that is, there exists a range of valuations $v_i > p_0$ such that $\pi_2^*(v_i; \mathcal{N}, p_0) = (v_i - p_0) F_0(p_0)^{\mathcal{N}-1}$. This of course will not impact the continuity, monotonicity, and invertibility properties of the upper bound $\overline{B}_2(\cdot; \mathcal{N}, p_0)$ for values $v_i \geq p_0$. Nothing can be said about rationalizable upper bounds for $v_i < p_0$, except that they lie strictly beneath p_0. Bounds for such a range of valuations are irrelevant for the optimal decision process of bidders. Proceeding iteratively, the *level-k bound for rationalizable bids* is given by

$$b_i \leq v_i - \underline{\pi}_k^*(v_i; \mathcal{N}, p_0) \equiv \overline{B}_k(v_i; \mathcal{N}, p_0), \quad \text{where}$$
$$\underline{\pi}_k^*(v_i; \mathcal{N}) = \max_{b \geq p_0} (v_i - b) F_0(\overline{S}_{k-1}(b; \mathcal{N}, p_0))^{\mathcal{N}-1} \tag{45}$$

and $\overline{S}_{k-1}(\cdot; \mathcal{N}, p_0)$ is the inverse function of $\overline{B}_{k-1}(\cdot; \mathcal{N}, p_0)$, well defined for all values and bids $\geq p_0$.

5.4.2 Identification With Level-k Rationality When Only Winning Bids Are Observed. If we replace Assumption B0 with B0′, then all of the results in Section 5.3.1 hold with a binding reserve price for all $v_i \geq p_0$ and $b_i \geq p_0$. Consider a semiparametric setting such as that described in (29), where the distribution of valuations is allowed to depend on the publicly observed reserve price p_0,

$$\mathcal{F}_v^{\Theta, p_0} = \big\{F(\cdot; \theta, p_0) : \theta \in \Theta, \text{ and } F_0(\cdot; p_0) = F(\cdot; \theta_0, p_0) \text{ for some } \theta_0 \in \Theta\big\}. \tag{46}$$

Let $\overline{B}_k(\cdot; \mathcal{N}|\theta, p_0)$ denote the level-k upper bound for rationalizable bids that would be induced by a given distribution $F(\cdot; \theta, p_0) \in \mathcal{F}_v^{\Theta, p_0}$, and let $\overline{B}_k(\cdot; \mathcal{N}|\theta, p_0)$ denote its inverse function. Let b^* denote the winning bid, and let $F_{b^*}(\cdot; p_0)$ denote its distribution function (given p_0). Note that b^* is $\max_{i=1,\ldots,\mathcal{N}} b_i$, truncated from below at p_0. This automatic truncation ensures that the bounds in (45) are satisfied. As mentioned previously, bids below p_0 may not satisfy these bounds. If bidders are level-k rational, then for any reserve price p_0, we must have

$$F_{b^*}(b; p_0) \geq F(\overline{S}_k(b; \mathcal{N}|\theta_0, p_0); \theta_0, p_0)^{\mathcal{N}} \qquad \forall b \in \mathbb{S}(b^*|p_0), \tag{47}$$

where $\mathbb{S}(b^*|p_0)$ is the support of b^* given p_0. This result is the equivalent to (39).

Proposition 4. Suppose that F_0 belongs to a space of distribution functions as described in (46). Moreover, suppose that we have a random sample of size L of auctions, each of which has $\mathcal{N}$ bidders and where we observe only the winning bid in every auction. Let the reservation price p_0 be known. Define

$$\begin{aligned} \Lambda(\theta|a, c; k, p_0) &= \int \big(1 - \mathbb{1}\big\{F_{b^*}(b; p_0) \geq F(\overline{S}_k(b; \mathcal{N}|\theta, p_0); \theta, p_0)^{\mathcal{N}}\big\}\big) \\ &\quad \times \mathbb{1}\{a \leq b \leq c\}\, dF_{b^*}(b; p_0), \\ \Gamma(\theta|k, p_0) &= \int \int \Lambda(\theta|a, c; k, p_0)\, dF_{b^*}(a; p_0)\, dF_{b^*}(c; p_0). \end{aligned} \tag{48}$$

Then, under the sole assumption that all bidders are level-k rational, the identified set is

$$\Theta(k, p_0) = \{\theta \in \Theta : \Gamma(\theta|k, p_0)^2 = 0\}.$$

Now suppose that we assume that winning bids satisfy Assumption B1 for some k_0. For any $\theta \in \Theta$, let

$$\begin{aligned} \mathcal{F}^c(\theta, p_0) &= \big\{\theta \in \Theta : \overline{B}_{k_0}(v_i; \mathcal{N}|\theta', p_0) < \overline{B}_{k_0}(v_i; \mathcal{N}|\theta, p_0) \\ &\qquad \text{with probability } 1\big\}, \\ \Theta_0^*(p_0) &= \big\{\theta \in \Theta(k_0, p_0) : \nexists \theta' \in \Theta(k_0, p_0) \\ &\qquad \text{such that } \theta' \in \mathcal{F}^c(\theta, p_0)\big\}. \end{aligned} \tag{49}$$

Then the identified set is

$$\Theta_0^* = \big\{\theta \in \Theta : \theta \in \Theta_0^*(p_0) \text{ with probability 1 (with respect to } p_0)\big\}. \tag{50}$$

Note that the identification result in (50) requires that we explicitly assume that Assumption B1 holds for winning bids. With a nonzero reserve price, the number of actual bidders in a given auction may differ from $\mathcal{N}$. The characterization of the identified set in Proposition 4 still can be constructive in this case if we assume that $\mathcal{N}$ is the same across all auctions in the population, and if the number of actual bidders is observed. For the ℓth auction, denote the latter by I_ℓ. Therefore, $I_\ell = \sum_{i=1}^{\mathcal{N}} \mathbb{1}\{v_i \geq p_{0_\ell}\}$ and

$$E[I_\ell] = \mathcal{N} E_{p_{0_\ell}}\big[F\big(p_{0_\ell}; \theta_0, p_{0_\ell}\big)\big] \quad \Rightarrow \quad \mathcal{N} = \frac{E[I_\ell]}{E_{p_{0_\ell}}[F(p_{0_\ell}; \theta_0, p_{0_\ell})]}, \tag{51}$$

where $E_{p_{0_\ell}}[\cdot]$ denotes the expectation taken with respect to the reserve price, which is assumed to be observed for any given auction. The foregoing result is the basis for identifying $\mathcal{N}$. It then follows that Proposition 4 is a constructive identification result.

5.4.3 Identification Results for the Rationality Level k_0 in Assumption B1. Suppose we assume that there exists a finite $k_0 \geq 2$ that satisfies the conditions of Assumption B1 (otherwise see Remark 3). The results in Proposition 3 are constructive when $k_0 \geq 2$ is assumed to be known. Naturally, we would be interested in having an identification result for both θ and k_0 simultaneously.

Proposition 5. Let $\Theta(k)$ and $\Gamma(\theta|k)$ be as defined in Proposition 2. Define

$$\underline{\Gamma}(k) = \min_{\theta \in \Theta(2)} \Gamma(\theta|k)^2. \tag{52}$$

Then, if Assumption B1 is satisfied with $k_0 \geq 2$, the following results hold:

a. $\underline{\Gamma}(k) = 0$ for all $k \leq k_0$; however, $\underline{\Gamma}(k) = 0$ does not imply that $k \leq k_0$.
b. $\underline{\Gamma}(k) > 0$ implies that $k > k_0$.

It follows from Proposition 5 that any k' such that $\underline{\Gamma}(k') > 0$ can be ruled out as the true k_0 described in Assumption B1, implying that there is a subset of bidders who are strictly less than level-k' rational. At the same time, the set $\{k \in \mathbb{N} : \underline{\Gamma}(k) = 0\}$ includes all $k \leq k_0$ and also includes some values $k > k_0$.

6. CONCLUSION

In structural econometrics models, assumptions are implicitly grouped into behavioral assumptions and other auxiliary assumptions. Behavioral assumptions usually are unchallenged in identification analysis, and thus econometricians focus on the robustness of estimation results to those auxiliary assumptions (which are not implied by theory, such as functional forms and distributional assumptions). In this article we have explored the identifying role that some behavioral assumptions play. Mainly, we examined the identification power of equilibrium in three simple games. We replaced equilibrium with a form of rationality (i.e., interim rationalizability) that includes equilibrium as a special case, and compared the identified features of the game under rationality and under equilibrium. The games that we studied are stylized versions of empirical models considered and applied in the literature, and thus insights provided here can be carried over to those empirical frameworks. We do not advocate dropping the equilibrium assumptions from empirical work, however; rather, we have simply examined the identifying power of equilibrium in these simple setups. For example, it is not clear that we would want to drop equilibrium in a first-price auction, because the underlying interim-rationalizability based model may not provide strong restrictions on the observed bids as they relate to the underlying valuations. Ultimately, the researcher faces the usual trade-off between robustness and predictive power, requiring a balancing act guided by the economics of the particular application at hand. We also do not advocate using rationalizability per se as the basis for strategic interaction. Other frameworks are available in the literature, but, we note that the form of rationalizability used here has received much attention from game theorists (see, e.g., Morris and Shin 2003; Dekel et al. 2007; and references cited therein). Moreover, interim rationalizability allows us to incorporate the concept of higher-order beliefs into the econometric analysis through what we have defined here as rationality levels.

Some questions remain to be answered, and we leave these for ongoing and future work. As far as our results here, we are concerned with identification. A natural extension would be to study the statistical properties of estimators proposed herein and apply those estimators in empirical examples. Another important question is the issue of sharpness and whether the inequality-based inference procedures implied by the model deliver wide identified sets for parameters (as compared to the sharp identified sets). These interference procedures are attractive because of they lead to simple to compute estimators. More work needs to be done to look for other estimators that deliver sharp inferences. Another avenue of research is to extend some of the ideas to dynamic setups. It is well known that inference in dynamic games is difficult when one tries to account for the presence of multiple equilibria. Recent important contributions to this field have been made by Aguiregabiria and Mira (2007), Bajari, Benkard, and Levin (2007), Pesendorfer and Schmidt-Dengler (2004), and Pakes, Ostrovsky, and Berry (2005) (see also Berry and Tamer 1996). The identification question is complicated mainly due to the complexity of the underlying economic model and say beliefs off the equilibrium, where no data are available. From a practical perspective, estimating dynamic games while allowing for generation of different data points by a different equilibrium is a difficult problem because it involves solving for multiple fixed points in a complicated nonlinear problem; thus relaxing equilibrium in this setting might lead to enormous computational advantages because there is not need to solve for these fixed points. It also may be possible to examine the identification power of other strategic concepts that would be natural in dynamic settings, such as the self-confirming equilibria of Fudenberg and Levine (1993). In addition to examining the robustness to equilibrium assumptions, these identification framework can be used to study whether inference under these different strategic frameworks is more practically useful for applied researchers. We leave these topics for future research.

ACKNOWLEDGMENTS

The authors thank seminar participants at various institutions and M. Engers, D. Fudenberg, F. Molinari, M. Sinischalchi, and A. Pakes for useful comments. They especially thank Serena Ng for her diligent editorial help and the discussants for their valuable comments. The work of Aradillas-Lopez was supported by National Science Foundation grant SES 0718409, the Gregory C. Chow Econometric Research Program at Princeton University. The work of Tamer was supported by the National Science Foundation.

APPENDIX: PROOFS

Proof of Theorem 1

From our previous analysis, we know that both players are level-1 rational if and only if, with probability one in $\mathbb{S}(X)$,

$$\begin{aligned}
&\Pr(\varepsilon_1 > X_1'\beta_1, \varepsilon_2 > X_2'\beta_2|X) \\
&\quad \leq \Pr(Y_1 = 0, Y_2 = 0|X) \\
&\quad \leq \Pr(\varepsilon_1 > X_1'\beta_1 + \alpha_1, \varepsilon_2 > X_2'\beta_2 + \alpha_2|X), \\
&\Pr(\varepsilon_1 \leq X_1'\beta_1 + \alpha_1, \varepsilon_2 > X_2'\beta_2|X) \\
&\quad \leq \Pr(Y_1 = 1, Y_2 = 0|X) \\
&\quad \leq \Pr(\varepsilon_1 \leq X_1'\beta_1, \varepsilon_2 > X_2'\beta_2 + \alpha_2|X), \\
&\Pr(\varepsilon_1 > X_1'\beta_1, \varepsilon_2 \leq X_2'\beta_2 + \alpha_2|X) \\
&\quad \leq \Pr(Y_1 = 0, Y_2 = 1|X)
\end{aligned} \tag{A.1}$$

$$\leq \Pr(\varepsilon_1 > X_1'\beta_1 + \alpha_1, \varepsilon_2 \leq X_2'\beta_2|X),$$
$$\Pr(\varepsilon_1 \leq X_1'\beta_1 + \alpha_1, \varepsilon_2 \leq X_2'\beta_2 + \alpha_2|X)$$
$$\leq \Pr(Y_1 = 1, Y_2 = 1|X)$$
$$\leq \Pr(\varepsilon_1 \leq X_1'\beta_1, \varepsilon_2 \leq X_2'\beta_2|X).$$

We denote the true parameter value by θ_0. To prove part a, take any $\widetilde{\beta}_1 \neq \beta_{1_0}$ such that $\widetilde{\beta}_{\ell,1} \neq \beta_{\ell,1_0}$. Given this and the support properties of $X_{\ell,1}$, for any scalar d, we can observe either of the following two events with positive probability: (a) $X_1'\widetilde{\beta}_1 + d > X_1'\beta_{1_0}$ or (b) $X_1'\widetilde{\beta}_1 < X_1'\beta_{1_0} + \alpha_{1_0}$. Take case (a) first. With $d = \alpha_1$ (arbitrary), if $\widetilde{\beta}_{2,\ell}\beta_{2,\ell_0} > 0$, then we can make $\widetilde{\beta}_2'X_2 \to +\infty$ and $\beta_{2_0}'X_2 \to +\infty$. By Assumption (A1), this yields $\Pr(\varepsilon_1 \leq X_1'\widetilde{\beta}_1 + \alpha_1, \varepsilon_2 \leq X_2'\widetilde{\beta}_2 + \alpha_2|X) \to \Pr(\varepsilon_1 \leq X_1'\widetilde{\beta}_1 + \alpha_1|X)$ and $\Pr(\varepsilon_1 \leq X_1'\beta_{1_0}, \varepsilon_2 \leq X_2'\beta_{2_0}|X) \to \Pr(\varepsilon_1 \leq X_1'\beta_{1_0}|X) < \Pr(\varepsilon_1 \leq X_1'\widetilde{\beta}_1 + \alpha_1|X)$. Therefore, with positive probability as X_2 explodes, $\Pr(\varepsilon_1 \leq X_1'\widetilde{\beta}_1 + \alpha_1, \varepsilon_2 \leq X_2'\widetilde{\beta}_2 + \alpha_2|X) > \Pr(\varepsilon_1 \leq X_1'\beta_{1_0}, \varepsilon_2 \leq X_2'\beta_{2_0}|X) > \Pr(Y_1 = 1, Y_2 = 1|X)$, which violates (A.1). If $\widetilde{\beta}_{2,\ell}\beta_{2,\ell_0} < 0$, then the result is easier to obtain by making $\widetilde{\beta}_2'X_2 \to +\infty$ and $\beta_{2_0}'X_2 \to -\infty$. For case (b), drive $\widetilde{\beta}_2'X_2 \to -\infty$ and $\beta_{2_0}'X_2 \to -\infty$ if $\widetilde{\beta}_2\beta_{2_0} > 0$, or $\widetilde{\beta}_2'X_2 \to -\infty$ and $\beta_{2_0}'X_2 \to +\infty$ if $\widetilde{\beta}_2\beta_{2_0} > 0$. In either case, we eventually obtain $\Pr(\varepsilon_1 > X_1'\widetilde{\beta}_1, \varepsilon_2 > X_2'\widetilde{\beta}_2|X) > \Pr(\varepsilon_1 > X_1'\beta_{1_0} + \alpha_{1_0}, \varepsilon_2 > X_2'\beta_{2_0} + \alpha_{2_0}|X) > \Pr(Y_1 = 0, Y_2 = 0|X)$, which violates (A.1). This establishes the identification of $\beta_{\ell,1}$; an analog proof shows that $\beta_{\ell,2}$ is identified, which proves part a.

To establish part b, focus on the worst-case scenario and take $\widetilde{\theta} \neq \theta_0$ where $\widetilde{\beta}_{d,p} \neq \beta_{d,p0}$ but $\widetilde{\beta}_{\ell,p} = \beta_{\ell,p_0}$ for $p = 1, 2$; the parameters of the unbounded-support shifters are fixed at their true values. Here identification must rely on the properties of $X_{d,p}$, the bounded-support shifters. The condition in the statement of the proposition ensures that (a) or (b) hold even if we fix $\widetilde{\beta}_{\ell,p} = \beta_{\ell,p_0}$. To complete the proof of b we proceed as in the previous paragraph. (Note that we now have $\widetilde{\beta}_{\ell,p}\beta_{\ell,p_0} > 0$.) The case where $\widetilde{\beta}_{d,p} \neq \beta_{d,p0}$ and $\widetilde{\beta}_{\ell,p} \neq \beta_{\ell,p_0}$ is straightforward along the same lines.

Now we proceeded to part c. Consider $\widetilde{\theta}$ to be equal to θ_0 element-by-element except for $\widetilde{\alpha}_1 \neq \alpha_{1_0}$, and recall that the parameter space of interest has $\alpha_p \leq 0$. Clearly, none of the lower bounds in (A.1) evaluated at $\widetilde{\theta}$ will ever be larger than the corresponding upper bounds evaluated at θ_0, and none of the upper bounds evaluated at $\widetilde{\theta}$ will ever be smaller than the corresponding lower bounds evaluated at θ_0. Therefore, without further assumptions, $\widetilde{\theta}$ and θ_0 are observationally equivalent and α_1 is not identified. The only way that we can proceed is by adding more structure on $\Pr(Y_1, Y_2|X)$. We have $\Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \alpha_{1_0}) \leq \Pr(Y_1 = 1|X) \leq \Pr(\varepsilon_1 \leq X_1'\beta_{1_0})$; therefore, $\Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \widetilde{\alpha}_1) > \Pr(Y_1 = 1|X)$ only if $\widetilde{\alpha}_1 > \alpha_{1_0}$. Thus $\widetilde{\theta}$ can violate (A.1) only if $\widetilde{\alpha}_1 > \alpha_{1_0}$. For any such $\widetilde{\alpha}_1$, let $\Delta = \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \widetilde{\alpha}_1) - \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \alpha_{1_0}) > 0$. By the assumption in part c, there exists a subset $\mathcal{X}_1 \in \mathbb{S}(X_1)$ such that $\Pr(Y_1 = 1|X) < \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \alpha_{1_0}) + \Delta = \Pr(\varepsilon_1 \leq X_1'\beta_{1_0} + \widetilde{\alpha}_1)$. Make $X_2'\beta_{2_0} \to +\infty$, and the lower bound on the fourth inequality in (A.1) will be violated. This establishes part c. Any $\widetilde{\theta} \neq \theta_0$ where $\widetilde{\alpha}_p \neq \alpha_{p0}$ and either $\widetilde{\beta}_{\ell,p} \neq \beta_{\ell,p_0}$ or $\widetilde{\beta}_{d,p} \neq \beta_{d,p0}$ can be shown to be not observationally equivalent to θ_0 using the same arguments as in the previous paragraphs given the assumptions in parts a and b.

Proof of Theorem 2

Suppose that there exists a subset of realizations in $\overline{\mathcal{X}}_1^* \subset \mathcal{X}_1^*$ such that

$$X_1'\beta_1 + \Delta_1 + \alpha_1 > X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} \qquad \forall X_1 \in \overline{\mathcal{X}}_1^*. \quad \text{(A.2)}$$

By continuity of the linear index and of the distribution H_1, for any $X_1 \in \overline{\mathcal{X}}_1^*$, we can find a pair $0 \leq \overline{p}^L(X_1) < \overline{p}^U(X_1) \leq 1$ such that

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\overline{p}^L(X_1)\big) < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\overline{p}^U(X_1)\big). \quad \text{(A.3)}$$

To see why $\overline{p}^L(X_1)$ and $\overline{p}^U(X_1)$ exist, fix $\overline{p}^U(X_1) = 1$. By continuity, there exists a small enough $\delta > 0$ such that $\overline{p}^L(X_1) \geq 1 - \delta$ satisfies (A.3). If condition (17) in Theorem 2 holds, then there exists $\mathcal{W}_1^* \subset \mathbb{S}(W_1)$ such that

$$\min\big\{E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1], E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0}\big)|\mathcal{I}_1\big]\big\} \geq \overline{p}^L(X_1) \qquad \forall W_1 \in \mathcal{W}_1^*, \quad \text{(A.4)}$$
$$\max\big\{E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1], E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0}\big)|\mathcal{I}_1\big]\big\} \leq \overline{p}^U(X_1) \qquad \forall W_1 \in \mathcal{W}_1^*.$$

Note trivially that because $\alpha_p \leq 0$ everywhere in Θ, we have that

$$\min\big\{E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1], E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0}\big)\big|\mathcal{I}_1\big]\big\} \leq \max\big\{E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1], E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0}\big)|\mathcal{I}_1\big]\big\}$$

with probability 1.

