The Identification Power of Equilibrium in Simple Games∗

Andres Aradillas-Lopez† Elie Tamer‡
Princeton University Northwestern University

March 2008

Abstract

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 where we drop the Nash equilibrium assumption and instead use rationalizability (Bernheim (1984) and Pearce (1984)) 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 $\geq 1$. 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 concept 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 role different assumptions play. We provide constructive identification results that lead naturally to consistent estimators.

∗We thank seminar participants at many institutions for comments. We also thank M. Engers, D. Fudenberg, F. Molinari, M. Siniscalchi, and A. Pakes for useful comments. We especially thank Serena Ng for her diligent editorial help and the discussants for their valuable comments.
†Department of Economics, Princeton University, Princeton NJ 08544. Support from the National Science Foundation grant # SES 0718409, and from the Gregory C. Chow Econometric Research Program at Princeton University are gratefully acknowledged.
‡Department of Economics, Northwestern University, Evanston IL 60208. Support from the National Science Foundation is gratefully acknowledged.
1 Introduction

This paper examines 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 parametrizes payoff functions. We compare what can be learned about this vector from 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). We find that in static discrete games with complete information, the identified features of the games with more than one level of rationality is similar to what one obtains 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) Nash Equilibrium