By definition, we have that

$$E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1] = \pi_2^L(\theta|k = 2; \mathcal{I}_1) \quad \text{and} \quad E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0}\big)|\mathcal{I}_1\big] = \pi_2^U(\theta_0|k = 2; \mathcal{I}_1). \quad \text{(A.5)}$$

Combining (A.4) and (A.5), we have that

$$\pi_2^L(\theta|k = 2; \mathcal{I}_1) \geq \overline{p}^L(X_1), \quad \pi_2^U(\theta_0|k = 2; \mathcal{I}_1) \leq \overline{p}^U(X_1) \qquad \forall W_1 \in \mathcal{W}_1^*. \quad \text{(A.6)}$$

Combining (A.3) and (A.6), we have that

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k = 2; \mathcal{I}_1)\big) \leq H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\overline{p}^L(X_1)\big) < H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\overline{p}^U(X_1)\big) \leq H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^U(\theta_0|k = 2; \mathcal{I}_1)\big) \quad \forall W_1 \in \mathcal{W}_1^*. \quad \text{(A.7)}$$

This corresponds to the case described in the first line of (18). Next, suppose that (A.2) does not hold but there exists a subset of realizations $\overline{X}_1^{**} \subset \mathcal{X}_1^*$ such that

$$X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} > X_1'\beta_1 + \Delta_1 + \alpha_1 \qquad \forall X_1 \in \overline{\mathcal{X}}_1^{**}. \quad \text{(A.8)}$$

Repeating the same arguments as before and exchanging θ and θ_0, we arrive at the equivalent of (A.7), namely

$$H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) < H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) \quad \forall W_1 \in \mathcal{W}_1^{**}. \tag{A.9}$$

This corresponds to the case described in the second line of (18). The last remaining possibility is that neither (18) nor (A.8) holds. In this case,

$$X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0} = X_1'\beta_1 + \Delta_1 + \alpha_1 \qquad \forall X_1 \in \mathcal{X}_1^*. \tag{A.10}$$

Because X_1 has full column rank in $\mathcal{X}_1^*$, (A.10) implies that $\beta_{1_0} = \beta_1$ and $\Delta_{1_0} + \alpha_{1_0} = \Delta_1 + \alpha_1$. Because $\theta_1 \neq \theta_{1_0}$, we must have either

$$\Delta_1 > \Delta_{1_0} \qquad \text{or} \qquad \Delta_1 < \Delta_{1_0}. \tag{A.11}$$

Suppose that $\Delta_1 > \Delta_{1_0}$. This immediately yields $X_1'\beta_1 + \Delta_1 > X_1'\beta_{1_0} + \Delta_{1_0}$ for all $X_1 \in \mathcal{X}_1^*$. By continuity, we can find a pair $0 \le \underline{p}^L(X_1) < \underline{p}^U(X_1) \le 1$ such that

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\underline{p}^U(X_1)\big) > H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\underline{p}^L(X_1)\big). \tag{A.12}$$

To see why $\underline{p}^L(X_1)$ and $\underline{p}^U(X_1)$ exist, fix $\underline{p}^L(X_1) = 0$. By continuity, there exists a small enough $\delta > 0$ such that $\overline{p}^U(X_1) \le \delta$ satisfies (A.12). If condition (17) in Theorem 2 holds, then there exists $\mathcal{W}_1^{***} \subset \mathbb{S}(W_1)$ such that

$$\min\big\{E[H_2(X_2'\beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1], E\big[H_2(X_2'\beta_{2_0} + \Delta_{2_0} + \alpha_{2_0})|\mathcal{I}_1\big]\big\} \ge \underline{p}^L(X_1) \qquad \forall W_1 \in \mathcal{W}_1^{***}, \tag{A.13}$$

$$\max\big\{E[H_2(X_2'\beta_2 + \Delta_2)|\mathcal{I}_1], E\big[H_2\big(X_2'\beta_{2_0} + \Delta_{2_0}\big)|\mathcal{I}_1\big]\big\} \le \underline{p}^U(X_1) \qquad \forall W_1 \in \mathcal{W}_1^{***}.$$

Using the definitions of $\pi_2^L(\theta|k=2;\mathcal{I}_1)$ and $\pi_2^U(\theta_0|k=2;\mathcal{I}_1)$ [e.g., eq. (A.5)], we obtain

$$\pi_2^U(\theta|k=2;\mathcal{I}_1) \le \underline{p}^U(X_1), \quad \pi_2^L(\theta_0|k=2;\mathcal{I}_1) \ge \underline{p}^L(X_1) \qquad \forall W_1 \in \mathcal{W}_1^{***}. \tag{A.14}$$

Using (A.12), this yields

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) > H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) \quad \forall W_1 \in \mathcal{W}_1^{***}. \tag{A.15}$$

This corresponds to a case like that described in the second line of (18). If $\Delta_1 < \Delta_{1_0}$, then the same arguments as before while exchanging θ with θ_0 would lead us to conclude that there exists a set $\mathcal{W}_1^{4*} \subset \mathbb{S}(W_1)$ such that

$$H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big) > H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big) \quad \forall W_1 \in \mathcal{W}_1^{4*}. \tag{A.16}$$

We have now established (18) in Theorem 2 for the where case $k = 2$. The cases where $k > 2$ follow immediately by recalling the monotonic property of rationalizable bounds, which says that, with probability 1,

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k+1;\mathcal{I}_1)\big) \le H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k;\mathcal{I}_1)\big) \qquad \forall k \ge 1$$

and

$$H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k+1;\mathcal{I}_1)\big) \ge H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k;\mathcal{I}_1)\big) \qquad \forall k \ge 1.$$

To see why this implies that the rationalizable bounds for player 1's conditional choice probabilities are disjoint with positive probability for all $k \ge 2$, recall that the level-2 bounds are given by

$$\begin{aligned} &\big[H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k=2;\mathcal{I}_1)\big), \\ &\quad H_1\big(X_1'\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k=2;\mathcal{I}_1)\big)\big] \quad (\text{for } \theta), \\ &\big[H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^U(\theta_0|k=2;\mathcal{I}_1)\big), \\ &\quad H_1\big(X_1'\beta_{1_0} + \Delta_{1_0} + \alpha_{1_0}\pi_2^L(\theta_0|k=2;\mathcal{I}_1)\big)\big] \quad (\text{for } \theta_0). \end{aligned} \tag{A.17}$$

It follows from our results that the level-2 rationalizable bounds for θ are disjoint from those of θ_0 with positive probability. Because the bounds for $k > 2$ are contained in those of $k = 2$ with probability 1, it follows immediately that these bounds are also disjoint for $k > 2$. It follows that if the population of player 1 agents are at least level-2 rational, then any θ with $\theta_1 \neq \theta_{1_0}$ will produce level-2 bounds that are violated with positive probability. Thus no such θ can be observationally equivalent to one that has $\theta_1 = \theta_{1_0}$, and, consequently, θ_{1_0} is identified. Naturally, if the same conditions of Theorem 2 hold when we exchange the subscripts "1" and "2," then θ_{2_0} will be identified.

[Received November 2007. Revised February 2008.]

REFERENCES

Aguiregabiria, V., and Mira, P. (2007), "Sequential Estimation of Dynamic Games," *Econometrica*, 75, 1–53.

Andrews, D., Berry, S., and Jia, P. (2003), "On Placing Bounds on Parameters of Entry Games in the Presence of Multiple Equilibria," working paper, Yale Department of Economics.

Aradillas-Lopez, A. (2005), "Semiparametric Estimation of a Simultaneous Game With Incomplete Information," working paper, Princeton University, Department of Economics.

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

Bajari, P., Hong, H., and Ryan, S. (2005), "Identification and Estimation of Discrete Games of Complete Information," working paper, Stanford University, Department of Economics.

Battigalli, P., and Sinischalchi, M. (2003), "Rationalizable Bidding in First-Price Auctions," *Games and Economic Behavior*, 45, 38–72.

Bernheim, D. (1984), "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007–1028.

Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

Berry, S., and Tamer, E. (1996), "Identification in Models of Oligopoly Entry," *Advances in Economics and Econometrics*, Vol. 2, eds. Blundell, Newey, and Persson, Cambridge University Press, pp.56–85.

Bjorn, P., and Vuong, Q. (1985), "Simultaneous Equations Models for Dummy Endogenous Variables: A Game-Theoretic Formulation With an Application

to Labor Force Participation," Working Paper 537, California Institute of Technology, Department of Economics.
Bresnahan, T., and Reiss, P. (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009.
Ciliberto, F., and Tamer, E. (2003), "Market Structure and Multiple Equilibria in Airline Markets," working paper, Northwessem University, Department of Economics.
Costa-Gomes, M., and Crawford, V. (2006), "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96, 1737–1745.
Costa-Gomes, M., Crawford, V., and Broseta, B. (2001), "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69, 1193–1235.
Crawford, V., and Iriberri, N. (2007), "Level-*k* Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica*, 75, 1721–1770.
Dekel, E., and Wolinsky, A. (2003), "Rationalizable Outcomes of Large Private-Value First-Price Discrete Auctions," *Games and Economic Behavior*, 43, 175–188.
Dekel, E., Fudenberg, D., and Levine, D. (2004), "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282–303.
Dekel, E., Fudenberg, D., and Morris, S. (2007), "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15–40.
Dominguez, M., and Lobato, I. (2004), "Consistent Estimation of Models Defined by Conditional Moment Restrictions," *Econometrica*, 72, 1601–1615.
Fudenberg, D., and Levine, D. (1993), "Self-Confirming Equilibrium," *Econometrica*, 61, 523–546.
Guerre, E., Perrigne, I., and Vuong, Q. (1999), "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–574.
Ho, T., Camerer, C., and Weigelt, K. (1998), "Iterated Dominance and Iterated Best Response in Experimental '*p*-Beauty Contests'," *American Economic Review*, 88, 947–969.
Honoré, B., and Tamer, E. (2006), "Bounds on Parameters in Panel Dynamic Discrete Choice Models," *Econometrica*, 74, 611–629.
Morris, S., and Shin, H. (2003), "Global Games: Theory and Applications," in *Advances in Economics and Econometrics. Proceedings of the Eighth World Congress of the Econometric Society*, eds. M. Dewatripont, L. Hansen, and S. Turnovsky, Cambridge, U.K.: Cambridge University Press, pp. 56–114.
Nagel, R. (1995), "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313–1326.
Pakes, A., Ostrovsky, M., and Berry, S. (2005), "Simple Estimators of the Parameters in Discrete Dynamic Games (With Entry/Exit)," working paper, Harvard Department of Economics.
Pakes, A., Porter, J., Ho, K., and Ishii, J. (2005), "Moment Inequalities and Their Applications," working paper, Harvard Department of Economics.
Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029–1050.
Pesendorfer, M., and Schmidt-Dengler, P. (2004), "Identification and Estimation of Dynamic Games," working paper, London School of Economics, Department of Economics.
Seim, K. (2002), "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," working paper, Stanford Business School.
Stahl, D., and Wilson, P. (1995), "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 218–254.
Tamer, E. T. (2003), "Incomplete Bivariate Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70, 147–167.
Weinstein, J., and Yildiz, M. (2007), "A Structure Theorem for Rationalizability With Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400.

Comment

Victor Aguirregabiria
Department of Economics, University of Toronto, Toronto, Ontario M5S 3G7, Canada
(*victor.aguirregabiria@utoronto.ca*)

In this article we study the identification of structural parameters in dynamic games when we replace the assumption of Markov perfect equilibrium (MPE) with weaker conditions, such as rational behavior and rationalizability. The identification of players' time discount factors is of especial interest. Identification results are presented for a simple two-period/two-player dynamic game of market entry-exit. Under the assumption of level-2 rationality (i.e., players are rational and know that they are rational), a exclusion restriction and a large-support condition on one of the exogenous explanatory variables are sufficient for point identification of all structural parameters.

1. INTRODUCTION

Structural econometric models of individual or firm behavior typically assume that agents are rational in the sense that they maximize expected payoffs given their subjective beliefs about uncertain events. Empirical applications of game-theoretic models have used stronger assumptions than rationality. Most of these studies apply the Nash equilibrium (NE) solution, or some of its refinements, to explain agents' strategic behavior. The NE concept is based on assumptions on players' knowledge and beliefs that are more restrictive than rationality. Although there is no set of necessary conditions for generating the NE outcome, the set of sufficient conditions typically includes the assumption that players' actions are common knowledge. For instance, Aumann and Brandenburger (1995) showed that mutual knowledge of payoff functions and of rationality, along with common knowledge of the conjectures (actions), imply that the conjectures form a NE. But this assumption on players' knowledge and beliefs may be unrealistic in some applications; therefore, it is relevant to study whether the *principle of revealed preference* can identify the parameters in players' payoffs under weaker conditions than NE. For instance, we would like to know whether rationality is sufficient for identification. It is also relevant to study the identification power of other assumptions that are stronger than rationality but weaker than NE, such as common knowledge rationality (e.g., everybody knows that players are rational; everybody knows that everybody knows that players are rational). Common knowledge rationality is

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000114

closely related to the solution concepts *iterated strict dominance* and *rationalizability* (see Fudenberg and Tirole 1991, chap. 2).

The article by Aradillas-Lopez and Tamer (2008) is the first study that deals with these interesting identification issues. The authors study the identification power of rational behavior and rationalizability in three classes of static games that have received significant attention in empirical applications: binary choice games with complete and incomplete information, and auction games with independent private values. Their article contributes to the literature on identification of incomplete econometric models, those models that do not provide unique predictions on the distribution of endogenous variables (see also Tamer 2003; Haile and Tamer 2003). Aradillas-Lopez and Tamer show that standard exclusion restrictions and large-support conditions are sufficient to identify structural parameters despite the nonuniqueness of the model predictions. Although structural parameters can be point identified, the researcher still faces an identification issue when using the estimated model to perform counterfactual experiments. Players' behavior under the counterfactual scenario is not point identified. This problem also appears in models with multiple equilibria. However, a nice feature of Aradillas-Lopez and Tamer's approach is that, at least for the class of models that they consider, it is quite simple to obtain bounds of the model predictions under the counterfactual scenario.

The main purpose of this article is to study the identification power of rational behavior and rationalizability in a class of empirical games that was not analyzed in Aradillas-Lopez and Tamer's article: dynamic discrete games. Dynamic discrete games are of interest in economic applications where agents interact over several periods and make decisions that affect their future payoffs. In static games of incomplete information, players form beliefs on the probability distribution of their opponents' actions. In dynamic games, players also should form beliefs on the probability distribution of players' future actions, including their own future actions, and also on the distribution of future exogenous state variables. The most common equilibrium concept in applications of dynamic games is Markov perfect equilibrium (MPE). As in the case of NE, the concept of MPE is based on strong assumptions of players' knowledge and beliefs. MPE assumes that players maximize expected intertemporal payoffs and have rational expectations, and that players' strategies are common knowledge. In this article, we maintain the assumption that every player knows his own strategy function and has rational expectations on his own future actions; however, we relax the assumption that players' strategies are common knowledge. We study the identification of structural parameters, including players' time discount factors, when we replace the assumption of common knowledge strategies with weaker conditions, such as rational behavior.

We present identification results for a simple two-period/two-player dynamic game of market entry-exit. Under the assumption of level-2 rationalizability (i.e., players are rational and they know that they are rational), a exclusion restriction and a large-support condition on one of the exogenous explanatory variables are sufficient for point identification of all of the structural parameters, including time discount factors.

2. DYNAMIC DISCRETE GAMES

2.1 Model and Basic Assumptions

Assume that two firms decide whether to operate or not in a market. We use the index $i \in \{1, 2\}$ to represent a firm and the index $j \in \{1, 2\}$ to represent its opponent. Time is discrete an indexed by $t \in \{1, 2, \ldots, T\}$, where T is the time horizon. Let $Y_{it} \in \{0, 1\}$ be the indicator of the event "firm i is active in the market at period t." Every period t the two firms decide simultaneously whether or not to be active in the market. A firm makes this decision to maximize its expected and discounted profits $E_t(\sum_{s=0}^{T-t} \delta_i^s \Pi_{i,t+s})$, where $\delta_i \in (0, 1)$ is the firm's discount factor and Π_{it} is its profit at period t. The decision to be active in the market has implications not only on a firm's current profits, but also on its expected future profits. More specifically, there is an entry cost that should be paid only if a currently active firm was not active at previous period. Therefore, a firm's incumbent status (or lagged entry decision) affects current profits. The one-period profit function is

$$\Pi_{it} = \begin{cases} \mathbf{Z}_i \boldsymbol{\eta}_{it} + \gamma_{it} Y_{i,t-1} + \alpha_{it} Y_{jt} - \varepsilon_{it} & \text{if } Y_{it} = 1 \\ 0 & \text{if } Y_{it} = 0, \end{cases} \tag{1}$$

where Y_{jt} represents the opponent's entry decision, $\mathbf{Z}_i$ is a vector of time-invariant exogenous market and firm characteristics that affect firm i's profits, and $\boldsymbol{\eta}_{it}$, γ_{it}, and α_{it} are parameters. The parameter $\gamma_{it} \geq 0$ represents firm i's entry cost at period t. The parameter $\alpha_{it} \leq 0$ captures the competitive effect. At period t, firms know the variables $\{Y_{1,t-1}, Y_{2,t-1}, \mathbf{Z}_1, \mathbf{Z}_2\}$ and the parameters $\{\boldsymbol{\eta}_{1t}, \boldsymbol{\eta}_{2t}, \gamma_{1t}, \gamma_{2t}, \alpha_{1t}, \alpha_{2t}\}$. For the sake of simplicity, we also assume that firms know future values of the parameters $\{\boldsymbol{\eta}, \gamma, \alpha\}$ without any uncertainty. The vector $\boldsymbol{\theta}$ represents the whole sequence of parameters form period 1 to T. The variable ε_{it} is private information of firm i at period t. A firm has uncertainty on the current value of his opponent's ε and also on future values of both his own and his opponent's ε's. The variables ε_{1t} and ε_{2t} are independent of $(\mathbf{Z}_1, \mathbf{Z}_2)$, independent of each other, and independently and identically distributed over time. Their distribution functions, H_1 and H_2, are absolutely continuous and strictly increasing with respect to the Lebesgue measure on $\mathbb{R}$.

2.2 Rational Forward-Looking Behavior

The literature on estimation of dynamic discrete games has applied the concept of MPE. This equilibrium concept assumes that (a) players' strategy functions depend only on payoff-relevant state variables; (b) players are forward-looking, maximize expected intertemporal payoffs, have rational expectations, and know their own strategy functions; and (c) players' strategy functions are common knowledge. The concept of rational behavior that we consider here maintains assumptions (a) and (b) but relaxes condition (c).

Let $\mathbf{X}_t$ be the vector with all of the payoff-relevant and common knowledge state variables at period t: $\mathbf{X}_t \equiv (Y_{i,t-1}, Y_{j,t-1}, \mathbf{Z}_i, \mathbf{Z}_j)$. The information set of player i is $\{\mathbf{X}_t, \varepsilon_{it}\}$. Let $\sigma_{it}(\mathbf{X}_t, \varepsilon_{it})$ be a strategy function for player i at period t. This is a function from the support of $(\mathbf{X}_t, \varepsilon_{it})$ into the binary set $\{0, 1\}$. Associated with any strategy function σ_{it}, we can define a probability function $P_{it}(\mathbf{X}_t)$ that represents the probability of $Y_{it} = 1$

conditional on $\mathbf{X}_t$ and on player i following strategy σ_{it}; that is, $P_{it}(\mathbf{X}_t) \equiv \int I\{\sigma_{it}(\mathbf{X}_t, \varepsilon_{it}) = 1\}\, dH_i(\varepsilon_{it})$, where $I\{\cdot\}$ is the indicator function. It will be convenient to represent players' behavior and beliefs using these *conditional choice probability* (CCP) functions. The CCP function $P_{jt}(\mathbf{X}_t)$ represents firm i's beliefs on the probability that firm j will be active in the market at period t if current state is $\mathbf{X}_t$. Here $\mathbf{P}_j$ represents the sequence of CCPs $\{P_{jt}(\cdot) : t = 1, 2, \ldots, T\}$. Therefore, $\mathbf{P}_j$ contains firm i's beliefs on his opponent's current and future behavior.

A strategy function $\sigma_{it}(\mathbf{X}_t, \varepsilon_{it})$ is *rational* if for every possible value of $(\mathbf{X}_t, \varepsilon_{it})$, the action $\sigma_{it}(\mathbf{X}_t, \varepsilon_{it})$ maximizes player i's expected and discounted sum of current and future payoffs, given his beliefs on the opponent's strategies.

For the rest of the article, we concentrate on a two-period version of this game: $T = 2$. Let $\mathbf{P}_j \equiv \{P_{j1}(\cdot), P_{j2}(\cdot)\}$ be firm i's beliefs on the probabilities that firm j will be active at periods 1 and 2. In the final period, firms play a static game, and the definition of a rational strategy is the same as in a static game. Therefore, $\sigma_{i2}(\mathbf{X}_2, \varepsilon_{i2})$ is a *rational strategy function* for firm i at period 2 if $\sigma_{i2}(\mathbf{X}_2, \varepsilon_{i2}) = I\{\varepsilon_{i2} \leq \Delta_{i2}^{\mathbf{P}_j}(\mathbf{X}_2)\}$, where the threshold function $\Delta_{i2}^{\mathbf{P}_j}(\mathbf{X}_2)$ is the difference between the expected payoff of being in the market and the payoff of not being in the market at period 2, that is,

$$\Delta_{i2}^{\mathbf{P}_j}(\mathbf{X}_2) \equiv \mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}Y_{i1} + \alpha_{i2}P_{j2}(\mathbf{X}_2). \tag{2}$$

Now consider the game at period 1. The strategy function $\sigma_{i1}(\mathbf{X}_1, \varepsilon_{i1})$ is rational if $\sigma_{i1}(\mathbf{X}_1, \varepsilon_{i1}) = I\{\varepsilon_{i1} \leq \Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)\}$, where the threshold function $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ represents the difference between the expected value of firm i if active at period 1 minus its value if not active, given that firm i behaves optimally in the future and that he believes that his opponent's CCP function is $\mathbf{P}_j$, that is,

$$\begin{aligned}\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1) &\equiv \mathbf{Z}_i\boldsymbol{\eta}_{i1} + \gamma_{i1}Y_{i0} + \alpha_{i1}P_{j1}(\mathbf{X}_1)\\ &\quad + \delta_i P_{j1}(\mathbf{X}_1)[V_{i2}^{\mathbf{P}_j}(1,1) - V_{i2}^{\mathbf{P}_j}(0,1)]\\ &\quad + \delta_i(1 - P_{j1}(\mathbf{X}_1))[V_{i2}^{\mathbf{P}_j}(1,0) - V_{i2}^{\mathbf{P}_j}(0,0)], \end{aligned}\tag{3}$$

where $V_{i2}^{\mathbf{P}_j}(\mathbf{X}_2)$ is firm i's value function at period 2 averaged over ε_{i2}, that is, $V_{i2}^{\mathbf{P}_j}(\mathbf{X}_2) \equiv \int \max\{0; \mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}Y_{i1} + \alpha_{i2}P_{j2}(\mathbf{X}_2) - \varepsilon_{i2}\}\, dH_i(\varepsilon_{i2})$. According to this definition of rational strategy function, we say that the CCP functions $P_{i1}(\cdot)$ and $P_{i2}(\cdot)$ are rational for firm i if, given beliefs $\mathbf{P}_j$, we have that

$$P_{it}(\mathbf{X}_t) = H_i(\Delta_{it}^{\mathbf{P}_j}(\mathbf{X}_t)) \quad \text{for } t = 1, 2. \tag{4}$$

In the final period, the game is static and has the same structure as that of Aradillas-Lopez and Tamer (2008); therefore, the derivation of rationalizability bounds on $P_{i2}(\mathbf{X}_2)$ and the conditions for set and point identification of $\{\boldsymbol{\eta}_{i2}, \gamma_{i2}, \alpha_{i2}\}$ are the same as in that article. Section 2.3 discusses two important properties of the threshold functions $\Delta_{it}^{\mathbf{P}_j}(\mathbf{X}_t)$. Section 2.4 derives rationalizability bounds on $P_{i1}(\mathbf{X}_1)$. Section 3 shows how these bounds can be used to identify the parameters $\{\delta_i, \boldsymbol{\eta}_{i1}, \gamma_{i1}, \alpha_{i1}\}$.

2.3 Two Important Properties of the Threshold Functions

The assumption of rationality (or of level-k rationality) implies informative bounds on players' behavior only if the effect of beliefs $\mathbf{P}_j$ on the threshold function $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ is bounded with probability 1. Otherwise, the best-response probability of an arbitrarily pessimistic (optimistic) rational player would be 0 (1) with probability 1. In Aradillas-Lopez and Tamer's static game, this condition holds if the parameters take finite values. In our finite-horizon dynamic model, this condition also is necessary and sufficient. If the parameters $\{\delta_i, \boldsymbol{\eta}_{i1}, \boldsymbol{\eta}_{i2}, \gamma_{i1}, \gamma_{i2}, \alpha_{i1}, \alpha_{i2}\}$ take finite values, then there are two finite constants, c_i^{low} and c_i^{high}, such that for any belief $\mathbf{P}_j$ and any finite value of $\mathbf{X}_1$, the threshold function $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ is bounded by the constants $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1) \in [c_i^{low}, c_i^{high}]$. For an infinite-horizon dynamic game (i.e., $T = \infty$), we also need the discount factor δ_i to be smaller than 1.

The recursive derivation of rationality bounds in Aradillas-Lopez and Tamer's static game is particularly simple because the expected payoff function is strictly monotonic in beliefs $\mathbf{P}_j$. This monotonicity condition is not really needed for identification, but it simplifies the analysis and also likely the estimation procedure. In our two-period game, $\Delta_{i2}^{\mathbf{P}_j}(\mathbf{X}_2)$ is a nonincreasing function of $P_{j2}(\mathbf{X}_t)$ if and only if $\alpha_{i2} \leq 0$. But the monotonicity of $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ with respect to $P_{j1}(\mathbf{X}_1)$ does not follow simply from the restrictions $\alpha_{i1} \leq 0$ and $\alpha_{i2} \leq 0$. Restrictions on other parameters, or on beliefs, are needed to satisfy this monotonicity condition. At period 1, we have that

$$\begin{aligned}\frac{\partial \Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)}{\partial P_{j1}(\mathbf{X}_1)} &= \alpha_{i1} + \delta_i\big(V_{i2}^{\mathbf{P}_j}(1,1) - V_{i2}^{\mathbf{P}_j}(0,1)\\ &\qquad - V_{i2}^{\mathbf{P}_j}(1,0) + V_{i2}^{\mathbf{P}_j}(0,0)\big). \end{aligned}\tag{5}$$

Clearly, $\alpha_{i1} \leq 0$ is not sufficient for $\Delta_{i1}^{\mathbf{P}_j}$ to be a nonincreasing function of $P_{j1}(\mathbf{X}_1)$. We also need the value function $V_{i2}^{\mathbf{P}_j}(Y_{i1}, Y_{j1})$ to be not too supermodular; that is, $V_{i2}^{\mathbf{P}_j}(1,1) - V_{i2}^{\mathbf{P}_j}(0,1) - V_{i2}^{\mathbf{P}_j}(1,0) + V_{i2}^{\mathbf{P}_j}(0,0)$ should be either negative (i.e., $V_{i2}^{\mathbf{P}_j}$ is submodular) or positive but not larger than $-\alpha_{i1}/\delta_i$ (i.e., $V_{i2}^{\mathbf{P}_j}$ is supermodular, but not too much). To derive sufficient conditions, it is important to take into account that $V_{i2}^{\mathbf{P}_j}(\mathbf{X}_2) \equiv G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}Y_{i1} + \alpha_{i2}\ P_{j2}(\mathbf{X}_2))$, where the function $G_i(a)$ is $E_{\varepsilon_i}(\max\{0; a - \varepsilon_i\})$. This function has the following properties: it is continuously differentiable, its first derivative is $H_i(a) \in (0, 1)$, it is convex, $\lim_{a\to-\infty} G_i(a) = 0$, $\lim_{a\to+\infty} G_i(a) - a = 0$, and for any positive constant b, we have that $G_i(a + b) - G_i(a) < b$. There are different sets of sufficient conditions for $\partial\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)/\partial P_{j1}(\mathbf{X}_1) \leq 0$, for instance, a simple set of conditions is $\alpha_{i1} \leq 0$, $\alpha_{i2} \leq 0$, and $\alpha_{i1} - 2\delta_i\alpha_{i2} \leq 0$. Another set of conditions is $\alpha_{i1} \leq 0$ and $\alpha_{i2} \leq 0$; here firm i believes that, ceteris paribus, it is more likely that the opponent's will be active at period 2 if it was active at period 1 [i.e., $P_{j2}(Y_{i1}, 1) \geq P_{j2}(Y_{i1}, 0)$ for $Y_{i1} = 0, 1$], and $(\alpha_{i1} - \delta_i/\alpha_{i2}) \leq 0$. For the rest of the article, we assume that $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ is nonincreasing in $P_{j1}(\mathbf{X}_1)$.

2.4 Bounds With Forward-Looking Rationality

Let $k \in \{0, 1, 2, \ldots\}$ be the index of the level of rationality of both players. We define $P_{it}^{L,k}(\mathbf{X}_t)$ and $P_{it}^{U,k}(\mathbf{X}_t)$ as the lower and upper bounds for player i's CCP at period t under level-k rationality. Level-0 rationality does not impose any restriction, and thus $P_{it}^{L,0}(\mathbf{X}_t) = 0$ and $P_{it}^{U,0}(\mathbf{X}_t) = 1$ for any state $\mathbf{X}_t$. For the last period, $t = 2$, the derivation of the probability bounds is exactly the same as in the static model; therefore, for $k \geq 1$,

$$\begin{aligned} P_{i2}^{L,k}(\mathbf{X}_2) &= H_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}Y_{i1} + \alpha_{i2}P_{j2}^{U,k-1}(\mathbf{X}_2)\big), \\ P_{i2}^{U,k}(\mathbf{X}_2) &= H_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}Y_{i1} + \alpha_{i2}P_{j2}^{L,k-1}(\mathbf{X}_2)\big). \end{aligned} \tag{6}$$

The rest of this section derives a recursive formula for the probability bounds at period 1. Let Π_j^k be the set of player j's CCPs (at periods 1 and 2) that are consistent with level-k rationality. By definition, level-k rationality bounds at period 1 are $P_{i1}^{L,k}(\mathbf{X}_1) = H_i(\Delta_{i1}^{L,k}(\mathbf{X}_1))$ and $P_{i1}^{U,k}(\mathbf{X}_1) = H_i(\Delta_{i1}^{U,k}(\mathbf{X}_1))$, where

$$\begin{aligned} \Delta_{i1}^{L,k}(\mathbf{X}_1) &\equiv \min_{\mathbf{P}_j \in \Pi_j^{k-1}} \{\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)\}, \\ \Delta_{i1}^{U,k}(\mathbf{X}_1) &= \max_{\mathbf{P}_j \in \Pi_j^{k-1}} \{\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)\}. \end{aligned} \tag{7}$$

Given the monotonicity of $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ with respect to $\mathbf{P}_j$, the minimum and the maximum of $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ are reached at the boundaries of the set Π_j^k. More specifically, it is possible to show that the value of (P_{j1}, P_{j2}) that minimizes $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ is

$$\begin{aligned} \big\{&P_{j1}^{U,k-1}(\mathbf{X}_1); P_{j2}^{U,k-1}(1,1); P_{j2}^{U,k-1}(1,0); \\ &P_{j2}^{L,k-1}(0,1); P_{j2}^{L,k-1}(0,0)\big\}; \end{aligned} \tag{8}$$

that is, the most pessimistic belief for firm i (i.e., the one that minimizes $\Delta_{i1}^{\mathbf{P}_j}$) is such that the probability that the opponent is active at period 1 takes its maximum value, and when firm i decides to be active (inactive) at period 1, the probability that the opponent is active at period 2 takes its maximum (minimum) value. Similarly, the value of (P_{j1}, P_{j2}) that maximizes $\Delta_{i1}^{\mathbf{P}_j}(\mathbf{X}_1)$ is

$$\begin{aligned} \big\{&P_{j1}^{L,k-1}(\mathbf{X}_1); P_{j2}^{L,k-1}(1,1); P_{j2}^{L,k-1}(1,0); \\ &P_{j2}^{U,k-1}(0,1); P_{j2}^{U,k-1}(0,0)\big\}. \end{aligned} \tag{9}$$

Firm i's most optimistic belief (i.e., the one that maximizes $\Delta_{i1}^{\mathbf{P}_j}$) is such that the probability that the opponent is active at period 1 takes its minimum value, and when firm i decides to be active (inactive) at period 1, the probability that the opponent is active at period 2 takes its minimum (maximum) value. Therefore, we have the following recursive formulas for the bounds $\Delta_{i1}^{L,k}(\mathbf{X}_1)$ and $\Delta_{i1}^{U,k}(\mathbf{X}_1)$; for $k \geq 1$,

$$\begin{aligned} \Delta_{i1}^{L,k}(\mathbf{X}_1) &= \mathbf{Z}_i\boldsymbol{\eta}_{i1} + \gamma_{i1}Y_{i0} + \alpha_{i1}P_{j1}^{U,k-1}(\mathbf{X}_1) \\ &\quad + \delta_i\big[P_{j1}^{U,k-1}(\mathbf{X}_1)W_{i2}^{L,k}(1) \\ &\qquad + (1 - P_{j1}^{U,k-1}(\mathbf{X}_1))W_{i2}^{L,k}(0)\big], \\ \Delta_{i1}^{U,k}(\mathbf{X}_1) &= \mathbf{Z}_i\boldsymbol{\eta}_{i1} + \gamma_{i1}Y_{i0} + \alpha_{i1}P_{j1}^{L,k-1}(\mathbf{X}_1) \\ &\quad + \delta_i\big[P_{j1}^{L,k-1}(\mathbf{X}_1)W_{i2}^{U,k}(1) \\ &\qquad + (1 - P_{j1}^{L,k-1}(\mathbf{X}_1))W_{i2}^{U,k}(0)\big], \end{aligned} \tag{10}$$

where

$$\begin{aligned} W_{i2}^{L,k}(1) &\equiv G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2} + \alpha_{i2}P_{j2}^{U,k-1}(1,1)\big) \\ &\quad - G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \alpha_{i2}P_{j2}^{L,k-1}(0,1)\big), \\ W_{i2}^{L,k}(0) &\equiv G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2} + \alpha_{i2}P_{j2}^{U,k-1}(1,0)\big) \\ &\quad - G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \alpha_{i2}P_{j2}^{L,k-1}(0,0)\big), \\ W_{i2}^{U,k}(1) &\equiv G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2} + \alpha_{i2}P_{j2}^{L,k-1}(1,1)\big) \\ &\quad - G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \alpha_{i2}P_{j2}^{U,k-1}(0,1)\big), \\ W_{i2}^{U,k}(0) &\equiv G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2} + \alpha_{i2}P_{j2}^{L,k-1}(1,0)\big) \\ &\quad - G_i\big(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \alpha_{i2}P_{j2}^{U,k-1}(0,0)\big). \end{aligned} \tag{11}$$

For instance, for level-1 rationality, we have

$$\begin{aligned} \Delta_{i1}^{L,1}(\mathbf{X}_1) &= \mathbf{Z}_i\boldsymbol{\eta}_{i1} + \gamma_{i1}Y_{i0} + \alpha_{i1} \\ &\quad + \delta_i[G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2} + \alpha_{i2}) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2})], \\ \Delta_{i1}^{U,1}(\mathbf{X}_1) &= \mathbf{Z}_i\boldsymbol{\eta}_{i1} + \gamma_{i1}Y_{i0} \\ &\quad + \delta_i[G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \gamma_{i2}) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2} + \alpha_{i2})]. \end{aligned} \tag{12}$$

An important implication of the monotonicity in $\mathbf{P}_j$ of the threshold function $\Delta_{i1}^{\mathbf{P}_j}$ is that the sequence of lower bounds $\{\Delta_{i1}^{L,k}(\mathbf{X}_1): k \geq 1\}$ is nondecreasing and the sequence of upper bounds $\{\Delta_{i1}^{U,k}(\mathbf{X}_1): k \geq 1\}$ is nonincreasing, that is, for any value of $\mathbf{X}_1$ and any $k \geq 1$,

$$\begin{aligned} \Delta_{i1}^{L,k+1}(\mathbf{X}_1) &\geq \Delta_{i1}^{L,k}(\mathbf{X}_1), \\ \Delta_{i1}^{U,k+1}(\mathbf{X}_1) &\leq \Delta_{i1}^{U,k}(\mathbf{X}_1). \end{aligned} \tag{13}$$

The bounds become sharper when we increase the level of rationality.

3. IDENTIFICATION

Suppose that we have a random sample of many (infinite) independent markets at periods 1 and 2. For each market in the sample, we observe a realization of the variables $\{Y_{i0}, Y_{i1}, Y_{i2}, \mathbf{Z}_i : i = 1, 2\}$. The realizations of the unobservable variables $\{\varepsilon_{it}\}$ are independent across markets. We are interested in using this sample to estimate the vector of structural parameters $\boldsymbol{\theta} \equiv \{\delta_i, \boldsymbol{\eta}_{it}, \gamma_{it}, \alpha_{it} : i = 1, 2; t = 1, 2\}$.

Let $P_{it}^0(\mathbf{X}_t)$ be the true conditional probability function $\Pr(Y_{it} = 1|\mathbf{X}_t)$ in the population, and let $\boldsymbol{\theta}^0$ be the true value of $\boldsymbol{\theta}$ in the population. We consider the following assumptions on the data-generating process for any player $i \in \{1, 2\}$ and any period $t \in \{1, 2\}$:

- (A1) The reduced-form probability $P_{it}^0(\mathbf{X}_t)$ is identified at any point in the support of $\mathbf{X}_t$.
- (A2) The variance–covariance matrix $\text{var}(\mathbf{Z}_i, Y_{i,t-1})$ has full rank.

(A3) The distribution function H_i is known to the researcher.
(A4) $\alpha_{it}^0 \leq 0$, and $\boldsymbol{\theta}^0$ belongs to a compact set Θ.

Assumptions (A1) and (A3) imply that the population threshold function $\Delta_{it}^0(\mathbf{X}_t) \equiv H_i^{-1}(P_{it}^0(\mathbf{X}_t))$ is identified at any point in the support of $\mathbf{X}_t$. We use $\Delta_{it}^0(\mathbf{X}_t)$ instead of $P_{it}^0(\mathbf{X}_t)$ in the following analysis.

Level-k rationality implies the following restrictions on the threshold functions evaluated at the true $\boldsymbol{\theta}^0$:

$$\Delta_{it}^{L,k}(\mathbf{X}_t, \boldsymbol{\theta}^0) \leq \Delta_{it}^0(\mathbf{X}_t) \leq \Delta_{it}^{U,k}(\mathbf{X}_t, \boldsymbol{\theta}^0). \tag{14}$$

Note that, by the monotonicity in k of the rationalizability bounds, if a value of $\boldsymbol{\theta}$ satisfies the restrictions for level-k rationality, then it also satisfies the restrictions for any level k' smaller than k. Let Θ^k be the identified set of parameters for level-k rational players. By definition,

$$\Theta^k = \big\{\boldsymbol{\theta} \in \Theta : \Delta_{it}^{L,k}(\mathbf{X}_t, \boldsymbol{\theta}) \leq \Delta_{it}^0(\mathbf{X}_t) \leq \Delta_{it}^{U,k}(\mathbf{X}_t, \boldsymbol{\theta}) \text{ for any } (i, t, \mathbf{X}_t)\big\}. \tag{15}$$

In the context of dynamic games, the discount factor δ_i is a particularly interesting parameter. Does the identified set Θ^k include the whole interval (0, 1) for the discount factor, or can we rule out some values for that parameter? For instance, can we rule out that players are myopic (i.e., $\delta_i = 0$)? Consider the case of level-1 rationality. Given the restriction $\Delta_{i1}^0(\mathbf{X}_1) \leq \Delta_{i1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)$, and assuming that $\gamma_{i2}^0 - \alpha_{i2}^0 \geq 0$, it is straightforward to show that

$$\delta_i^0 \geq \sup_{\mathbf{X}_1}\left\{\frac{\Delta_{i1}^0(\mathbf{X}_1) - \mathbf{Z}_i\boldsymbol{\eta}_{i1}^0 - \gamma_{i1}^0 Y_{i0}}{G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \alpha_{i2}^0)}\right\}. \tag{16}$$

This expression illustrates several aspects on the identification of δ_i^0. Level-1 rationality implies informative restrictions on the set of parameters, such that Θ^1 does not contain the whole parameter space. In particular, given some values of the other parameters, we can guarantee that the lower bound on δ_i^0 (the right side of the inequality) is strictly positive. Expression (16) also illustrates that we can rule out some values of the discount factor in the interval (0, 1) only if we impose further restrictions, either on the other parameter, or exclusion and support restrictions on the observable explanatory variables.

The rest of the article presents sufficient conditions for point identification of the parameters in $\boldsymbol{\theta}^0$. To prove point identification, we need to establish that for any vector $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0$, there are values of $\mathbf{X}_t$ with positive probability mass such that the inequality $\Delta_{it}^{L,k}(\mathbf{X}_t, \boldsymbol{\theta}) \leq \Delta_{it}^0(\mathbf{X}_t) \leq \Delta_{it}^{U,k}(\mathbf{X}_t, \boldsymbol{\theta})$ does not hold, that is, either $\Delta_{it}^{L,k}(\mathbf{X}_t, \boldsymbol{\theta}) > \Delta_{it}^0(\mathbf{X}_t)$ or $\Delta_{it}^{U,k}(\mathbf{X}_t, \boldsymbol{\theta}) < \Delta_{it}^0(\mathbf{X}_t)$. The following exclusion restriction and large-support assumption is key to the point identification results that we present later:

(A5) There is a variable $Z_{i\ell} \subset \mathbf{Z}_i$ such that $\eta_{i1\ell}^0 \neq 0$, $\eta_{i2\ell}^0 \neq 0$, and conditional on any value of the other variables in $(\mathbf{Z}_i, \mathbf{Z}_j)$, denoted by $\mathbf{Z}_{(-i\ell)}$, the random variable $\{Z_{i\ell}|\mathbf{Z}_{(-i\ell)}\}$ has unbounded support.

Theorem 1 (Point identification under level-1 rationalizability). Suppose that players are level-1 rational and that assumptions (A1)–(A5) hold. Let $\eta_{i1\ell}^0$ and $\eta_{i2\ell}^0$ be the parameters associated with the exclusion restrictions in assumption (A5). Then $\eta_{i1\ell}^0$ and $\eta_{i2\ell}^0$ are point identified.

Proof. For notational simplicity, in this proof we omit the subindex i, but it should be understood that all variables and parameters are player i's. First, we prove the identification of $\eta_{2\ell}^0$. Suppose that $\boldsymbol{\theta}$ is such that $\eta_{2\ell} \neq \eta_{2\ell}^0$. Given $\boldsymbol{\theta}$ and an arbitrary value of $(\mathbf{Z}_{(-\ell)}, Y_1)$, let Z_ℓ^* be the value of Z_ℓ that makes the lower bound function evaluated at $\boldsymbol{\theta}$ equal to the upper bound function evaluated at $\boldsymbol{\theta}^0$, that is, $\Delta_2^{L,1}(Z_\ell^*, \mathbf{Z}_{(-\ell)}, Y_1; \boldsymbol{\theta}) = \Delta_2^{U,1}(Z_\ell^*, \mathbf{Z}_{(-\ell)}, Y_1; \boldsymbol{\theta}^0)$. Given the form of these functions, this value is

$$Z_\ell^* \equiv (\eta_{2\ell} - \eta_{2\ell}^0)^{-1}\big(\mathbf{Z}_{(-\ell)}\big[\boldsymbol{\eta}_{2(-\ell)}^0 - \boldsymbol{\eta}_{2(-\ell)}\big] + Y_1[\gamma_2^0 - \gamma_2] - \alpha_2\big). \tag{17}$$

Here Z_ℓ^* is a finite value that belongs to the support of Z_ℓ. Suppose that $\eta_{2\ell} > \eta_{2\ell}^0$. Then, for values of Z_ℓ greater than Z_ℓ^*, we have that

$$\Delta_2^{L,1}(\mathbf{X}_2, \boldsymbol{\theta}) = \mathbf{Z}\boldsymbol{\eta}_2 + \gamma_2 Y_1 + \alpha_2 > \mathbf{Z}\boldsymbol{\eta}_2^0 + \gamma_2^0 Y_1 = \Delta_2^{U,1}(\mathbf{X}_2, \boldsymbol{\theta}^0), \tag{18}$$

which contradicts the restrictions imposed by level-1 rationality. By assumption (A5), the probability $\Pr(Z_\ell > Z_\ell^*|\mathbf{Z}_{(-\ell)}, Y_1)$ is strictly positive. Because the previous argument can be applied for any possible value of $(\mathbf{Z}_{(-\ell)}, Y_1)$, the result holds with a positive probability mass $\Pr(Z_\ell > Z_\ell^*)$; therefore, we can reject any value of $\eta_{2\ell}$ strictly greater than $\eta_{2\ell}^0$. Similarly, if $\eta_{2\ell} < \eta_{2\ell}^0$, then for values of Z_ℓ smaller than Z_ℓ^*, we have that $\Delta_2^{L,1}(\mathbf{X}_2, \boldsymbol{\theta}) > \Delta_2^{U,1}(\mathbf{X}_2, \boldsymbol{\theta}^0)$. We can reject any value of $\eta_{2\ell}$ strictly smaller than $\eta_{2\ell}^0$. Thus $\eta_{2\ell}^0$ is identified.

Now consider the identification of $\eta_{1\ell}^0$. Note that the proof that follows does not assume that $\eta_{2\ell}^0$ is known. Identification of $\eta_{1\ell}^0$ does not require identification of $\eta_{2\ell}^0$. Given the form of the functions $\Delta_1^{L,1}$ and $\Delta_1^{U,1}$, we have that

$$\begin{aligned}\Delta_1^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}) &- \Delta_1^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)\\ &= \mathbf{Z}(\boldsymbol{\eta}_1 - \boldsymbol{\eta}_1^0) + Y_0(\gamma_1 - \gamma_1^0) + \alpha_1\\ &\quad + \delta\big[G(\mathbf{Z}\boldsymbol{\eta}_2 + \gamma_2 + \alpha_2) - G(\mathbf{Z}\boldsymbol{\eta}_2)\big]\\ &\quad - \delta^0\big[G(\mathbf{Z}\boldsymbol{\eta}_2^0 + \gamma_2^0) - G(\mathbf{Z}\boldsymbol{\eta}_2^0 + \alpha_2^0)\big].\end{aligned} \tag{19}$$

Suppose that $\boldsymbol{\theta}$ is such that $\eta_{1\ell} > \eta_{1\ell}^0$. By the properties of function $G(\cdot)$, the values $\delta[G(\mathbf{Z}\boldsymbol{\eta}_2 + \gamma_2 + \alpha_2) - G(\mathbf{Z}\boldsymbol{\eta}_2)]$ and $\delta^0[G(\mathbf{Z}\boldsymbol{\eta}_2^0 + \gamma_2^0) - G(\mathbf{Z}\boldsymbol{\eta}_2^0 + \alpha_2^0)]$ are bounded within the intervals $[0, \delta(\gamma_2 + \alpha_2)]$ and $[0, \delta^0(\gamma_2^0 - \alpha_2^0)]$. Because the parameter space Θ is a compact set, both $\delta(\gamma_2 + \alpha_2)$ and $\delta^0(\gamma_2^0 - \alpha_2^0)$ clearly are finite values. This implies that for any arbitrary value of $(\mathbf{Z}_{(-\ell)}, Y_1)$, we can always find a finite value of Z_ℓ, say $\bar{Z}_\ell$, such that for $Z_\ell > \bar{Z}_\ell$, we have that $\Delta_1^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}) - \Delta_1^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) > 0$, which contradicts the restrictions imposed by level-1 rationality. By assumption (A5), the probability $\Pr(Z_\ell > \bar{Z}_\ell|\mathbf{Z}_{(-\ell)}, Y_0)$ is strictly positive; therefore, we can reject any value of $\eta_{1\ell}$ strictly greater than $\eta_{1\ell}^0$. We can apply a similar argument to show that we can reject any value of $\eta_{1\ell}$ strictly smaller than $\eta_{1\ell}^0$. Thus $\eta_{1\ell}^0$ is identified.

Point identification of all of the parameters of the model requires at least level-2 rationality. Furthermore, in this dynamic game, at least two additional conditions are needed. First, identifying the discount factor requires that the last period entry cost, γ_{i2}^0, be strictly positive. If this parameter is zero, then the dynamic game becomes static at period 1, and the discount factor does not play any role in the decisions of rational players. Second, the parameters $\eta_{i1\ell}^0$ and $\eta_{i2\ell}^0$ in assumption (A5) should have the same sign.

Theorem 2 (Point identification under level-2 rationalizability). Suppose that assumptions (A1)–(A5) hold, players are level-2 rational, the parameters $\eta_{i1\ell}^0$ and $\eta_{i2\ell}^0$, in assumption (A5) have the same sign, and $\gamma_{i2}^0 > 0$. Then all of the structural parameters in $\boldsymbol{\theta}^0$ are point identified.

Proof. Aradillas-Lopez and Tamer (2008) show that under the conditions of this theorem, all of the parameters in the static game are identified. Therefore, this proof considers that the vector $(\boldsymbol{\eta}_{i2}^0, \gamma_{i2}^0, \alpha_{i2}^0)$ is known, and it concentrates on the identification of $(\delta_i^0, \boldsymbol{\eta}_{i1}^0, \gamma_{i1}^0, \alpha_{i1}^0)$. The proof goes through four cases, which cover all the possible values of $\boldsymbol{\theta} \neq \boldsymbol{\theta}^0$.

Case (a). Suppose that $\boldsymbol{\theta}$ is such that $\eta_{i1\ell} \neq \eta_{i1\ell}^0$. Theorem 1 shows that we can reject this value of $\boldsymbol{\theta}$.

Case (b). Suppose that $\boldsymbol{\theta}$ is such that $\eta_{i1\ell} = \eta_{i1\ell}^0$ but $\boldsymbol{\eta}_{i1(-\ell)} \neq \boldsymbol{\eta}_{i1(-\ell)}^0$ or/and $\gamma_{i1} \neq \gamma_{i1}^0$. We prove here that, given this $\boldsymbol{\theta}$, there is a set of values of $\mathbf{X}_1$, with positive probability mass, such that $\Delta_{i1}^{L,2}(\mathbf{X}_1, \boldsymbol{\theta}) > \Delta_{i2}^{U,2}(\mathbf{X}_1, \boldsymbol{\theta}^0)$, which contradicts the restrictions of level-2 rationality. By definition,

$$\begin{aligned}
\Delta_{i1}^{L,2}(\boldsymbol{\theta}) &- \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \\
&= \mathbf{Z}_i(\boldsymbol{\eta}_{i1} - \boldsymbol{\eta}_{i1}^0) + Y_{i0}(\gamma_{i1} - \gamma_{i1}^0) \\
&\quad + \alpha_{i1} P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) - \alpha_{i1}^0 P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) \\
&\quad + \delta_i\big[P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) W_{i2}^{L,2}(1) \\
&\qquad + (1 - P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta})) W_{i2}^{L,2}(0)\big] \\
&\quad - \delta_i^0\big[P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) W_{i2}^{U,2}(1) \\
&\qquad + (1 - P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)) W_{i2}^{U,2}(0)\big]. \qquad (20)
\end{aligned}$$

Given $\boldsymbol{\theta}$, let $(\mathbf{Z}_{i(-\ell)}, Y_{i0})$ be a vector such that $\mathbf{Z}_{i(-\ell)}(\boldsymbol{\eta}_{i1} - \boldsymbol{\eta}_{i1}^0) + Y_{i0}(\gamma_{i1} - \gamma_{i1}^0) > 0$. By the noncollinearity assumption in (A2) and the exclusion restriction in (A5), for any pair $(Z_{i\ell}, Z_{j\ell})$, the set of values $(\mathbf{Z}_{i(-\ell)}, Y_{i0})$ satisfying the previous inequality has positive probability mass. Now, given the monotonicity of the probabilities $P_{j1}^{L,1}$, $P_{j1}^{U,1}$, $P_{j2}^{L,1}$, and $P_{j2}^{U,1}$ with respect to $Z_{j\ell}$, and given that $\text{sign}(\eta_{j1\ell}^0) = \text{sign}(\eta_{j2\ell}^0)$, we can find values of $Z_{j\ell}$ large enough (or small enough, depending on the sign of the parameter) such that these probabilities are arbitrarily close to 0. That is the case both for the probabilities evaluated at $\boldsymbol{\theta}$ and for those evaluated at $\boldsymbol{\theta}^0$, because in both cases the values of $\eta_{j1\ell}$ and $\eta_{j2\ell}$ are the true ones, $\eta_{j1\ell}^0$ and $\eta_{j2\ell}^0$. Therefore, for these values of $Z_{j\ell}$, we have that

$$\begin{aligned}
\Delta_{i1}^{L,2}(\boldsymbol{\theta}) &- \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \\
&\simeq \mathbf{Z}_{i(-\ell)}(\boldsymbol{\eta}_{i1} - \boldsymbol{\eta}_{i1}^0) + Y_{i0}(\gamma_{i1} - \gamma_{i1}^0) \\
&\quad + (\delta_i - \delta_i^0)[G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0)]. \qquad (21)
\end{aligned}$$

By the definition of the function $G_i(\cdot)$, as $Z_{i\ell}\eta_{i\ell 2}^0$ goes to $-\infty$, both $G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0)$ and $G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0)$ go to 0. Therefore, for these pairs of $(Z_{i\ell}, Z_{j\ell})$, we have that $\Delta_{i1}^{L,2}(\boldsymbol{\theta}) - \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \simeq \mathbf{Z}_{i(-\ell)}(\boldsymbol{\eta}_{i1} - \boldsymbol{\eta}_{i1}^0) + Y_{i0}(\gamma_{i1} - \gamma_{i1}^0) > 0$,which contradicts the restrictions of level-2 rationality. Thus $\boldsymbol{\eta}_{i1(-\ell)}^0$ and γ_{i1}^0 are identified.

Case (c). Suppose that $\boldsymbol{\theta}$ is such that $\boldsymbol{\eta}_{i1} = \boldsymbol{\eta}_{i1}^0$ and $\gamma_{i1} = \gamma_{i1}^0$, but $\alpha_{i1} \neq \alpha_{i1}^0$. Now

$$\begin{aligned}
\Delta_{i1}^{L,2}(\boldsymbol{\theta}) &- \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \\
&= \alpha_{i1} P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) - \alpha_{i1}^0 P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) \\
&\quad + \delta_i\big[P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) W_{i2}^{L,2}(1) \\
&\qquad + (1 - P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta})) W_{i2}^{L,2}(0)\big] \\
&\quad - \delta_i^0\big[P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) W_{i2}^{U,2}(1) \\
&\qquad + (1 - P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)) W_{i2}^{U,2}(0)\big]. \qquad (22)
\end{aligned}$$

Suppose that $\alpha_{i1} > \alpha_{i1}^0$. There are values of $Z_{j\ell}$ large enough (or small enough) such that the probabilities $P_{j1}^{L,1}$, $P_{j1}^{U,1}$, $P_{j2}^{L,1}$, and $P_{j2}^{U,1}$ are arbitrarily close to 1. For these values,

$$\begin{aligned}
\Delta_{i1}^{L,2}(\boldsymbol{\theta}) &- \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \\
&\simeq \alpha_{i1} - \alpha_{i1}^0 + (\delta_i - \delta_i^0) \\
&\quad \times [G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0 + \alpha_{i2}^0) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \alpha_{i2}^0)]. \qquad (23)
\end{aligned}$$

As $Z_{i\ell}\eta_{i\ell 2}^0$ goes to $-\infty$, $G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0 + \alpha_{i2}^0)$ and $G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \alpha_{i2}^0)$ go to 0. Therefore, for these pairs of $(Z_{i\ell}, Z_{j\ell})$, we have that $\Delta_{i1}^{L,2}(\boldsymbol{\theta}) - \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \simeq \alpha_{i1} - \alpha_{i1}^0 > 0$, which contradicts the restrictions of level-2 rationality. Similarly, when $\alpha_{i1} < \alpha_{i1}^0$, we can show that there is a set of values of $\mathbf{X}_1$ with positive probability mass such that $\Delta_{i1}^{U,2}(\mathbf{X}_1, \boldsymbol{\theta}) < \Delta_{i2}^{L,2}(\mathbf{X}_1, \boldsymbol{\theta}^0)$, which also contradicts the restrictions of level-2 rationality. Thus α_{i1}^0 is identified.

Case (d). Suppose that $\boldsymbol{\theta}$ is such that $\boldsymbol{\eta}_{i1} = \boldsymbol{\eta}_{i1}^0$, $\gamma_{i1} = \gamma_{i1}^0$, and $\alpha_{i1} = \alpha_{i1}^0$, but $\delta_i \neq \delta_i^0$. Then

$$\begin{aligned}
\Delta_{i1}^{L,2}(\boldsymbol{\theta}) &- \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \\
&= \alpha_{i1}^0[P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) - P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)] \\
&\quad + \delta_i\big[P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta}) W_{i2}^{L,2}(1) \\
&\qquad + (1 - P_{j1}^{U,1}(\mathbf{X}_1, \boldsymbol{\theta})) W_{i2}^{L,2}(0)\big] \\
&\quad - \delta_i^0\big[P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0) W_{i2}^{U,2}(1) \\
&\qquad + (1 - P_{j1}^{L,1}(\mathbf{X}_1, \boldsymbol{\theta}^0)) W_{i2}^{U,2}(0)\big]. \qquad (24)
\end{aligned}$$

Suppose that $\delta_i > \delta_i^0$. There are values of $Z_{j\ell}$ large enough (or small enough) such that the probabilities $P_{j1}^{L,1}$, $P_{j1}^{U,1}$, $P_{j2}^{L,1}$ and $P_{j2}^{U,1}$ are arbitrarily close to zero. For these values,

$$\Delta_{i1}^{L,2}(\boldsymbol{\theta}) - \Delta_{i2}^{U,2}(\boldsymbol{\theta}^0) \simeq (\delta_i - \delta_i^0)[G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0)] > 0, \quad (25)$$

which contradicts the restrictions of level-2 rationality. Now, consider the difference between $\Delta_{i1}^{U,2}(\boldsymbol{\theta})$ and $\Delta_{i2}^{L,2}(\boldsymbol{\theta}^0)$. We have that

$$\begin{aligned}\Delta_{i1}^{U,2}(\boldsymbol{\theta}) - \Delta_{i2}^{L,2}(\boldsymbol{\theta}^0) &= \alpha_{i1}^0[P_{j1}^{L,1}(\mathbf{X}_1,\boldsymbol{\theta}) - P_{j1}^{U,1}(\mathbf{X}_1,\boldsymbol{\theta}^0)] \\ &\quad + \delta_i\big[P_{j1}^{L,1}(\mathbf{X}_1,\boldsymbol{\theta})W_{i2}^{U,2}(1) \\ &\qquad + (1 - P_{j1}^{L,1}(\mathbf{X}_1,\boldsymbol{\theta}))W_{i2}^{U,2}(0)\big] \\ &\quad - \delta_i^0\big[P_{j1}^{U,1}(\mathbf{X}_1,\boldsymbol{\theta}^0)W_{i2}^{L,2}(1) \\ &\qquad + (1 - P_{j1}^{U,1}(\mathbf{X}_1,\boldsymbol{\theta}^0))W_{i2}^{L,2}(0)\big]. \quad (26)\end{aligned}$$

Suppose that $\delta_i < \delta_i^0$. There are values of $Z_{j\ell}$ large enough (or small enough) such that the probabilities $P_{j1}^{L,1}$, $P_{j1}^{U,1}$, $P_{j2}^{L,1}$, and $P_{j2}^{U,1}$ are arbitrarily close to zero. For these values:

$$\Delta_{i1}^{U,2}(\boldsymbol{\theta}) - \Delta_{i2}^{L,2}(\boldsymbol{\theta}^0) \simeq (\delta_i - \delta_i^0)[G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0 + \gamma_{i2}^0) - G_i(\mathbf{Z}_i\boldsymbol{\eta}_{i2}^0)] < 0, \quad (27)$$

which contradicts the restrictions of level-2 rationality. Thus δ_i^0 is identified.

ACKNOWLEDGMENTS

The author thanks Andres Aradillas-Lopez, Arvind Magesan, Martin Osborne, and Elie Tamer for helpful comments.

ADDITIONAL REFERENCES

Aradillas-Lopez, A., and Tamer, E. (2008), "The Identification Power of Equilibrium in Games," *Journal of Business & Economic Statistics*, in press.
Aumann, R., and Brandenburger, A. (1995), "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161–1180.
Fudenberg, D., and Tirole, J. (1991), *Game Theory*, Cambridge, MA: MIT Press.
Haile, P., and Tamer, E. (2003), "Inference With an Incomplete Model of English Auctions," *Journal of Political Economy*, 111, 1–51.
Tamer, E. (2003), "Incomplete Simultaneous Discrete Response Model With Multiple Equilibria," *Review of Economic Studies*, 70, 147–167.

Comment

Patrick Bajari
Department of Economics, University of Minnesota, Minneapolis, MN 55455 (*bajari@umn.edu*)

In this short note, most of my efforts are directed toward open research questions that exist in the literature, rather than critiquing the particulars of the current article. As a researcher, I am always more interested in the question of where can we possibly go from here rather than dwelling on the limitations of a particular work. I start by describing the contribution of this article to the growing literature on the econometric analysis of games. Next I make some suggestions for some possible, fairly immediate extensions of the current research. Finally, I discuss some general outstanding issues that exist in the literature on estimating games.

Most of my comments are made from the perspective of an applied economist. I view myself as a potential end user of the methods that Aradillas-Lopez and Tamer are creating. I hope that my comments may suggest some directions for extensions that will be useful to econometric theorists, who serve as upstream developers of these exciting new tools.

1. CONTRIBUTION OF THE ARTICLE

This paper is part of a growing literature at the intersection of econometrics and game theory. An important class of models in this literature comprises generalizations of standard discrete choice models, such as the conditional logit, which allow for strategic interactions. In applied microeconomics, and especially empirical industrial organization, we are frequently confronted with problems where the discrete choices of agents are determined simultaneously. For example, starting with Bresnahan and Reiss (1990, 1991), entry in spatially separated markets typically has been modeled as a simultaneous system of discrete choice models. Bresnahan and Reiss noted that entry is naturally modeled as a discrete choice. Economic theory suggests that firms should enter if profits are greater than zero and not enter if profits are less than zero. Therefore, it is natural to include demand shifters, such as the number of consumers in the market and their income and cost shifters, such as the wage rate across markets in the discrete choice model. But in many applied problems, the market structure is concentrated; therefore, the entry decisions of the agents cannot be modeled in isolation. For example, a reasonable model of the decision by Wells Fargo to open a branch in mid-sized city should not be viewed in isolation of Bank of America's decision. To account for this interaction, Bresnahan and Reiss would have us

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000123

include Bank of America's entry decision as a right-side variable in Wells Fargo's utility, and vice versa. Consistent with economic theory, we assume that the observed choices are a Nash equilibrium (NE) to payoffs defined in this manner.

Econometrically, these entry games boil down to a simultaneous system of discrete choice models. Researchers quickly realized that estimating these models was a difficult problem. For a given set of parameters, covariates, and random preference shocks, these models may have multiple NE. Even worse, for certain specifications of the game, there may be no NE in pure strategies. Therefore, these discrete games may generate one, zero, or many predicted outcomes (in pure strategies). Obviously, a straightforward application of maximum likelihood or generalized method of moments is not possible in these settings.

A novel approach to estimating these models was proposed by Tamer (2002), who noted that although discrete entry games do not predict unique equilibrium, the assumption of NE could be used to bound the probability of observing various patterns of entry. Using these bounds, Tamer proposed an estimator that would allow us to estimate the set of parameters that are consistent with the observed frequencies of entry decisions.

This article extends earlier work by Tamer (2002) on the analysis of entry games and by Aradillas-Lopez (2007) on games of private information. But these earlier works make the assumption that the observed behavior is an NE, or Bayes–Nash in the case of private information games. Assuming that agents are playing a NE in applied work is controversial. First, it assumes that agents are able to engage in rational, self-interested behavior. Although most economists agree that this is a useful starting point for our analysis, we also realize that it is an imperfect approximation to the complexities of human behavior. Second, even assuming that agents are rational, the assumption of NE may be controversial. NE assumes that agents act "as if" they are able to make a best response to their common prior about the equilibrium actions. But it is unclear how agents coordinate on which equilibrium to play. It is well known that discrete games can have many NE, particularly as the number of strategies and players grow large. It is difficult to believe that agents are able to formally solve the game from first principals, as in economic theory, and determine which equilibrium is "correct." After all, the leading theorists have not been able to determine a process for selecting a unique equilibrium to a game (and it is possible that they never will). Therefore, it is a strong assumption to assume that agents from first principals have been able to resolve a problem that has escaped solution by our some of our best theorists despite considerable effort.

Aradillas-Lopez and Tamer argue that it may be worthwhile to impose a less-restrictive notion of rationality proposed by Bernheim (1984) and Pearce (1984), which has been referred to as rationalizability in the literature. Level-k rationalizability can be understood as behavior that can be rationalized by some beliefs that survive at least $k-1$ steps of iterated deletion of dominated strategies; for example, level-1 rationalizability imposes the restriction that an agent's actions must be a best response to some set of beliefs. This relaxes the NE assumption that agents act as if they have coordinated on a common set of beliefs about the equilibrium actions and make a set of best responses. Rationalizability is arguably the weakest solution concept that can be imposed on a game. Studying the implications of the weakest modeling assumptions from economic theory seems consistent with the spirit of the bounds literature, where the econometrician takes a very conservative approach to imposing assumptions on the data-generating process. Overall, this is an excellent research question, and the authors deserve much credit for opening this avenue of research.

More broadly, I am very pleased to see Aradillas-Lopez, Tamer, and other young econometricians working on econometric models of strategic interactions. Empirical industrial organization economists typically conceptualize the data-generating process in most markets as being a game. We view the data-generating process in this way because there is a wealth of evidence from previous descriptive work, antitrust cases, and talking with industry participants that agents act strategically. Firms certainly realize that their profits and other outcomes are interdependent and worry a great deal about the actions of their competitors. They may or may not behave in perfect accordance with the hyperrational models of economic theory. But there is something to be said for working with the models that are available, if for no other reason than to take game theory to the data very directly to see whether it works.

It is my hope that the entry by these talented young scholars will bring greater econometric precision and formality to the analysis of strategic interaction by empirical Industrial Organization (IO) economists and applied microeconomists more generally. As the current article demonstrates, there are a wealth of leading questions and models from economic theory that have yet to be explored econometrically. I am very excited to see where this literature will lead over the next decade.

2. POSSIBLE EXTENSIONS

The article is well done and is largely self-contained; however, there is clearly work to be done on extending its contributions. In this section I describe what I believe are "low-hanging fruit" in terms of extensions. I conjecture that some of these extensions have occurred to Aradillas-Lopez and Tamer; nonetheless, I believe that it is useful to include them in my discussion in the hope of stimulating further research.

First, the article cries out for an application. The authors are econometricians, and theories of specialization suggest that their comparative advantage is in proving econometric theorems about the difference between rationalizability and NE at a formal level. However, as a practitioner, I would like to see these methods used with data in a substantive application. Could we ever clearly reject the assumption of rationalizability from Nash behavior for any set of games (not just entry problems)? How does the behavior systematically differ from Nash? Is the difference between Nash and rationalizability of importance for an applied policy or welfare analysis question? In passing, I note that Goldfarb and Yang (2007) have independently made some first steps in this direction and have found evidence suggesting that large markets exhibit "more" rationality in a sense not unlike having a high level-k of rationalizability. But clearly, more work is needed. This question seems like low-hanging fruit for a young, entrepreneurial applied microeconomist.

Second, I would prefer to see an analysis of rationalizability for games other than entry games. As an applied researcher, I am somewhat skeptical that entry can be plausibly modeled as a static decision. Entry typically requires firms to incur fixed costs, and many investments are plausibly modeled as irreversible. Furthermore, the firms may have a learning curve. As a result of these factors, entering firms may not always expect to make profits in the first period. A more reasonable model of their entry decision should involve forward-looking behavior in which firms attempt to forecast their expected discounted profits over their lifetime in a particular market. Finally, entry is almost never simultaneous. There is no date zero at which markets open and firms must decide once and for all whether to enter. More commonly, entry data sets have a panel structure, and we observed the entry (and exit) of firms over a period of years. The one case in which entry may reasonably be modeled as a static, simultaneous problem is in auctions, as discussed by Athey, Levin, and Seira (2004) and Krasnokutskaya and Seim (2007).

This suggests two possible extensions. First, it would be useful to think about games that more reasonably might be taken to data. Second, if we do wish to study entry, perhaps we could extend our analysis of rationalizability to a dynamic setting.

Third, much of the work in the article focuses on games with two players and two strategies. Data sets in substantive applications often have a larger number of players and strategies. For example, Krasnokutskaya and Seim (2007) identified the set of potential entrants as the plan holders to the contracts. There are often many more than two potential plan holders. The methods proposed in the article may require the researcher to analytically characterize the subset of the payoff space in which different actions are rationalizable. This becomes increasingly difficult as the number of players or the cardinality of the choice set increases. For example, for games with 4 players and 2 strategies, the payoff matrix corresponds to a vector with $4 \cdot 2^4 = 64$ elements.

Finding the subspaces, analogous to figure 1 in a 64-dimensional subspace, is a daunting task that likely will be infeasible. To operationalize this procedure in higher dimensions, an alternative approach that exploits numerical methods likely will be required.

Finally, this article, and the bounds estimation literature more generally, needs to develop methods for dealing with unobserved heterogeneity in these models. In our applications, detection of a positive relationship between the actions of two agents could be because we failed to include all of the relevant information about demand, costs, or other payoff relevant variables in our analysis. Our ability to control for these factors is almost always imperfect outside of laboratory experiments. The recent article by Pakes, Porter, Ho, and Ishii (2005) takes important steps in this direction, but clearly more work needs to be done.

3. BROADER ISSUES

More broadly, a skeptical observer might ask: Who cares about the econometric analysis of games? Isn't this just an entirely academic exercise? We know that NE, and probably rationalizability as well, assumes too much rationality on the part of agents. Given that these assumptions strain credibility, who cares about the results?

I think that we care about these models for three reasons. First, in empirical industrial organizations and regulations, we commonly need to conduct applied welfare analysis; for example, an economist at the Federal Trade Commission would like know how much a proposed merger will increase or decrease consumer surplus, and a regulator at the Federal Communications Commission would like to know whether allowing for the portability of cell phone numbers will increase competition and consumer welfare sufficiently to justify the increased cost for the industry. In most markets, it is not plausible to assume that all actors are perfectly rational. On the other hand, as an economist, I do not know of a practical option for measuring welfare other than carefully calculating consumer and producer surplus.

In these applied policy problems, it is tempting to use only qualitative implications from applied theory to inform policy. Frequently, the economists survey the applied IO theory literature and ask what are the qualitative predictions of various theoretical models, and then use only these qualitative predictions.

I believe that structural estimation of these models adds value beyond consulting the theory literature. Forcing the model to confront the data is a learning experience. It is often surprisingly difficult to make a particular structural econometric model derived from an a priori reasonable theory fit the data with reasonable parameter values. Making the data confront the model in this fashion has always helped me to learn about the usefulness of competing theories.

Moreover, if we believe that firms are very irrational, then in principal, we should be able to help them improve profits considerably. Although there is a market for academic economists to consult firms on pricing and other strategic decisions, to be fair, it is a reasonably thin market. In markets with sufficient feedback (i.e., stupid decisions result in the loss of profits and potential exit), I believe that rationality is a reasonable starting point. One lesson of the literature on learning and evolution in games is that if feedback from the economic environment pushes firms in the right direction, then the only thing that behavior can converge to (assuming it converges) is a NE.

A second reason to care about these models is that as matter of positive economics, it is important to force game theory to confront the data. In applied microeconomics, it is common to test the implications of game theory by first deriving a comparative static of the game (e.g., margins should fall if more players enter) and then test this using standard statistical tools, such as regression (e.g., regress margins on the number of firms). Although I believe that this sort of descriptive work is useful, it often is a fairly weak test of the theory.

A far more demanding task of the theory is to find reasonable parameter values that are capable of matching the data. Even if we prove that our model is just identified or only slightly underidentified, this does not mean that we can rationalize the data with parameter values that are within the realm of plausibility. For example, we might find that we can rationalize entry only with profit margins $> 60\%$ and residual demand elasticities near 1. In a highly competitive industry, this would be evidence against even an underidentified model. In substantive applied problems, researchers are frequently surprised to find how well they can fit some aspects of the data and how poorly they can fit others. I believe that this exercise informs us about how we should extend and modify our theories; however, we can learn this only by being bold and taking our models to the data in the most demanding way possible by structurally estimating them.

A third reason why these models are important is because we need them to do counterfactuals. For example, in the merger analysis or number portability questions discussed earlier, at the time the regulators made the decisions, there was an available experiment in which a treatment group and a control group had received the regulation at random.

To make these counterfactual predictions in structural applied work, we use the model to make these extrapolations. Such exercises are by necessity a fairly bold extrapolation; however, the fact that we have bothered to write down a model and carefully estimate its parameters adds intellectual clarity and rigor to this exercise. Antitrust agencies and market regulators, if acting to maximize welfare, must either implicitly or explicitly make assumptions that allow them to forecast the impact of out of sample policy experiments.

Taking that structural models of games are useful to estimate as a given, I believe that it is important that we use these models on important and substantive applications. As academics, we frequently find it too tempting to examine problems of little economic importance but where the data are clean and the identifying assumptions reasonably uncontroversial.

As an applied researcher, I find that the major limitation of the bounds approach is that it cannot simulate the models given the parameter estimates. For a fixed payoff matrix, many games will give us multiple NE, and the bounds literature, by design, sidesteps the question of which equilibrium is played in the data. But without taking a stance on equilibrium selection, I cannot simulate behavior using the model.

As a result, it is not clear how to use the bounds estimator to conduct applied welfare analysis or conduct policy counterfactuals. Moreover, I cannot use the parameter estimates to conduct out-of-sample testing and validation, as done by Todd and Wolpin (2002). This limits my ability as a researcher to test my theories to assess their strong and weak points.

Clearly, these issues are extremely difficult, and a series of investigations by multiple researchers may be needed to find a satisfactory solution. By no means am I the first to raise this concern. However, I hope that upstream developers, such as Aradillas-Lopez and Tamer, along with other young econometricians, may provide some guidance on these practical issues.

ADDITIONAL REFERENCES

Athey, S., Levin, J., and Seira, E. (2004), "Comparing Open and Sealed Bid Auctions: Theory and Evidence From Timber Auctions," working paper, Stanford University.

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

——— (1991), "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81.

Goldfarb, A., and Yang, B. (2007), "Are All Managers Created Equal?" working paper, University of Toronto, Rotman School of Business.

Krasnokutskaya, E., and Seim, K. (2007), "Determinants of the Participation Decision in Highway Procurement Auctions," working paper, University of Pennsylvania, Department of Economics.

Todd, P., and Wolpin, K. (2002), "Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico," working paper, University of Pennsylvania, Department of Economics.

Comment

Han Hong

Department of Economics, Stanford University, Stanford, CA 94022 (*hanhong@stanford.edu*)

This is a very insightful and interesting article by Professors Aradillas-Lopez and Tamer, both experts in the topic of empirical estimation of game-theoretic models. They propose an innovative econometric identification methodology based on the notion of rationalizability of the observed strategic plays by individual agents in the model. They found that general level-k rationalizability places less stringent restrictions on the econometric model than the notation of Nash equilibrium (NE) outcomes. Rationalizability allows for an estimation method that can be used for both point-identified and set-identified models and is easier to implement in most empirical applications than NE-based estimation methods.

They then apply their rationalizability-based estimation method to three classes of games that recently have been extensively used in the empirical literature: bivariate static games of complete information and incomplete information, and first-price sealed-bid auctions. Rationalizability provides a powerful set of analysis tools for each. They also find that the relation between rationalizability and NE generates very different implications for games with complete information versus games with incomplete information. For complete-information games, rationalizability predicts a singleton outcome profile when a unique dominant-strategy NE exists but otherwise makes no prescriptions on the set of observationally nonequivalent outcomes. This is similar to the prediction on the support of strategic plays when mix-strategy equilibrium is allowed for constructing NE predictions.

The implications of rationalizability on incomplete-information games differ substantially. In these games, any NE is contained in all levels of rationalizable outcomes. When a unique NE exists for these games, it coincides with the limit of level-k

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000132

rationalizability iterations as k goes to ∞. When multiple equilibria are present in the game, rationalizability iteration will converge to a strict subset of all of the possible probabilistic predictions of the model that contains all of the multiple equilibria.

The distinction between complete-information games and incomplete-information games for rationalizability identification is particularly enlightening. The difference between these two classes of games is not only theoretically important, but also empirically very relevant. The different mathematical structure between complete- and incomplete-information games also figure prominently in estimation methods based on NE solution concepts. The software Gambit, developed under the direction of McKelvy and McLennan, uses a polynomial homotopy continuation equation solver to find the list of all possible interior and corner solutions to a game. The system of equations that must to be solved in the complete-information case is a system of polynomial equalities and inequalities. Under suitable regularity conditions, the all solution homopoty is known to be able to find all-solutions to such polynomial systems as long as the degree of the initial polynomial on the unit simplex is higher that the degree of the polynomial in the end system of interest. Importantly, polynomial equations are analytic on the complex domain, a property that underlies the validity of the all-solution homopoty method.

As noted by Bajari, Hong, Krainer, and Nekipelov (2006), the system of equations in the private information case depends crucially on the parametric assumptions imposed on the error terms of the utility functions. When the error terms come from the extreme value distribution, the system of equation in the private information case is a system of logistic functions. Such functions are not analytic in the complex domain. The all-solution homopoty method has not yet been shown to find all solutions in such a system of equations. A system polynomial response probability equation may be tried instead of a system of logistic functions in the private information game; however, what joint distribution of the latent utility index can give rise to a system of multinomial polynomial response probabilities is not clear. Alternatively, the degree of the initial unit simplex polynomial can be increased until all of the solutions of the logistic system of equations eventually will be found. But although the total number of real solutions to the logistic system is known to be finite, the total number of complex solutions is not.

Bajari, Hong, and Ryan (2004) also found that simulation-based estimation methods work very differently for games with complete and incomplete information. Complete-information games allow for direct application of the importance sampling scheme of Ackerberg and Gowrisankaran (2001) to separate the equilibrium computation stage and the parameter estimation stage. This is because computation of equilibria in the complete-information model depends only on the total latent utilities, which has a random distribution given each value of the parameters. Changing these parameters simply changes the weights on the parameter distribution during optimization estimation. On the other hand, in incomplete-information games, no randomness buffers the parameter value and the key computation step. A random-coefficient model creates the desired randomness between the parameter value and the computation step, so that changing the parameter values can be used to reweigh the computations in the first stage.

In contrast to the difficulty and complexity of computing all Nash equilibria in both complete-information and incomplete-information games as just described, the innovative econometric methodology of Aradillas-Lopez and Tamer also provides a great deal of computational flexibility. For example, their estimation method for complete-information games requires only applying fast linear programming methods to each point in the parameter set, whereas their estimator for incomplete-information games and first-price auctions requires only iterative calculation of nonparametric conditional expectations. Their methods require neither numerical optimization nor complex solutions of systems of nonlinear equations. Using level-k rationalibility as the driving behavioral assumption in the estimation of games provides significant computational advantages over NE-based solution methods. Using the approach of Aradillas-Lopez and Tamer, a researcher need only to characterize the lower and upper bounds for level-k rationalizable beliefs, which involves straightforward iterative procedures, and need not compute all of the equilibria in the game.

There also might be room for additional improvements in the computational flexibility of their method. In the complete-information game, parameter values in the parameter set must be sampled stochastically to verify whether the rationalizability conditions are satisfied at those parameter values. A Markov chain Monte Carlo approach can potentially improve the speed and the precision of sampling from the set of parameter values.

The authors are to be congratulated for pioneering the rationalizability approach to econometric identification of games, which should open a new exciting arena of both theoretical and empirical research. Many recent works, including those by the authors have addressed the statistical inference properties of set-identified models. A natural venue of continuing research is to investigate the statistical properties of the estimation and identification methods developed in this article. It also may be of potential interest to compare the bounds on the econometric model obtained through the rationalizability condition of this article to some of the more ad hoc methods that have been used previously. For example, in games with multiple equilibria, it is possible under suitable assumptions to rank all of the possible equilibria in a particular order. The model then can be estimated using data under both the "highest" and "lowest" equilibria as a sensitivity analysis. Whether this type of sensitivity analysis can produce meaningful results in the context of the more formal approach of bounding observable outcomes using the notation of rationalizability, as proposed in the article, is a question whose answer can provide useful guidance to many empirical researchers. This is a good example of questions that can be best understood in the context of an application driven by an empirical data set. I look forward to seeing more exciting research from the fruitful collaboration between these two respected scholars.

ACKNOWLEDGMENT

The author acknowledges research support from National Science Foundation grant SES 0721015.

ADDITIONAL REFERENCES

Ackerberg, D., and Gowrisankaran, G. (2001), "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation," working paper, University of California Los Angeles, Department of Economics.

Bajari, P., Hong, H., Krainer, J., and Nekipelov, D. (2006), "Estimating Static Models of Strategic Interactions," working paper, Duke University, Department of Economics, Stanford University, Department of Economics, and University of Minnesota, Department of Economics.

Comment

Shakeeb Khan
Department of Economics, Duke University, Durham, NC (*shakeebk@duke.edu*)

This article attempts to explore the role of imposing equilibrium conditions in models in identifying the structural parameters in certain static models. As the authors point out, equilibrium conditions in noncooperative strategic settings are not necessarily implications of assumptions of rational behavior. Consequently, the identified features of games generally are not robust to the strong assumption of Nash equilibrium (NE) behavior of agents. Instead, the article studies the identification question in simple games under the weaker notion of k-level rationality. It does so by demonstrating how to construct identified regions for parameters of interest under these weaker conditions. These sets generally are larger than those attained under the Nash assumptions, and the article shows how the sets shrink as the level of rationality increases. The article also proposes methods for estimating and conducting inference on these sets and demonstrates how their procedures can be much simpler than finding fixed points, as would be required when imposing equilibrium conditions. All of these ideas are illustrated within the context of three simple games: a discrete game with complete information, a discrete game with incomplete information, and a first-price auction with independent private values.

This article is very novel in many ways and has the potential to move the direction of future research in empirical work into many areas in which applied game theory can be used in modelling, such as industrial organization, international trade, labor, and public finance. In that sense, this article has the potential to be as seminal as the work of Tamer (2003), which was path-breaking in how it merged identification issues in econometrics with equilibrium conditions developed in game theory. Specifically, the article may provide tools to further generalize results that Tamer (2003) spawned, that is, the assumption of equilibrium play can be tested for with data instead of being assumed, because the notion of rationalizability can nest NE as a special case.

The work here suggests some important areas for further research, which can be divided into two categories. The first pertains to the general problem of set identification and inference, which arise not only in this article but also in other important published works. The second deals specifically with some of the new ideas raised in this article on level-k rationality.

An important area that needs to be addressed in this article and other related work (including other work by the same authors) is the notion of relative sharpness of the proposed bounds. In this article and others, sharpness of the bounds is too readily dismissed due to computational concerns. But sharpness is no less important than the notion of efficiency of finite-dimensional parameters in (point identified) semiparametric models, which has (justifiably) spawned an important literature in microeconometrics. Sharpness appears particularly important for the main issues that the authors are attempting to address in this article. Specifically, they are interested in the information lost by relaxing a Nash assumption to level-k rational. A very sharp bound may indicate that very little identification power is lost, whereas an unsharp bound, such as those provided in the article, may indicate that much power is lost.

Having said that, the notion of sharpness may be a very daunting task for models like those studied in this article (and, consequently, a potentially very fruitful exercise). To illustrate how difficult the problem may be, we note that the article itself points out that the inequalities used to construct identification sets follow from the model and its assumptions, not vice versa. This may immediately imply unsharpness of both the sets constructed by these inequalities and sharper sets attained by combining these inequalities. A sharp set would have to be based on all moment inequalities implied by the model.

The problem becomes even more challenging with the presence of covariates in the model, usually resulting in *conditional moment inequalities*, as mentioned by, for example, Khan and Tamer (2007). Analogous to the generalized method-of moments setting, this can imply infinitely many unconditional moment inequalities, each of which would need to be exploited in some way to attain a sharp bound.

Focusing attention on one of the models without covariates, my point on sharpness can be well illustrated in the context of the differing levels of rationality studied so thoroughly in this article. Should the bounds be constructed for a fixed level of rationality k, or is there information that can be exploited by how the bounds change as k varies? The authors themselves suggest this may be the case in their example, where they note that the bounds are indistinguishable for $k \geq 5$ compared with $k = 4$. This is clearly a feature of the assumptions of an exponential distribution, suggesting that there may be information on the rate of shrinkage of the sets as a function of k.

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000097

Another important issue in the context of set inference that arises in the article as well other works is the notion of *uniformity*. I bring this up in the specific context of allowing for point identification, both for the models considered here and those in previous work. In such work (i.e., this article and, e.g., Hong and Tamer 2003; Khan and Tamer 2007), assumptions are stated sequentially in such a way that the identified set eventually reduces to a point under which is usually a strong support condition. Therefore, it will be of interest to test this (strong) condition, conditional on the previous conditions being valid. This is precisely the case where a set inference procedure that is uniform (i.e., does not break down when there is point identification), such as that of Andrews and Soares (2007), can be particularly useful, and I would encourage the authors to formally consider this.

A final comment pertains to the authors' clever identification strategy for the level of rationality k. This is clearly a new innovation that does not exist in the set inference literature. Although not formally suggesting an estimation/testing procedure for k, the article does seem to suggest a sequential inference procedure, for which I would advise caution. After all, this sequential procedure seems loosely analogous to sequential t tests for inferring the lagged order p in AR(p) processes, and such a procedure has a positive probability of overfitting, even in large samples.

In conclusion, I emphasize that this is a very important article that can be and should be extended in many ways. My suggestions are just some of the ways in which the results can be improved on in both empirical and theoretical econometric settings.

ADDITIONAL REFERENCES

Andrews, D. W. K., and Soares, G. (2007), "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," mimeo, Yale University.

Hong, H., and Tamer, E. (2003), "Inference in Censored Models With Endogenous Regressors," *Econometrica*, 71, 905–932.

Khan, S., and Tamer, E. (2007), "Inference in Censored Regression Models Using Conditional Moment Inequalities," *Journal of Economics*, forthcoming.

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

Comment

Thierry MAGNAC
Toulouse School of Economics, Toulouse, France (*magnac@cict.fr*)

The article by Andres Aradillas-Lopez and Elie Tamer clarifies how identification can be obtained in games and provides techniques to do so. In other words, it offers constructive ways of analyzing either point or set identification of parameters governing the preferences and beliefs of players in different games. In this short comment, I highlight the key issues addressed by the article.

First, as sometimes alluded to by the authors, one way of reading their results is to recognize that concepts of Nash equilibrium (NE) or Bayesian NE lead to fixed-point equations. It is then a quite natural consequence to use iterations to arrive at equilibrium, because fixed-point problems are usually solved using a contraction argument and iterating the fixed-point equation an infinite number of times (Rust, Traub, and Wozniakowski 2002). This iteration technique sets two kinds of questions, questions of the interpretation of the iterations and also of the uniqueness of solutions.

Iterations may have structural interpretation. For instance, in dynamic choice models they are viewed as the steps in a real or mental process of backward induction by a decision maker. In this article Aradillas-Lopez and Tamer assume strategically sophisticated players who maximize their expected utility and for whom iterations are interpreted as *level-k rationalizability*. Behavior can be rationalized by some beliefs that can survive at least $k-1$ steps of iteration deletion of dominated strategies. At the limit when k tends to infinity and when this iterated process converges to a unique equilibrium, it is the NE. Nevertheless, iterations need not have structural interpretation. The convergence to a stable equilibrium in a simple demand-and-supply setup using a cobweb iteration technique might not be structural. A standing question regarding the gain from having a structural interpretation remains.

The question of uniqueness of equilibrium is very well described by the three examples worked out in the article. The first example is the simplest, because we arrive at equilibrium after one step. The second example is more sophisticated, because an infinite number of steps is needed before converging to the equilibrium if the equilibrium is unique. If it is not unique, then the limit set is larger than the set of equilibria. The last example in first-price auctions is more involved, because we cannot have convergence to the equilibrium by following Battigalli and Siniscalchi (2003) in their construction of iterations. Thus multiplicity of solutions is a more significant issue than what is obtained under equilibrium assumptions, and the approximation of the strength that an equilibrium hypothesis provides by a set of iterations is less clear.

A second way of reading the article is to examine the relationship, investigated by the authors, between the set of rationalizable strategies and the set of structural parameters (a set that can be a singleton). The tools that the authors use come from the literature dealing with set identification, to which Elie Tamer has contributed a lot (e.g., Tamer 2003; Honoré and

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000141

Tamer 2006). The formal exercise of writing down the transformation between the space of strategies and the space of parameters is not unified in the article yet. To clarify the issue (without taking much care with what these spaces are), I now attempt to discuss it more formally.

Let us first consider the example of a single decision maker. A standard structural model would say that the decision, say y, is a function of observed and exogenous characteristics, say x, of unobserved characteristics ε and of a finite-dimensional parameter β,

$$y = d(x, \varepsilon; \beta). \tag{1}$$

Semiparametric identification involves characterizing the binary relationship between the reduced form given by the probability distribution, $P(y \mid x)$, and the structural form, given by the parameter $\theta = (\beta, F_\varepsilon(\cdot \mid x))$, where $F_\varepsilon(\cdot \mid x)$ is the distribution of ε. The former characterization of the data, $P(y \mid x) = P_{\theta,d}(y \mid x)$, is induced by the latter structural parameter θ and the structural equation (1). The reduced form may have none, one, or many images in the space of structural parameters, in which cases we say that parameters are overidentified, point-identified, or set-identified. Examples of the last instance are not frequent, but are not uncommon. Multicolinearity in linear models or measurement error models (Leamer 1987) provides such instances.

When there are two decision makers, such as in a game, the foregoing setting applies without modification if the economic model $d(\cdot)$ in (1) remains a function. It suffices to reinterpret y as applying to decisions for both or all players. What changes in the games studied by Aradillas-Lopez and Tamer is that the structural model delivers structural *correspondences* instead of functions. We now have that

$$Y = d(x, \varepsilon; \beta), \tag{2}$$

where Y is a set. For instance, in Example 1 using NE in mixed strategies, the "central square" of the (t_1, t_2) space (the authors' ε) yields all solutions, $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. Other examples that the authors study can be framed in this manner using either an equilibrium concept or a rationalizability concept.

The formal discussion may help us derive some general results. For instance, the structural assumption of level-k rationality is weaker than the corresponding one at level $k + 1$, so that correspondences at levels k and $k + 1$ satisfy inclusion properties $Y_k \supseteq Y_{k+1}$. As a consequence, identified sets under successive levels of rationality might be shown to satisfy the same relationship.

Defining the structural model as a correspondence demonstrates that an aspect of what the authors are proposing is not specific to games. Return to the single decision maker case by considering identification in dynamic discrete choice models (Magnac and Thesmar 2002), where we might have a correspondence instead of a function. Probabilities of future events, anticipated by the decision maker, could be ambiguous for instance, because they satisfy inequality restrictions instead of being point-identified. The imaginative setting of Aradillas-Lopez and Tamer can be extended to these general frameworks to analyze set identification in these models.

Furthermore, dealing with (2), in which Y is a set, was the objective of a recent thoughtful article by Galichon and Henry (2007), whose results could be used to establish asymptotic properties of estimators. Specifically, the restrictions used by Aradillas-Lopez in their Examples 1 and 2 are parametric. This is the framework of Galichon and Henry, for which function $d(\cdot)$ is parameterized by a finite vector of parameters, β_1, and where ε follows a distribution indexed by parameters β_2, which also is finite. The use of parametric restrictions might help improve inference with respect to semiparametric approaches (Chernozhukov, Hong, and Tamer 2007).

Nevertheless, the additional twist with games is related to the standing question in the literature regarding dynamic choices about the separate identification of preferences and expectations (Manski 2004). What is interesting in games is that beliefs are a function of preferences and beliefs held by the other agents, so that they are intertwined. It is this property that Aradillas-Lopez and Tamer use, because rationalizability delivers constraints on beliefs. This would not be the case with single decision maker models.

Another useful question for applied researchers is to clarify the normalizations of the parameters needed and the space of parameters in which we can hope to get identification. The difference between a normalization and an identifying restriction might not be clear at first glance. A possible definition would be that normalizations do not affect the family of probability distributions despite a reduction in the parameter space. As before, let $\theta \in \Theta$, the parameter space. A normalization would be defined formally as follows: For any $\theta \in \Theta$, there exists $\theta_0 \in \Theta_0$, where $\Theta_0 \subset \Theta$, with the inclusion being strict, such that $\Pr(y \mid x, \theta) = \Pr(y \mid x, \theta_0)$ almost everywhere (y, x).

For instance, in the simple game with complete information, is the assumption about the discrete nature of the distribution function of (t_1, t_2) a normalization? How should this function be normalized? This case is simple, because we can summarize the problem by the probability weights of the nine regions of the parameter space, which are defined as lying below or above 0 or the parameters α. We thus need only eight unknowns to describe the problem, so that, because α_1 and α_2 are the structural parameters of interest, we need just six unknown probability weights. This is obtained by normalization, that is, there exists a transformation of the parameter space that can be written as a function of six unknowns and α_1 and α_2. I believe that this is a question to be solved before running the linear program (1) proposed by the authors, to avoid complications. The same remark applies to the other examples in the article, where normalization and identification assumptions might be distinguished more clearly.

There are also more specific questions. Some index restrictions are used, although we do not immediately understand their necessity. More deeply, the rôle of unobserved heterogeneity in beliefs and the common prior assumption is clearly high on the research agenda, and the article takes useful steps toward stating sufficient conditions.

In conclusion, this is a nice article because it sets up more questions than it solves, even though it solves many. It is certainly an ambitious research agenda that the authors are proposing for those working in this field.

ADDITIONAL REFERENCES

Chernozhukov, V., Hong, H., and Tamer, E. (2007), "Inference on Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284.
Galichon, A., and Henry, M. (2007), "A Test of Non-Identifying Restrictions and Confidence Regions for Partially Identified Parameters," unpublished manuscript, submitted to *Journal of Economics*.
Leamer, E. E. (1987), "Errors in Variables in Linear Systems," *Econometrica*, 55, 893–909.
Magnac, T., and Thesmar, D. (2002), "Identifying Dynamic Discrete Choice Processes," *Econometrica*, 70, 801–816.
Manski, C. F. (2004), "Measuring Expectations," *Econometrica*, 72, 1329–1376.
Rust, J., Traub, J. F., and Wozniakowski, H. (2002), "Is There a Curse of Dimensionality for Contraction Fixed Points in the Worst Case?" *Econometrica*, 70, 285–329.

Comment

Francesca MOLINARI
Department of Economics, Cornell University, Ithaca, NY 14853 (*fm72@cornell.edu*)

Adam M. ROSEN
Department of Economics, University College London, London WC1E 6BT, U.K. (*adam.rosen@ucl.ac.uk*)

This article discusses how the analysis of Aradillas-Lopez and Tamer (2008) on the identification power of equilibrium in games can be extended to supermodular games. These games embody models that exhibit strategic complementarity, an important and empirically relevant class of economic models. In these games, the extreme points of the Nash equilibrium and rationalizable strategy sets coincide. We discuss how this result facilitates a comparative analysis of the relative identification power of equilibrium and weaker notions of rational behavior. As an illustrative example, we consider a differentiated product oligopoly pricing game in which firms' prices are strategic complements.

1. INTRODUCTION

To date, Nash equilibrium (NE) has been the primary solution concept used in econometric models of games. But the theory of games need not rely on the use of NE or any of its refinements as a solution concept. Indeed, the seminal work of von Neumann and Morgenstern (1944) preceded the introduction of NE (Nash 1950). In later work that introduced rationalizability to the literature, Bernheim (1984) and Pearce (1984) showed that common knowledge of rationality yields a different solution concept than NE—rationalizability—and they argued thus that in some settings, deductive reasoning may not be sufficient to imply Nash behavior.

This point clearly has implications for econometricians who use game-theoretic models to infer agents' payoffs, behavior, and strategic interactions. Thus in settings where it is not clear how players settle on Nash equilibria, and in particular when there are multiple equilibria, it would seem desirable to deduce what inferences can be made robust to the assumption of NE. For this reason, we see the study of the identification power of equilibrium undertaken by Aradillas-Lopez and Tamer (2008; henceforth ALT) as an important step in the econometrics literature. The notion of rationalizability seems a natural one to adopt, in that it is a deductive implication of common knowledge of rationality.

That said, the task of drawing inferences in econometric models of games based only on rationalizability or only finite levels of rationality rather than equilibrium is difficult. Even when equilibrium is used as a solution concept, multiple equilibria are known to significantly complicate identification and inference. Issues raised by multiple equilibria would seem to be only exacerbated by weakening the NE assumption. Given these considerations, we see the constructive identification results provided by ALT in the games that they study as positive findings.

As we elaborate here, the results derived also relate to the literature on supermodular games. The applicability of supermodular games or similar models that exhibit strategic complementarity has been demonstrated in many economic contexts, including models of search (Diamond 1982), technology adoption (Farrell and Saloner 1986; Katz and Shapiro 1986), R&D competition (Reinganum 1981), oil drilling (Katz and Shapiro 1986), and firm production (Beresteanu 2005), as well as in additional references given by Milgrom and Roberts (1990) and Vives (1999, 2005). Such games have additional structure, often commensurate with economic theory, that can be exploited in conjunction with the analysis of both NE and rationalizable strategies, as shown by Milgrom and Roberts (1990). In supermodular games, Milgrom and Roberts (1990) showed that the sets of pure strategy Nash equilibria and rationalizable strategies have identical lower and upper bounds. Moreover, the equilibria of supermodular games exhibit monotone comparative statics (Milgrom and Shannon 1994), for which Echenique and Komunjer (2006) developed a formal test that is robust to multiple equilibria. Although the games considered by ALT are not supermodular, we show here that a transformed version of their entry games exhibits supermodularity, and that this can be exploited to apply the result of Milgrom and Roberts. Then we

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000088

discuss how these results can be used to study the identification power of NE relative to rationalizability or lower levels of rationality in supermodular games more generally, with a differentiated product oligopoly pricing game as our leading illustrative example.

2. THE ENTRY MODEL

Consider the complete information entry game of ALT, but redefine the strategy space as $s_1 = a_1$ and $s_2 = -a_2$. Thus $s_1 = 1$ if firm 1 is present in the market and $s_1 = 0$ otherwise, whereas $s_2 = -1$ if firm 2 is in the market and $s_2 = 0$ otherwise. This simple reformulation of the usual entry game strategy space amounts only to a relabeling of strategies without modifying the game's economic content and is based on the reformulation used by Vives (1999, chap. 2.2.3) in the analysis of Cournot duopoly. Given this reformulation of player 2's strategy, the firms' profit functions are given by

$$\pi_1 = s_1(t_1 - \alpha_1 s_2)$$

and

$$\pi_2 = -s_2(t_2 + \alpha_2 s_1),$$

where t_1 and t_2 are, as in ALT, each players' type. The redefined game is supermodular because (a) players' strategy sets are complete lattices, (b) the profit functions are order-continuous and bounded from above, (c) each profit function π_j is supermodular in s_j for fixed s_{-j}, and (d) each π_j has increasing differences in s_j and s_{-j}. These conditions are sufficient for the game to be supermodular [see, e.g., conditions (A1)–(A4) of Milgrom and Roberts 1990], which guarantees that each firm's best response function is increasing in its rival's strategy. Unfortunately, this reformulation cannot be used to obtain a supermodular representation of the entry game (or the Cournot game) when there are more than two players and entry decisions are strategic substitutes, that is, the α's are negative. When instead entry decisions are complementary (as in, e.g., Borzekowski and Cohen 2005), then the game is supermodular without reparameterization.

Now consider the *incomplete-information* entry game of ALT, with the redefined action space given earlier. The result that the transformed game is supermodular carries through here as well. We follow along the lines of Vives (1999, chap. 2.7.3) to justify this result, adopting similar notation in the context of this Bayesian game. Given the vector of parameters $\alpha = (\alpha_1, \alpha_2)$, let $\pi_i(\mathbf{s}, \mathbf{t}; \alpha)$ be the ex post payoff of player i when the vector of actions are $\mathbf{s} = (s_1, s_2)$ and the realized types are $\mathbf{t} = (t_1, t_2)$. Action spaces, profit functions, type sets, and the prior distribution of types are all common knowledge. A pure strategy for player i is a measurable function that assigns an action to every possible type of the player. Let Σ_i be the strategy space of player i, and identify strategies σ_i and τ_i if they are equal with probability 1. Let $\Pi_i(\sigma_1, \sigma_2) = E[\pi_i(\sigma_1(t_1), \sigma_2(t_2), \mathbf{t}; \alpha)]$ be the expected payoff to player i when player $-i$ uses strategy σ_{-i}, and let $\beta_i : \Sigma_{-i} \to \Sigma_i$ denote player i's best response correspondence. Then a Bayesian NE (BNE) is a strategy profile σ such that $\sigma_i \in \beta_i(\sigma_{-i})$ for $i = 1, 2$. Vives (1999, chap. 2.7.3) defined a natural order in the strategy space as $\sigma_i \le \sigma_i'$ if $\sigma_i(t_i) \le \sigma_i'(t_i)$ componentwise with probability 1. It then follows that the strategy spaces in the Bayesian game are complete lattices, and that $\Pi_i(\sigma_1, \sigma_2)$ is supermodular in σ_i and has increasing differences in (σ_i, σ_{-i}), because these properties of $\pi_i(\cdot, \cdot)$ are preserved under integration.

This shows that once the strategy space is reformulated, both the complete-information and incomplete-information entry games are supermodular. Thus, we can apply Milgrom and Roberts' theorem 5 or, equivalently, theorem 6.1 of Vives (1990), which we repeat here in the notation of the present game for clarity.

Theorem 5 (Milgrom and Roberts 1990). Let Γ be a supermodular game. For each player j, there exist largest and smallest serially undominated strategies, $\bar{s}_j$ and $\underline{s}_j$. Moreover, the strategy profiles $(\underline{s}_1, \underline{s}_2)$ and $(\bar{s}_1, \bar{s}_2)$ are pure NE profiles.

This theorem implies that in supermodular games, both NE and rationalizability yield identical lower and upper bounds on observable strategy profiles. This has two consequences. First, when the NE is unique, it gives the *only* rationalizable strategy profile. In this case the game is dominance-solvable, as iterated elimination of dominated strategies picks out the unique NE. Second, when there are multiple equilibria, the largest and smallest equilibrium strategy profiles also are the largest and smallest rationalizable strategy profiles. In the context of the complete-information entry game, this means that in the region of multiplicity where $(-t_1/\alpha_1, -t_2/\alpha_2) \in [0, 1]^2$, the equilibrium profiles $(s_1, s_2) = (0, -1)$ and $(s_1, s_2) = (1, 0)$ are both rationalizable and also are the smallest and largest such profiles. On the other hand, $(0, 0)$ and $(1, -1)$ also are rationalizable outcomes, but they are not pure strategy equilibria. They may, however, result from a mixed strategy equilibrium, yielding the conclusion that in the region of multiplicity, any outcome is observable when firms play either NE or rationalizable strategies. For the complete-information game, these implications coincide with the analysis of ALT illustrated in their figure 2.

The theorem of Milgrom and Roberts applies to all supermodular games. It does not, however, provide a characterization of feasible outcomes under level-k rationality for finite k. Thus it is not sufficient to provide a characterization of the identifiable features of the model with *level-k* rationality, $k < \infty$. Rather, we must use iteration of levels of rationality to determine the set of feasible outcomes under this restriction, as was done by ALT. Clearly, the model predictions under these two different solution concepts may differ even for $k \to \infty$, unless the game has a unique NE. Nonetheless, the property that the level-k largest and smallest rational strategies shrink to the largest and smallest Nash equilibria as $k \to \infty$ means that imposing level-k rationality should have some identifying power for some finite k, at least in those games where equilibrium restrictions do. In contrast, in games that are not supermodular, it is possible that the bounds on rationalizable outcomes are much wider than those implied by equilibrium behavior. If this is the case, then imposing level-k rationality may not have much identification power. For this reason, we believe that the study of the relative identifying power of equilibrium, rationalizability, and level-k rationality in supermodular games in the same vein as ALT is rather promising.

3. A DIFFERENTIATED PRODUCT PRICING GAME

To illustrate the applicability of this type of analysis in the context of supermodular games, we consider a simple linear duopoly model of price competition in the sale of differentiated products, although similar analysis also can be applied to nonlinear models with more than two firms. There are two firms, denoted $j = 1, 2$, that each produce their own version of a differentiated product in a given market. The firms set their prices simultaneously, and consumers in the market then choose how much of each product to buy. Market demand is assumed to be given by the linear specification for the differentiated products

$$Q_1(\mathbf{p}) = \alpha_1 - \beta_{11}p_1 + \beta_{12}p_2 + u_1 \quad (1)$$

and

$$Q_2(\mathbf{p}) = \alpha_2 - \beta_{22}p_2 + \beta_{21}p_1 + u_2, \quad (2)$$

where p_j denotes the price set for product j, and $Q_j(\mathbf{p})$ denotes the quantity demanded for product j given prices $\mathbf{p} \equiv (p_1, p_2)'$. We assume that the differentiated products are substitutes, and thus impose the following assumption:

Assumption A0. $\beta_{11} > \beta_{12} > 0$, $\beta_{22} > \beta_{21} > 0$.

Assumption A0 guarantees that each product's demand is downward-sloping in its own price and increasing in its competitors price. It also stipulates that demand for each product is more sensitive (in absolute terms) to price changes in its own price than price changes of the competing product. It also guarantees that the dominant diagonal condition is satisfied, which is sufficient for uniqueness of the NE of the pricing game.

Assume that each firm has marginal costs given by

$$mc_j = \gamma_{j0} + \gamma_{j1}w_1 + \gamma_{j2}w_2 + \epsilon_j \equiv \mathbf{w}\boldsymbol{\gamma}_j + \epsilon_j, \quad (3)$$

where w_1 and w_2 are marginal cost shifters, such as factor prices, observable to both the econometrician and the firms, and $\mathbf{w} = [1 \quad w_1 \quad w_2]$ and $\boldsymbol{\gamma}_j = [\gamma_{j0} \quad \gamma_{j1} \quad \gamma_{j2}]'$. (ϵ_1, ϵ_2) are additional firm-specific cost shifters that are assumed to be perfectly observed by the firms but unobserved by the econometrician. We have assumed that firms have the same observable cost shifters, but this is not crucial to the analysis.

Each firm seeks to maximize their profits,

$$\pi_1(\mathbf{p}) = (p_1 - mc_1)Q_1(\mathbf{p})$$

and

$$\pi_2(\mathbf{p}) = (p_2 - mc_2)Q_2(\mathbf{p}).$$

Prices are strategic complements, as each firm's profits are increasing in its rival's price. Using the associated first-order conditions, the firms' reaction functions are given by

$$p_1^*(\tilde{p}_2) = \frac{\alpha_1 + \beta_{12}\tilde{p}_2 + u_1 + \beta_{11}mc_1}{2\beta_{11}} \quad (4)$$

and

$$p_2^*(\tilde{p}_1) = \frac{\alpha_2 + \beta_{21}\tilde{p}_1 + u_2 + \beta_{22}mc_2}{2\beta_{22}}, \quad (5)$$

where $\tilde{p}_2$ ($\tilde{p}_1$) is firm 1's (firm 2's) conjecture for p_2 (p_1). In NE, firms correctly anticipate their rivals' equilibrium prices, which delivers the unique vector (p_1, p_2) such that $p_1 = p_1^*(p_2)$ and $p_2 = p_2^*(p_1)$. If on the other hand, firms are only level-k rational, then they need not correctly anticipate their opponents' strategies; rather, they each will play a best reply to some level-$(k-1)$ rational strategy of their opponent.

Under Assumption A0, theorem 5 of Milgrom and Roberts implies immediately that if the firms play rationalizable strategies (i.e., they are k-rational for all k), then the only rationalizable outcome is the unique NE. If instead one assumes that each firm is only level-k rational for some fixed k, then iterated elimination of dominated strategies can be used in conjunction with (4) and (5) to deliver the set of (p_1, p_2) outcomes consistent with level-k rationality for any parameter vector and realization of unobservables. In similar fashion to the analysis of ALT, this can then be used to formulate a consistent set estimate for model parameters, which in turn can be compared with estimates based on a NE assumption. We illustrate this later. Before proceeding, we make explicit six additional assumptions that we impose throughout:

Assumption A1. The econometrician observes a random sample of observations of (w_1, w_2, p_1, p_2) from a large cross-section of markets drawn from a population distribution that satisfies (1), (2), and (3), as well as Assumptions A0 and A2–A6.

Assumption A2. $E[\epsilon_j] = E[\epsilon_j \cdot w_l] = E[u_j \cdot w_l] = E[u_j] = 0$, $\forall j \in \{1, 2\}$, $\forall l \in \{1, 2\}$.

Assumption A3. $E[(1, w_1, w_2)' \cdot (1, p_1, p_2)]$ has full rank.

Assumption A4. $E[(1, w_1, w_2)' \cdot (1, w_1, w_2)]$ has full rank.

Assumption A5. Firms' strategies are bounded from above, $\exists(\bar{p}_1, \bar{p}_2)$ such that $p_1 \leq \bar{p}_1$ and $p_2 \leq \bar{p}_2$. Both firms' prices are bounded below at 0.

Assumption A6. The support of w_1 and w_2 are subsets of $\mathbb{R}_+$.

Assumption A1 is a typical random sampling assumption. Assumption A2 specifies that econometric unobservables are uncorrelated with cost shifters, a standard exogeneity assumption. Note that it explicitly rules out the special (and somewhat pathological) case where firm j's beliefs about its rival's price p_i depend on w in such a way as to exactly cancel out the effect of variation in w on p_j. Assumption A3 is standard as well; it is the typical rank condition that guarantees that there is sufficient exogenous variation in cost shifters $\mathbf{w}$ to identify the coefficients on prices in the demand specification. In NE, the reasoning used to support this is that variables that change marginal costs, $\mathbf{w}$, result in changes in prices through firms' equilibrium strategies. The same reasoning applies under the weaker rationality assumptions imposed here, because best-reply functions still depend on marginal costs. Assumption A4 is a rank condition that yields identification of the marginal cost parameters when the assumption of NE is imposed. Assumption A5 guarantees that the firms' action space is a lattice, which is necessary to ensure that the game is supermodular. A fixed upper bound on the pricing space can be assumed for simplicity, but more generally this could be a function of observables and model parameters determined endogenously in the model. We return to this issue in Section 3.3.1. Assumption A6 is motivated by the observation that factor prices, which are logical cost shifters, are necessarily nonnegative. This assumption will facilitate the use of the covariance restrictions of Assumption A2 in the construction of moment inequalities by preserving the direction of inequalities when multiplied by w_1 or w_2. If Assumption A6 does not hold, then additional care must be taken when constructing moment inequalities from the bounds implied by level-k rational behavior. Alternatively, a stronger conditional mean restriction on econometric unobservables with respect to exogenous covariates could be invoked.

3.1 Identification of Demand

Assumptions A1–A3 are sufficient conditions for point identification of the demand function parameters, regardless of whether firms play NE (which is equivalent to rationalizable strategies in this game by Milgrom and Roberts 1990, thm. 5) or level-k rational strategies. This is because they uniquely solve the moment conditions

$$
\begin{aligned}
&E\big[(Q_1 - (\alpha_1 - \beta_{11}p_1 + \beta_{12}p_2))w_l\big] = 0, \\
&E\big[(Q_2 - (\alpha_2 - \beta_{22}p_2 + \beta_{21}p_1))w_l\big] = 0,
\end{aligned} \tag{6}
$$

$l = 0, 1, 2$, where $w_0 \equiv 1$.

3.2 Implications of Nash Equilibrium

If firms play NE, then they play the solution to the first-order conditions (4) and (5), which can be written as

$$mc_1 = 2p_1 - \frac{1}{\beta_{11}}(\alpha_1 + \beta_{12}p_2 + u_1)$$

and

$$mc_2 = 2p_2 - \frac{1}{\beta_{22}}(\alpha_2 + \beta_{21}p_1 + u_2).$$

Combining this with the specification of the marginal cost functions, we can isolate the econometric unobservables ϵ_1 and ϵ_2,

$$
\begin{aligned}
\epsilon_1 &= 2p_1 - \frac{1}{\beta_{11}}(\alpha_1 + \beta_{12}p_2 + u_1) - \mathbf{w}\boldsymbol{\gamma}_1, \\
\epsilon_2 &= 2p_2 - \frac{1}{\beta_{22}}(\alpha_2 + \beta_{21}p_1 + u_2) - \mathbf{w}\boldsymbol{\gamma}_2.
\end{aligned} \tag{7}
$$

The orthogonality conditions of Assumption A2 yield a set of moment restrictions that identify the marginal cost parameters $\boldsymbol{\gamma}$ under Assumption A4. These moment conditions can be combined with those of (6) to consistently estimate model parameters by generalized method of moments.

3.3 Levels of Rationality

Now we suppose that instead of playing NE, firms play level-k rational strategies. The moment conditions (6) derived from the demand specification continue to hold, so that the demand parameters, the α's and β's, remain identified. The equality restrictions (7) on marginal costs no longer will be satisfied but can be replaced by moment inequality restrictions derived from level-k rational behavior. To illustrate this, we proceed to derive bounds on observed prices implied by level-k rational behavior following in the same vein as ALT, then show how these bounds can be used to construct moment inequality restrictions. The implied moment inequalities potentially can be used for estimation and inference on marginal cost parameters according to various methods proposed in the recent literature (e.g., Andrews, Berry, and Jia 2004; Andrews and Guggenberger 2007; Beresteanu and Molinari 2008; Chernozhukov, Hong, and Tamer 2007; Galichon and Henry 2006; Pakes, Porter, Ho, and Ishii 2004; Romano and Shaikh 2006; Rosen 2006).

First, suppose that firms are only level-1 rational. Then they will only play undominated strategies. It can be shown that the set of undominated strategies for each firm j are $p_j \in [p^L_{j,1}, p^U_{j,1}]$ where

$$p^L_{j,1} = \max(p^*_j(0), mc_j), \qquad p^U_{j,1} = \min(\bar{p}_j, p^*_j(\bar{p}_{-j})). \tag{8}$$

The rationale for these lower and upper bounds is as follows. Clearly, $p_j < mc_j$ is dominated, because that will yield negative profits. It remains to show that any $p_j < p^*_j(0)$ also is dominated. By definition, $p^*_j(0)$ is a best response to $p_{-j} = 0$. Because the best-response functions are increasing in rival's price, then indeed for any $p_{-j} > 0$, $p^*_j(0) < p^*_j(p_{-j})$. Because the profit function is strictly concave, $p^*_j(0)$ always will give firm j higher profits than any $p_j < p^*_j(0)$, so that any such p_j indeed is dominated by $p^*_j(0)$. Similar reasoning gives that any $p_j > p^*_j(\bar{p}_{-j})$ also is dominated, so the upper bound on level-1 rational strategies is the minimum of $\bar{p}_j$ and $p^*_j(\bar{p}_{-j})$.

With level-1 rational strategies in hand, bounds on firms' prices for higher levels of rationality are defined recursively as

$$
\begin{aligned}
p^L_{j,k} &= \max(p^*_j(p^L_{-j,k-1}), mc_j), \\
p^U_{j,k} &= \min(\bar{p}_j, p^*_j(p^U_{-j,k-1})).
\end{aligned} \tag{9}
$$

These bounds apply because a level-k rational firm always will play a best reply to a feasible level-$(k-1)$ strategy for its opponent. This is precisely the reasoning of ALT (Claim 1) applied in the setting of the Bertrand pricing game. Because each best-reply function is increasing in rival's strategy, the interval bounded by $p^*_j(p^L_{-j,k-1})$ and $p^*_j(p^U_{-j,k-1})$ is precisely the set of best replies to level-$(k-1)$ rational strategies. Note that Assumption A5 plays a key role in that it delivers a nontrivial upper bound on level-k rational behavior by bounding the effect of a rival's action on player j's payoff. In a nonlinear model, this potentially could be achieved without a bounded strategy space.

To see how these inequality restrictions on prices from level-k rationality have identifying power, note that $p^L_{j,k}$ and $p^U_{j,k}$ are functions of model parameters, data, and econometric unobservables. Following ALT, this implies, for any level k, a set of moment inequality restrictions on model parameters. Next we illustrate the set of implied inequalities for the cases where $k = 1$ and $k = 2$.

3.3.1 Moment Inequalities From Level-1 Rationality. Here we derive the inequality restrictions implied by level-k rational behavior for firm 1, for $k = 1$. The bounds for firm 2 have a symmetric derivation and thus are omitted. Level-1 rationality implies that the observed price p_1 must lie within the interval of firm 1's undominated strategies, that is, the bounds (8). The restriction that $p^L_{1,1} \le p_1$ is equivalent to the two inequality restrictions

$$p^*_1(0) \le p_1 \qquad \text{and} \qquad mc_1 \le p_1,$$

whereas the restriction from the upper bound, $p_1 \le p^U_{1,1}$, is given by the two inequalities

$$p_1 \le \bar{p}_1 \qquad \text{and} \qquad p_1 \le p_1^*(\bar{p}_2).$$

Multiplying these inequality restrictions by each component of $\mathbf{w}$, applying the definitions of $p_1^*(\cdot)$ and mc_1, and then taking expectations yields the following set of moment inequalities for $l = 0, 1, 2$, where $w_0 \equiv 1$:

$$E[w_l(2\beta_{11}p_1 - \alpha_1 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1)] \ge 0,$$
$$E[w_l(p_1 - \mathbf{w}\boldsymbol{\gamma}_1)] \ge 0,$$
$$E[w_l(\bar{p}_1 - p_1)] \ge 0,$$

and

$$E[w_l(\alpha_1 + \beta_{12}\bar{p}_2 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1 - 2\beta_{11}p_1)] \ge 0.$$

The third inequality allows for the case where $\bar{p}_1$ is a function of observables and model parameters determined endogenously in the model. If $\bar{p}_1$ is assumed to be exogenously given, then it does not depend on model parameters, and this condition will not prove informative and thus can be dropped.

Similar inequality restrictions also can be derived for firm 2. These moment inequalities then can be used for estimation and inference.

3.3.2 Moment Inequalities From Level-2 Rationality. When $k = 2$, each firm will play best responses to some level-1 rational strategies of their opponent. Using the level-k rational bounds of (9), we have the following lower bounds for p_1:

$$p_1^*(p^L_{2,1}) \le p_1 \qquad \text{and} \qquad mc_1 \le p_1,$$

where $p^L_{2,1}$ is the lower bound on firm 2's level-1 rational strategies. This is given by $p^L_{2,1} = \max\{p_2^*(0), mc_2\}$. Because $p_1^*(\cdot)$ is monotone, it follows that

$$p_1^*(p_2^*(0)) \le p_1, \qquad p_1^*(mc_2) \le p_1, \qquad \text{and} \qquad mc_1 \le p_1. \tag{10}$$

Upper bounds on p_1 from (9) are

$$p_1 \le \bar{p}_1 \qquad \text{and} \qquad p_1 \le p_1^*(p^U_{2,1}).$$

Because $p^U_{2,1} = \min\{\bar{p}_2, p_2^*(\bar{p}_1)\}$, these restrictions are equivalent to

$$p_1 \le \bar{p}_1, \qquad p_1 \le p_1^*(\bar{p}_2), \qquad \text{and} \qquad p_1 \le p_1^*(p_2^*(\bar{p}_1)). \tag{11}$$

Multiplying both sides of all inequalities of (10) and (11) by each component of $\mathbf{w}$ and taking expectations gives, for $l = 0, 1, 2$,

$$E\big[w_l\big(4\beta_{22}\beta_{11}p_1 - 2\beta_{22}[\alpha_1 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1] - \beta_{12}[\alpha_2 + \beta_{22}\mathbf{w}\boldsymbol{\gamma}_2]\big)\big] \ge 0,$$
$$E[w_l(2\beta_{11}p_1 - \alpha_1 + \beta_{12}\mathbf{w}\boldsymbol{\gamma}_2 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1)] \ge 0,$$
$$E[w_l(p_1 - \mathbf{w}\boldsymbol{\gamma}_1)] \ge 0,$$
$$E[w_l(\bar{p}_1 - p_1)] \ge 0,$$
$$E[w_l(\alpha_1 + \beta_{12}\bar{p}_2 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1 - 2\beta_{11}p_1)] \ge 0,$$

and

$$E\big[w_l\big(2\beta_{22}[\alpha_1 + \beta_{11}\mathbf{w}\boldsymbol{\gamma}_1] + \beta_{12}[\alpha_2 + \beta_{21}\bar{p}_1 + \beta_{22}\mathbf{w}\boldsymbol{\gamma}_2] - 4\beta_{11}\beta_{22}p_1\big)\big] \ge 0.$$

Level-2 rationality for firm 2 yields a symmetric set of moment inequalities. As was the case for level-1 rationality, the moment inequalities can be used to provide a basis for estimation and inference.

These moment equalities appear to be more notationally cumbersome than those for level-1 rationality. For this reason, it is worth restating that the demand parameters (i.e., the α's and β's) are identified separately from the moment equalities (6). The foregoing moment inequalities are useful for their restrictions on the marginal cost parameters (the γ's), and the moment functions on the left side of the inequalities are linear in these. Similar analysis could be applied to nonlinear models as well, although of course the moment functions no longer would be linear.

3.4 Extension to Differentiated Products With Incomplete Information

In the previous section, we assumed that each firm j perfectly observes its rival's type, that is, their rival's unobservable cost-shifter ϵ_{-j}. This might not be the case in practice, resulting in a game of incomplete information. Our analysis can also be applied to the differentiated product model with incomplete information. By the same argument as in Section 2, the incomplete-information version of this game also is supermodular. Each firm maximizes expected profit conditional on its own information set and their (rationalizable) beliefs. Suppose that each firm j observes ϵ_j but not ϵ_{-j}. Let $G(\epsilon_1, \epsilon_2)$ be the joint distribution of firm 1 and firm 2 types, and assume that this distribution is common knowledge. Then firm 1, for example, solves

$$\max_{p_1} \int [(p_1 - mc_1)Q_1(p_1, \tilde{p}_2(\epsilon_2))]\, dG_1(\epsilon_2|\epsilon_1),$$

where $\tilde{p}_2(\epsilon_2)$ is firm 1's conjecture for firm 2's price as a function of ϵ_2. If firm 1 is level-1 rational, then $\tilde{p}(\cdot)$ simply maps firm 2's strategy space. If firm 1 is level-2 rational, then the range of $\tilde{p}_2(\cdot)$ lies in the set of firm 2's strategies that are level-1 rational (i.e., undominated), because firm 1 knows that firm 2 will not play a dominated strategy. If the firms are playing rationalizable strategies, then they are level-∞ rational (common knowledge of rationality), and the range of $\tilde{p}_2(\cdot)$ is a subset of firm 2's rationalizable strategies. In this case, due to the result of Milgrom and Roberts, the extreme points of the set of rationalizable strategies are pure strategy BNE.

Maximizing expected profit, we obtain that the best replies are

$$p_1(\epsilon_1) = \frac{\alpha_1 + \beta_{12}E_{G_1}(p_2(\epsilon_2)|\epsilon_1) + u_1 + \beta_{11}mc_1}{2\beta_{11}}$$

and

$$p_2(\epsilon_2) = \frac{\alpha_2 + \beta_{21}E_{G_2}(p_1(\epsilon_1)|\epsilon_2) + u_2 + \beta_{22}mc_2}{2\beta_{22}}.$$

In some special cases, including those where types are independent or where $G(\epsilon_1, \epsilon_2)$ is bivariate normal, the BNE is unique (Vives 1999, chap. 8.1.2). Thus, by theorem 5 of Milgrom and

Roberts (1990), the set of rationalizable strategies is given by the unique BNE. The same analysis as in the previous section then can be conducted, comparing the conclusions drawn under the assumption of Nash behavior versus level-k rationality.

Irrespective of whether the set of BNE is a singleton, the analysis of ALT applies. In particular, if firm j is level-1 rational, then $p_j \in [p_{j,1}^L, p_{j,1}^U]$, where $p_{j,1}^L$ and $p_{j,1}^U$ are as defined in (8). If firm j is level-2 rational, then $G_j(\cdot|\epsilon_j)$ assigns zero probability to $p_{-j} \notin [p_{-j,1}^L, p_{-j,1}^U]$. An iterative argument can be applied for level-k rationality.

4. CONCLUSION

In the familiar context of models of entry and auctions, ALT have shown how in empirical work the assumption of NE can be weakened to rationalizability or even lower levels of rationality. Specifically, they have shown how these weaker solution concepts have identifying power that can be used for estimation and inference in these models. These findings are appealing in that the weaker solution concepts are direct implications of various levels of agents' rationality, whereas common knowledge of rationality is not sufficient to guarantee Nash behavior.

We have shown how their insight can be fruitfully applied to the analysis of supermodular games. These games are useful for modeling settings with strategic complementarity, which are an important and empirically relevant class of economic models, well studied in microeconomic theory. Specifically, Milgrom and Roberts (1990) and Vives (1990) have shown that in supermodular games, the largest and smallest rationalizable strategies coincide with the largest and smallest Nash equilibria. This relationship between the two solution concepts implies that considering the implications of level-k rational behavior for increasing levels of k, the bounds on rational strategies will shrink to those of the Nash set. Although the full sets of model predictions based on the two solution concepts may still differ when the equilibrium is not unique, this relationship between the two sets is useful for characterizing their relative identification power. For this reason, supermodular games provide appealing grounds on which to compare the identification power of various levels of rationality relative to that of an equilibrium assumption. To demonstrate the applicability of such analysis, we consider an oligopoly pricing game. We illustrate how the assumption of level-k rational behavior can be used to deliver moment equalities and inequalities, which then can be used for estimation and inference, where some parameters are possibly set-identified rather than point-identified.

ACKNOWLEDGMENTS

The authors thank Arie Beresteanu and participants at the *2007 Joint Statistical Meetings* session on "Inference in Simple Dynamic Games With Multiple Equilibria" for comments and discussion. They also thank the editor, Serena Ng, and a reviewer for their comments. All remaining errors are the authors' own. Molinari gratefully acknowledges financial support from National Science Foundation grant SES-0617482. Rosen gratefully acknowledges financial support from the Economic and Social Research Council through its Centre for Microdata Methods and Practice grant RES-589-28-0001.

ADDITIONAL REFERENCES

Andrews, D. W. K., and Guggenberger, P. (2007), "Validity of Subsampling and Plug-in Asymptotic Inference for Parameters Defined by Moment Inequalities," working paper, Cowles Foundation, Yale University.

Andrews, D. W. K., Berry, S., and Jia, P. (2004), "Confidence Regions for Parameters in Discrete Games With Multiple Equilibria, With an Application to Discount Chain Store Location," working paper, Cowles Foundation, Yale University.

Aradillas-Lopez, A., and Tamer, E. (2008), "The Identification Power of Equilibrium Simple in Games," *Journal of Business Economic Statistics*, 26, 261–283.

Beresteanu, A. (2005), "Nonparametric Analysis of Cost Complementarities in the Telecommunications Industry," *Rand Journal of Economics*, 36, 870–889.

Beresteanu, A., and Molinari, F. (2008), "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, To appear.

Bernheim, B. D. (1984), "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007–1028.

Borzekowski, R., and Cohen, A. M. (2005), "Estimating Strategic Complementarities in Credit Unions' Outsourcing Decisions," working paper, Federal Reserve Board of Governors.

Chernozhukov, V., Hong, H., and Tamer, E. (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284.

Diamond, P. (1982), "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 90, 881–894.

Echenique, F., and Komunjer, I. (2006), "A Test for Monotone Comparative Statics," working paper, UCSD, Department of Economics and Caltech Division of the Humanities and Social Sciences.

Farrell, J., and Saloner, G. (1986), "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation," *American Economic Review*, 76, 940–955.

Galichon, A., and Henry, M. (2006), "Inference in Incomplete Models," working paper, Columbia University, Department of Economics.

Katz, M., and Shapiro, C. (1986), "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94, 822–841.

Milgrom, P., and Roberts, J. (1990), "Rationalizability, Learning, and Equilibrium in Games With Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Milgrom, P., and Shannon, C. (1994), "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Nash, J. (1950), "Equilibrium Points in N–Person Gamers," *Proceedings of the National Academy of Sciences*, 36, 48–49.

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2004), "The Method of Moments With Inequality Constraints," working paper, Harvard University, Department of Economics.

Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1007–1028.

Reinganum, J. (1981), "Dynamic Games of Innovation," *Journal of Economic Theory*, 25, 21–41.

Romano, J. P., and Shaikh, A. M. (2006), "Inference for the Identified Set in Partially Identified Econometric Models," working paper, Stanford University, Department of Economics.

Rosen, A. (2006), "Confidence Sets for Partially Identified Parameters That Satisfy a Finite Number of Moment Inequalities," Working Paper CWP25/06, Cemmap.

Vives, X. (1990), "Nash Equilibrium With Strategic Complementarities," *Journal of Mathematical Economics*, 19, 305–321.

——— (1999), *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press.

——— (2005), "Complementarities and Games: New Developments," *Journal of Economic Literature*, 43, 437–479.

von Neumann, J., and Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.

Comment

Allan COLLARD-WEXLER
Department of Economics, Stern School of Business, New York University, New York, NY 10012
(*wexler@nyu.edu*)

1. A SIMPLE ARADILLAS–LOPEZ AND TAMER ESTIMATOR

To illustrate the power and ease of Aradillas-Lopez and Tamer's approach, I estimate a simple entry model in the spirit of Bresnahan and Reiss (1991), using only the restrictions that players use rationalizable strategies, on data from the ready-mix concrete industry. Although this empirical exercise is fairly stripped down, it can be adapted for greater realism, such as allowing for different types of entrants or correlation in the unobserved component of firm profits. In the second section, I discuss the realism of Nash equilibrium (NE) in applied work.

1.1 Data

I use data on entry patterns of ready-mix concrete manufacturers in isolated towns across the United States. In previous work (e.g., Collard-Wexler 2006), I have studied entry patterns in the ready-mix concrete market. Concrete is a material that cannot be transported for much more than an hour, and thus it makes sense to study entry in local markets. I construct "isolated markets" by selecting all cities in the United States that are at least 20 miles away from any other city of at least 2,000 inhabitants. I then count the number of ready-mix concrete establishments in the U.S. Census Bureau's Zip Business Patterns for zip codes at most 5 miles away from the town. A description of the Zip Business Patterns data set is available at *http://www.census.gov/epcd/www/zbp_base.html* (accessed September 1, 2007). More information on the construction of data on isolated towns as well as the set of towns and zip codes used to construct the data set used in this discussion is available at *http://pages.stern.nyu.edu/∼acollard/Data%20Sets%20and%20Code.html*.

I designate the number of ready-mix concrete firms in a market as N_j and the number of potential entrants in a market as N^e. I set the number of potential entrants to six, the maximum number of firms in any market in the data. I estimate the probability of entry using nonparametric regression,

$$\hat{P}(x_i) = \frac{1}{N^e} \sum_{x_j \neq x_i} N_j K\left(\frac{x_j - x_i}{h}\right). \tag{1}$$

I use a normal density as a kernel $K(\cdot)$ and choose smoothing parameter $h = .43$ to minimize the sum of squared errors from the regression. Figure 1 presents the number of ready-mix concrete establishments in a town plotted against town population, along with a nonparametric regression of this relationship where $\hat{N}(x_i) = N^e \hat{P}(x_i)$.

1.2 Estimator

I use a entry model similar to that discussed by Aradillas-Lopez and Tamer. Firms are ex ante identical but receive different private information shocks to the profits that they will receive on entry. I parameterize a firm's profits as

$$\pi_i = \beta \underbrace{x_i}_{\text{Log population}} + \alpha \underbrace{N_i}_{\text{Number of entrants}} + \underbrace{\epsilon_i}_{\text{Private information shock}},$$

where β measures the effect of population on profits and α is the effect of an additional competitor on profits. Initially, the highest prior that I can assign to the entry probability of my opponents is that all opponents enter, that is, $\bar{e}(x_i)^0 = 1$. Likewise, the lowest possible prior that I can have is that no other firms enter the market, that is, $\underline{e}(x_i)^0 = 0$. Given these upper and lower bounds on beliefs, a firm chooses to enter if it makes positive profits. Aradillas-Lopez and Tamer define level-k rationality or level-k rationalizability as behavior that can be rationalized by some beliefs that survive at least $k-1$ steps of iterated deletion of dominated strategies. Thus the first stage of this process of iterated deletion of dominated strategies assigns possible entry probabilities at 0 and 1. Bounds on a firm's expected profits π_i^k (where k denotes the level-k of rationality) given the assumption that effect of additional firms is to decrease profits are

$$x_i\beta + \alpha N^e \bar{e}(x_i)^0 + \epsilon_i \leq \pi_i^0 \leq x_i\beta + \alpha N^e \underline{e}(x_i)^0 + \epsilon_i.$$

The bounds on the probability that a firm will enter given $K = 0$, denoted as $e_i(\theta)$, follow directly,

$$F_\epsilon(x_i\beta + \alpha N^e \bar{e}(x_i)^0) \leq e_i(\theta) \leq F_\epsilon(x_i\beta + \alpha N^e \underline{e}(x_i)^0),$$

where F_ϵ is the cdf of ϵ.

From here, it is straightforward to iterate on the upper and lower bounds for entry probabilities for levels of rationality greater than $K = 0$. This process of iteration is equivalent to deleting dominated strategies. The upper bound on the entry probability for a firm is denoted by $\bar{e}(x_i)^k$, and the lower bound is denoted by $\underline{e}(x_i)^k$; these are given recursively by

$$\bar{e}(x_i)^{k+1} = F_\epsilon(x_i\beta + \alpha N^e \underline{e}(x_i)^k) \tag{2}$$

and

$$\underline{e}(x_i)^{k+1} = F_\epsilon(x_i\beta + \alpha N^e \bar{e}(x_i)^k). \tag{3}$$

In my application I just assume that ϵ has a standard normal distribution, that is, $\epsilon \sim \mathrm{N}(0, 1)$.

Aradillas-Lopez and Tamer define the identified set for level-k rationality as the set of parameter values that satisfy the level-k' conditional moment inequalities for each $k \geq k' \geq 1$ with

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000150

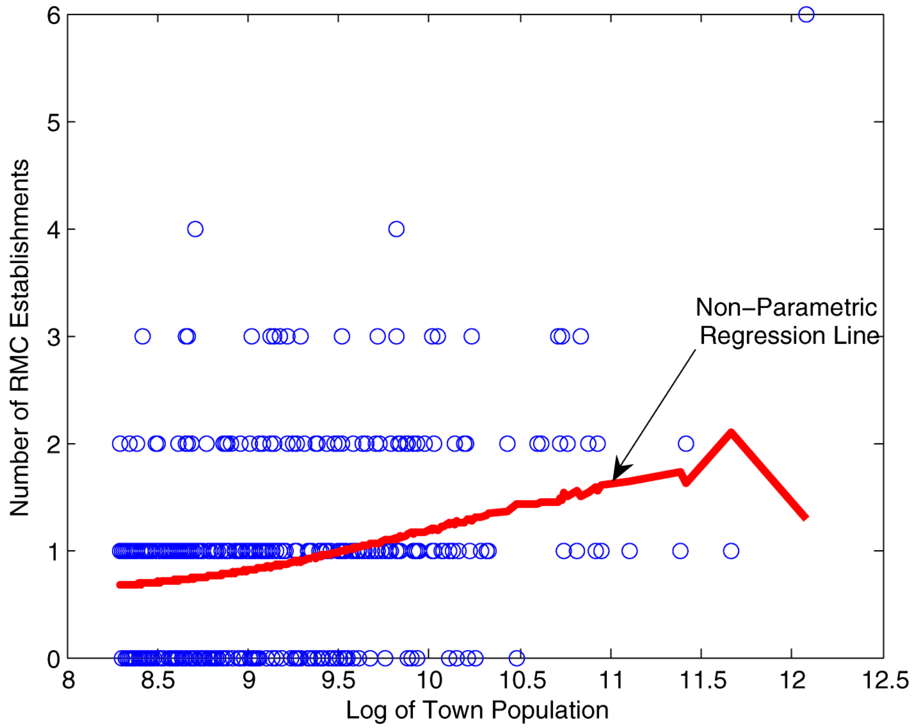

Figure 1. Entry patterns of ready-mix concrete plants in isolated markets.

probability 1. Thus a natural estimator for this model can be derived from looking for cases when the entry probabilities in the data are outside of the upper and lower bounds. The criterion for one such estimator is

$$Q^k(\theta) = \sum_i \left([\hat{P}(x_i) - \bar{e}^k(\theta, x_i)]^+\right)^2 + \left([\underline{e}^k(\theta, x_i) - \hat{P}(x_i)]^+\right)^2. \tag{4}$$

The identified set is just the set of parameters θ for which there are no violations of the upper and lower bounds,

$$\hat{\Theta}^I = \{\theta \in \Theta : Q^k(\theta) = 0\}. \tag{5}$$

In contrast, the standard estimator using a symmetric NE, such as the model of Seim (2005), would look for an entry probability e^* that is a fixed point to the best-response mapping, that is, e^* such that

$$e^* = F_\epsilon(x_i\beta + \alpha N^e e^*).$$

This would lead to an estimator with the following criterion function, the distance between the symmetric Nash solution and the data:

$$Q^N(\theta) = \sum_i [\hat{P}(x_i) - e^*(\theta, x_i)]^2. \tag{6}$$

Note that the estimator that minimizes the Nash criterion in equation (6) generally will be point-identified. The estimated parameter for the Nash criterion is $\hat{\theta}^N = \text{argmin}_\theta\, Q^N(\theta)$, the (generically) unique parameter that minimizes the deviations of the Nash prediction from the data.

Figure 2 presents the prediction of both the rationalizable model for up to 100 levels of iterated deletion of dominated strategies and the NE model for parameters $\alpha = -.42$ and $\beta = .1$. The top lines show the upper bound on the number of firms that enter, corresponding to the smallest possible belief about the number of other firms that might enter. Likewise, the bottom lines correspond to the lower bound on the number of expected entrants if I held the greatest belief about the entry probability of opponents. The middle dashed line shows the prediction from the symmetric Nash model. Note that whereas the upper and lower bounds grow closer to each other as we increase the k-level of iterated deletion of strategies, they do not necessarily converge to the symmetric Nash model. Indeed, it is possible to sustain asymmetric equilibria in this model of the type 1, firms enter because they expect other firms not to enter, and 2, firms stay out of the market because they expect other firms to enter. The larger the competitive interaction parameter α, the larger the split between the upper and lower bounds. In fact, it is this effect of competitive interaction α on the spread between the upper and lower bounds that makes it difficult to reject very highly competitive interactions.

Figure 3 presents the identified set described by (5) for the model of Aradillas-Lopez and Tamer using ready-mix concrete data where I let k go from 0 to 100. As k increases above 30, the blue shaded area in the top left disappears from the identified set, indicating that assuming a higher level-k of rationality shrinks the identified set. In particular, the highest possible α in the identified set decreases from $-.2$ to $-.5$ as k goes from 0 to about 30. Above $k = 30$, the identified set stays about the same, as would be expected given that in a finite number of iterated deletion of dominated strategies gives the set of rationalizable strategies. The upper bound on the effect of competition on profits is about $\alpha = -.50$, so we can state conclusively that there is an effect of competition on profits in the ready-mix concrete industry. But there is no lower bound on the effect of competition on profits, so we cannot reject the assertion that

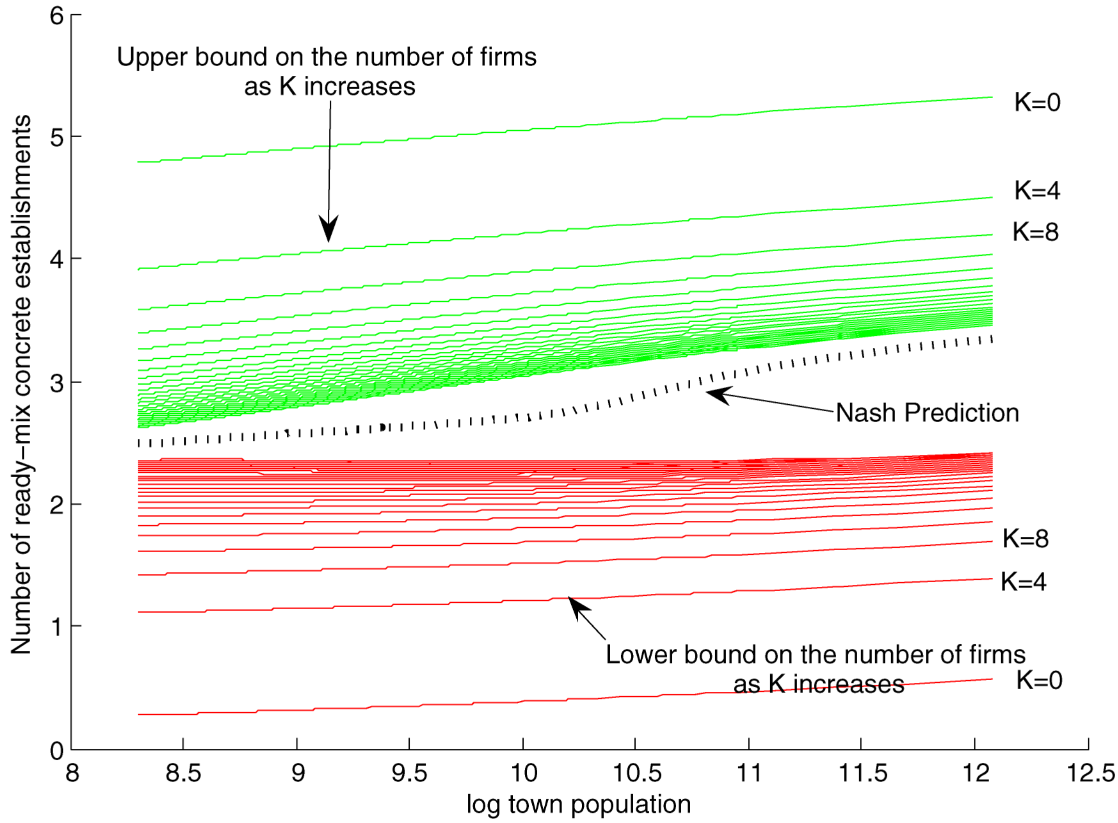

Figure 2. Model predictions for k levels of iterated deletion of dominated strategies and NE ($\alpha = -.42$ and $\beta = .1$).

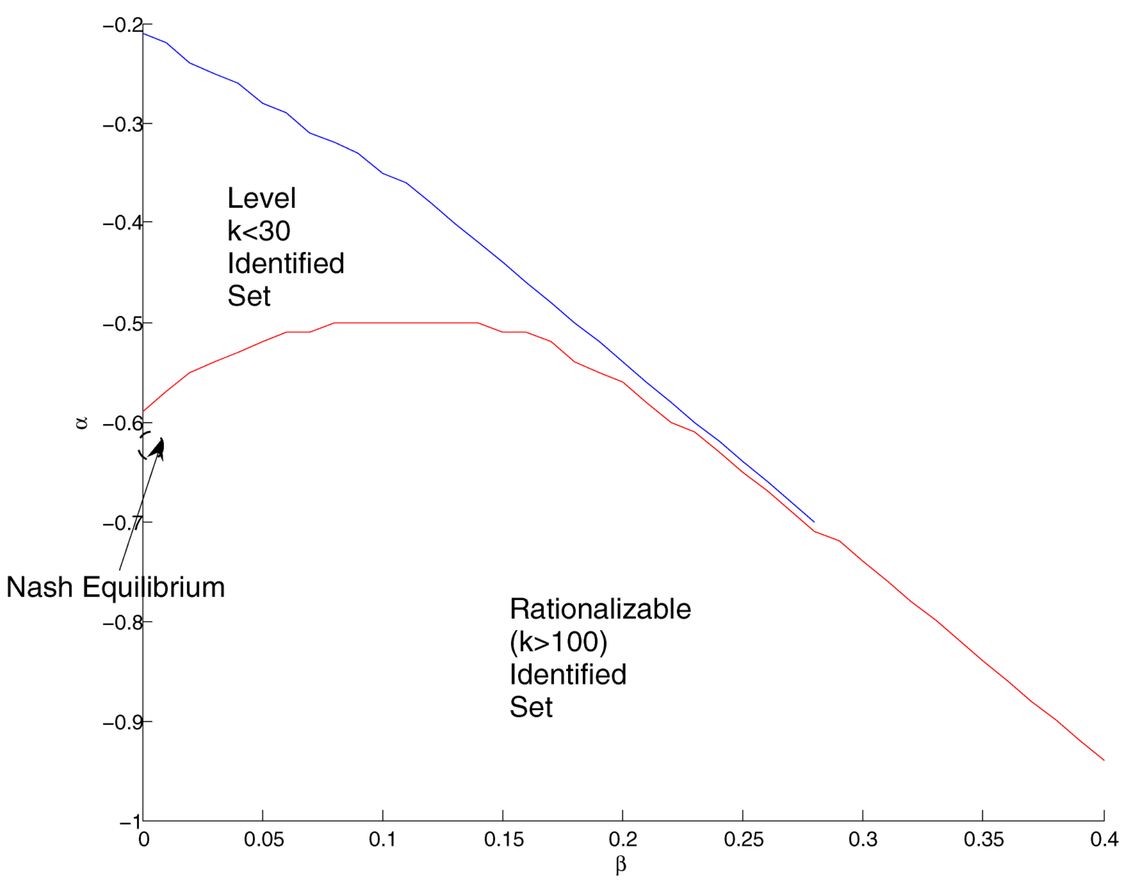

Figure 3. Identified set for the model of Aradillas-Lopez and Tamer using ready-mix concrete data [α, β such that $Q(\alpha, \beta) = 0$].

competition reduces profits by an arbitrarily large amount. To understand this result, it is worth remembering the increasing the effect of competition on profits pushes out the upper and lower bounds in Figure 2, because a large effect of competition on profits makes it possible to sustain asymmetric equilibria of the type where if I expect no other firm to enter, then I will enter for sure, and if I expect other firms to enter, then I will choose not to enter myself. Thus increasing the competitive parameter α will enlarge the set of permissible entry probabilities, which makes it impossible to form a lower bound on α. This seems to be a fairly generic result that casts some doubt on estimates from the entry literature on the strength of competition. Figure 3 also shows the location of the parameter that minimizes the Nash criterion function presented in (6). Note that this parameter gives no information on the size of the identified set.

2. DYNAMICS, LEARNING, AND STATIC MODELS

In the previous section I gave a sketch of an empirical implementation of the ideas of Aradillas-Lopez and Tamer. In this section I take a step back and discuss how closely the model that Aradillas-Lopez and Tamer propose matches common applications in empirical industrial organization.

2.1 Dynamic Oligopoly and Two-Period Models

The main gap between the model of Aradillas-Lopez and Tamer and the most recent work studying entry (e.g., Ryan 2006; Sweeting 2007; Collard-Wexler 2006) is the use of a two-period or static model instead of an explicitly dynamic model. In a two-period model, firms first make their entry decision and then receive a continuation value given the actions of other firms. These models were first used by Bresnahan and Reiss (1991) and Berry (1992) to study entry into dental and airline markets. Two-period models can be considered a "reduced-form" for a fully dynamic model, because airlines or dentists enter at different times and have the option to exit and reenter in the future. There are also a limited number of situations in which a simultaneous entry game corresponds exactly to the application in mind. I will discuss one particular example in the next section.

Using a static model to estimate behavior that is the outcome of a dynamic entry process may lead us to misunderstand the importance of multiple rationalizable strategies. For instance, much of the uncertainty in a static model come from the fact that either I do not know what my opponents will do or there are many possible actions that my opponents can take that are rationalizable. But in a dynamic context, this problem should be mitigated substantially, because entry and exit rates in most industries tend to be fairly low; for instance, in the ready-mix concrete industry, there is a 5% probability that a firm will enter or exit over the next year, and the probability of two firms entering simultaneously is $<1\%$. Given how low these entry and exit rates are, it is not clear how big a difference in expected payoffs that I would expect if I were using the most pessimistic or the most optimistic belief about the entry probabilities of my rivals. Moreover, given the high sunk costs in many industries, why not wait for other firms to enter before deciding to exit the market? Note that there are still multiple rationalizable strategies in dynamic games, but this is not the same problem as multiplicity in the static reduced form.

2.2 Learning and Nash

NE is frequently justified as the outcome from a learning process (see, e.g., Fudenberg and Levine 1998). In cases where agents have extensive experience with a market, NE can be justified as the outcome of a learning process. Thus rationalizeability seems to be too weak.

Alternatively, Nash can be conceived as a best response to the historical distribution of strategies used by firms. Suppose that I am a construction firm trying to bid in a first-price auction to build a highway. I might try to use introspection about the strategies of other players to form my beliefs about what other firms will bid. Alternatively, I could just review similar auctions to get an idea of the distribution of the bids of my opponents. In the context of entry into the ready-mix concrete industry, I could simply estimate the entry policies of my rivals $\hat{e}(x_i)$ in similar markets. I could then use these estimated probabilities to compute my expected value should I choose to enter.

Aradillas-Lopez and Tamer's model is an exact fit in a limited set of empirical applications. First, we need to look for situations in which firms enter at the same time. Second, an ideal application is to an industry where there is little experience of entry, which would allow firms to simply look at the distribution of the strategies of their opponents. Perhaps, the best example of this is the study of Augereau, Greenstein, and Rysman (2005) of the adoption of one of two types of 56-K modem technology by Internet service providers (ISPs). Between May and June 1997, ISPs had to choose one of two incompatible 56-K modem technologies. Choosing the same technology as rival ISPs lowered profits. This is clearly a one-shot game, because reversing the modem choice was expensive and there was no previous experience with the technology on which to draw to try to understand what opponents would do. Goldfarb and Yang (2007) extended the work Augereau, Greenstein, and Rysman (2005) by looking at weakening the Nash assumption in the 56-K modem adoption decision. They estimated the model of Camerer, Ho, and Chong (2004), which assumes that agents are playing against a Poisson distribution of level-k types. They found evidence that managers use limited introspection into the actions of their rivals, and that the degree of introspection is positively correlated with the education level within the market.

3. CONCLUSION

Aradillas-Lopez and Tamer have written a provocative article that challenges applied researchers in industrial organization to rethink the behavior that we use in empirical work. In particular, the assumption that agents are playing Nash can be quite strong in situations where there is limited experience with a market. Moreover, in many cases Nash is more than we need to estimate parameters. As I hope that this empirical implementation has demonstrated, I believe that Aradillas-Lopez and Tamer's approach will be adopted in empirical work.

ADDITIONAL REFERENCES

Augereau, A., Greenstein, S., and Rysman, M. (2005), "Coordination versus Differentiation in a Standards War: The Adoption of 56K Modems," *RAND Journal of Economics*, 37, 887–909.
Berry, S. T. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60, 29.
Camerer, C. F., Ho, T.-H., and Chong, K. (2004), "A Cognitive Hierarchy Model of Behavior in Games," *Quarterly Journal of Economics*, 19, 861.
Collard-Wexler, A. (2006), "Demand Fluctuations and Plant Turnover in the Ready-Mix Concrete Industry," working paper, New York University, Stern School of Business, Economics Department.
Fudenberg, D., and Levine, D. (1998), *The Theory of Learning in Games*, Cambridge, MA: MIT Press.
Goldfarb, A., and Yang, B. (2007), "Are All Managers Created Equal?" working paper, University of Toronto, Marketing Department.
Ryan, S. P. (2006), "The Costs of Environmental Regulation in a Concentrated Industry," working paper, Massachusetts Institute of Technology, Economics Department.
Seim, K. (2005), "An Empirical Model of Firm Entry With Endogenous Product-Type Choices," *Rand Journal of Economics*, 37, 619–640.
Sweeting, A. T. (2007), "The Costs of Product Repositioning: The Case of Format Switching in the Commercial Radio Industry," working paper, Duke University, Economics Department.

Rejoinder

Andres Aradillas-Lopez
Department of Economics, Princeton University, Princeton, NJ 08544

Elie Tamer
Department of Economics, Northwestern University, Evanston, IL 60208 (*tamer@northwestern.edu*)

We wish to thank all of the discussants for their thoughtful and interesting comments. We would like to take this opportunity to briefly address some of the issues raised implicitly and explicitly in their comments. In doing so, we would like to classify these comments into two broad categories: *implementation* and *extensions*.

1. IMPLEMENTATION

The aim of the article was to provide identification results under rationalizable behavior, leaving for future and ongoing research the problem of estimation and inference using these identification results. A number of comments were aimed at the issue of implementation and usefulness of these results for practitioners in more complicated game-theoretic settings than those studied in the article. We begin with the question of computational issues.

1.1 General Issues for Applied Researchers

One of the potential advantages of our approach is the ability to bypass the need to compute the equilibria in the game. But our article provided no insight into the actual computational complexity of doing inference based on rationalizable behavior. This was addressed through a nice example by Collard-Wexler, who estimated a static model of entry in the ready to mix cement industry using level-k rationality and compared those with estimates obtained assuming Nash. It turns out that with Nash, the model point identifies the parameter vector of interest, whereas with rationalizability, even as $k \to \infty$, the model puts one-sided bounds on the parameters. This is not surprising and in fact demonstrates the power of Nash in simple empirical setups where there is no heterogeneity. We also agree with Collard-Wexler's point that in some entry settings, static games of incomplete information might not be the preferred framework as a model and that a model with dynamics can be a better approximation. His comment also highlights the need to incorporate correlation in unobservables into the method. As we mention in the article, our constructive identification results can be extended to such cases, then the joint distribution of unobservables is assumed known, possibly up to a finite-dimensional parameter. Conceptually, our setup allows for correlation among the unobservables.

The need to develop a computational methodology based on rationalizable bounds that deals more realistically with unobserved heterogeneity was also addressed by Bajari. He correctly points out that a positive relationship between the actions of two agents could be the result of correlation across agents' unobserved heterogeneity instead of a strategic interaction effect. As we point out in the article, at least in cases where normal-form payoffs have a simple structure, we retain the ability to partially disentangle unobserved beliefs from unobserved heterogeneity and even would be able to deal with the case where beliefs are conditioned on unobservables as long as the joint distribution of these unobservables is assumed known. Work on cases where this distribution is unknown, although conceptually straightforward, is left for future research.

Bajari also suggests the need for an empirical application in which issues of rationalizability versus equilibrium can be examined. This is reasonable because identifying power of equilibrium is ultimately an empirical question. Bajari then questions the use of entry games as a basis for our model. Although it is true that a firm's decision to enter a market is dynamic in nature, static entry models can be used as models to shed light on long-term market structure. In addition, our framework for studying identification under rationality as we define it still can be used to study static games other than entry. It also is certainly possible to generalize the current setup to handle a large set of

Journal of Business & Economic Statistics
July 2008, Vol. 26, No. 3
DOI 10.1198/073500108000000169

players and a large strategy space. (We address this issue here in Sec. 2.2.) We do not pursue this here, because it is notationally cumbersome. On the broader issue of answering policy questions using partially identified structural models, our answer is fairly simple. Partially identified models can be viewed as providing a set of parameters, all of which are consistent with the economic model consistent with the economic model (or net of models). Thus doing policy analysis with those estimated parameters is no different, conceptually, than doing policy analysis with point-identified parameters, because here if interest lies in accounting for sample uncertainty, then one would want to map confidence regions for the point-identified parameter vector into a *set* of policy effects. When conducting inference or policy question using a partially identified model, one basically uses a similar map from a confidence region (on the set or the parameter) into a confidence region for policy effects.

The computational part of the article is lacking and needs to be addressed. These issues are raised in the comment by Hong, who illustrates fundamental differences between the computational strategies needed to estimate complete-information and incomplete-information games under the assumption of Nash equilibrium behavior. Although we speculate that these substantive differences would be mitigated if the assumption of Nash equilibrium were replaced with that of level-k rationalizability, we find his suggestion of using simulation-based estimation procedures a very interesting area of research, because it might enable researchers to treat the "rationality level-k" explicitly as an additional structural parameter to estimate. It also might be possible to exploit the structure of the iterative procedure to make the computations simpler and more efficient. Homotopy-based computational methods appear particularly promising in some models. We leave this as a project for further research.

Magnac stresses the need to provide guidance for applied researchers as to the type of parameter normalizations that are needed before implementing our procedure. The identification results in the article are based on the same type of parameter normalizations that one would use in Nash equilibrium–based models. This is possible because of three key assumptions: (1) The distribution of unobservables is known (up to finite-dimensional parameters) to the researcher; (2) beliefs are conditioned on observables; and (3) the sign of the strategic interaction effect is known ex ante, which enables us to characterize the lower and upper rationalizable bounds for any given parameter value and any level-k. It can be shown (see Aradillas-Lopez 2007) that knowing the signs of the strategic interaction parameters is not essential to establish identification results based on rationalizability even in larger games; however, it also can be shown that under some conditions (e.g., if there exists asymmetric information across players), to obtain well-defined asymptotic properties from sample analog estimators of rationalizable bounds, the researcher needs to rule out the case in which there is no strategic interaction. Loosely speaking, in such a scenario the parameter space must rule out the case in which the strategic interaction effects are zero for all players. Returning to our previous discussion about the case in which beliefs are conditioned on unobservables, we contend that the parameter normalization analogy with equilibrium models also would follow through. To be precise, we speculate that assuming a joint distribution for unobservables where each marginal scale is normalized would enable us to include a parameter that measures interdependence across players' unobserved heterogeneity in the parameter space. Magnac also discusses the rationality level-k as a structural parameter in the model. In the article we present some identification results for the smallest level-k in the population of players, but our setting is unable to identify the rationality level of each individual. Constructing a model capable of producing identification results for each agent's rationality level-k requires adding more structure. It appears to be difficult to use economic theory to add such a structure into one-shot games such as those that we study. In contrast, using a learning theory in a dynamic game setting seems like a natural application where this can be done.

1.2 Sharpness

This issue was raised by Molinari and Rosen as well as Khan, who also pointed out that θ also contains information about the relationship between the L_k- and the L_{k-1}–rational bounds. This is important because any L_k-rational player is also L_{k-1}–rational. The concept of level-k rationalizability enables us to characterize the sharp identified set for θ in a straightforward way, by looking at all decision rules produced by some well-defined beliefs that belong in the space of L_k-rational beliefs. Although this characterization is not constructive if such a space contains infinitely many elements (e.g., when conditioning signals have continuous support), it is possible to find a characterization of the identified set that is sharper than the conditional moment inequality description used in the article. For example, in the game with complete information, incorporating a reduced-form selection rule that depends on both observed and unobserved regressors delivers sharp estimates (conditional on the selection rule being correct); see the discussion in section 3.1 and 4.2 and definition (11) for example.

In the article we focused on conditional moment inequalities due to their simplicity and the fact that they can be generalized to larger games. We also showed that an identification-at-infinity result holds for θ using the moment inequalities for L_2-rational players if the conditioning signals have sufficiently rich support.

1.3 Counterfactual Analysis With Rationalizable Behavior

Counterfactual analysis for strategic-interaction models is challenging even if Nash equilibrium behavior is assumed. As Aguirregabiria points out, counterfactual experiments may be only partially identified even if the payoff parameters of the underlying game are point-identified. In the presence of multiple equilibria, point-identification of a counterfactual experiment would require the introduction of an equilibrium selection mechanism. An analogous device might enable us to obtain point predictions for counterfactual analysis with rationalizable behavior. This device would be a *belief-selection mechanism*. Suppose that all players are assumed to be (at least) L_k-rational in the 2×2 incomplete information game studied in the article. Instead of being completely agnostic about the way in which players choose their beliefs within the L_k-rational space, we could represent their optimal decision rules as

$$Y_p = \mathbf{1}\{X_p'\beta_p + \alpha_p[\xi_p \cdot \pi_{-p}^U(\theta|k;\mathcal{I}) + (1-\xi_p)\cdot \pi_{-p}^L(\theta|k;\mathcal{I})] \geq \varepsilon_p\}. \quad (1)$$

The connected nature of the space of L_k-rational beliefs implies that every L_k-rational player follows a decision rule of the type described in (1). If we add structure on the weights $\xi_p \in [0, 1]$, then we can treat them as a belief-selection mechanism. Suppose, for example, that $\xi_p = \xi_p(V_p, \phi)$, where V_p is observable and $\xi_p(\cdot)$ is of known functional form, except for a finite-dimensional parameter ϕ. This yields

$$Y_p = \mathbf{1}\{X_p'\beta_p + \alpha_p[\xi_p(V_p,\phi) \cdot \pi_{-p}^U(\theta|k;\mathcal{I}) + (1-\xi_p(V_p,\phi))\cdot \pi_{-p}^L(\theta|k;\mathcal{I})] \geq \varepsilon_p\}. \quad (2)$$

The expression in (2) is analogous to introducing an equilibrium selection mechanism in a model that assumes BNE behavior. Introducing an equilibrium selection theory is not uncommon in econometric models of games with multiple equilibria. Let $\eta \equiv (\theta', \phi')$. Under our assumptions, the lower and upper bounds $\pi_{-p}^L(\theta|k;\mathcal{I})$ and $\pi_{-p}^U(\theta|k;\mathcal{I})$ are identified for any given k and θ. If $\alpha_p \neq 0$ for both players and if $k \geq 2$ (i.e., every player performs at least one round of deletion of dominated strategies), then it would be possible to point-identify η from (2) *for a given k* if there exists an exclusion restriction between W_p and $(X_p, Z_p, \mathcal{I})$. Bajari points out the need to devise procedures that enable the practitioner to do counterfactuals. In the context of equilibrium models, he stresses the need to take a stance on equilibrium selection. Here belief selection mechanisms can be used to do counterfactuals with rationalizable behavior.

2. EXTENSIONS

The commentators have gone to great lengths to discuss extensions of the methods in our article to more complicated and interesting games, as well as to individual decision making models in which there is no strategic interaction.

2.1 Dynamic and Supermodular Games

Aguirregabiria extends our analysis to dynamic discrete choice games. This is a very important direction for research. He proposes to drop the requirement in Markov games that players' strategies are common knowledge but to maintain the assumptions that strategies are constant over time (Markov) and that players are forward-looking expected utility maximizers. He then derives rationalizable bounds on conditional choice probabilities that can be used to obtain point-identification results by exploiting regressor variation. Overall, this is an important direction for future research, because it is much harder to deal with multiple equilibria in dynamic models, and thus moving to a weaker concept for dynamic strategic behavior might allow for easier setups. Although the strategic setup that Aguirregabiria proposes is interesting, whether it is the only one is not clear, and we need to explore further, for example, a direct generalization of our rationalizability concept (i.e., the extensive form version of rationalizability studied by Pearce). As we mentioned previously, a dynamic setting appears to be ideal for introducing a theory of learning and enhance the model's identification power for each player's rationality level-k. This is explored in more detail in the comment by Collard-Wexler, although his discussion is not aimed at ways to use learning to identify individual rationality levels. It appears that even a simple characterization where each agent's k increases with time might be an interesting starting point. More work is needed in this area.

In careful detail, Molinari and Rosen extend the analysis to sets of supermodular games. This is a rich set that includes entry games as well as Cournot and other pricing games. They show how for a given parameterization of a pricing game, it is possible to derive the identification power of rationality of level-k and compare it with equilibrium. They provide a set of usable inequality restrictions that can be used to estimate the parameters of interest. This is a very interesting direction because the set of supermodular games is wide and more amenable to empirical applications. Their results show that the results and methods in our article have applications well beyond simple two-player games, and we believe that supermodularity appears to be the best approach for adapting our methods to games with multinomial (or continuous) actions. Even though it is not explicit in their results, a key feature shared by all extensions discussed by the commentators is the fact that the effect of other agents' choices on each particular player's payoffs is bounded with probability 1. It follows from here that the most optimistic (pessimistic) L_k assessment for expected utility is well defined. The extent to which this can be extended to games in which actions have an unbounded support is an area of future and ongoing research. Beyond the game-theoretic setting, we agree with Magnac's discussion that it would be fruitful to extend the ideas here to inference in dynamic single-agent models as done by Rust, who established nonpoint identification results. Perhaps an extension of our approach could be combined with behavioral theoretical results about self-control and procrastination, where a single agent's dynamic decision is modeled as a game played against one's "future self."

2.2 Games With Multiple Players

The viability of extending the methods proposed here to multiple-player settings was discussion by several commentators. The results in our article can be extended to games with multiple players if the effect of opponents' actions on each player's payoffs is bounded with probability 1 and if the computational task of characterizing the most pessimistic (optimistic) beliefs is well defined. For example, Aradillas-Lopez (2007) studied an incomplete-information game with $\mathcal{P} \geq 2$ players, each of which faces a binary choice $Y_p \in \{0, 1\}$. The effect of opponents' actions is summarized by an aggregate variable that is bounded with probability 1. Specifically, normal-form payoffs are given by

$$Y_p \times \left(X_p'\beta_p + \sum_{M=0}^{\mathcal{P}-1} \Delta_p^M \mathbf{Y}_{-p}^M - \varepsilon_p\right), \quad \text{where } \mathbf{Y}_{-p}^M = \mathbf{1}\left\{\sum_{q\neq p} Y_q = M\right\}, \quad (3)$$

where X_p denotes a vector of observable (to the econometrician) covariates; ε_p denotes an unobservable, privately observed, continuously distributed shock; and $(\beta_p, \Delta_p^0, \ldots, \Delta_p^{\mathcal{P}-1}) \equiv \theta_p$ denotes a vector of constant parameters. Alternatively, a payoff parameterization in which opponents' individual identities matter can be handled easily. The 2×2 game studied in our article is a special case of (3). The key feature that allows us to extend the methods that we analyze to the game described in (3) is the fact that the aggregate variable $\sum_{M=0}^{\mathcal{P}-1} \Delta_p^M \mathbf{Y}_{-p}^M$ is bounded with probability 1. Furthermore, its linear nature also ensures that the most pessimistic (optimistic) assessment for expected utility is well defined as the solution to a linear programming problem. This also preserves the dominance-solvable nature of the game and the limiting relationship between BNE and rationalizability as $k \to \infty$. (We refer the interested reader to Aradillas-Lopez 2007 for details.)

3. CONCLUSION

We concur with the discussants that this article is a starting point, and more needs to be done to address some of the concerns of applied researchers when applying the methods that we propose. As some of the comments illustrate, extensions to more realistic settings, such as dynamic models, as well as games with multiple players and actions, can be done in a straightforward way. The issue of sharpness deserves special attention since in cases with partial identification, one ought to obtain the identified set. It is also important to address the case in which beliefs are conditioned on unobservables. Here we have argued that these cases can be analyzed using simulation-based methods as long as the distribution of these unobservables is assumed known, possibly up to a finite-dimensional parameter. We hope to adequately address some of the comments and concerns in future research, and we thank the discussants once again for their thoughtful input.

ADDITIONAL REFERENCE

Aradillas-Lopez, A. (2007), "Using Rationalizable Bounds for Conditional Choice Probabilities as Control Variables," working paper, Princeton University, Department of Economics.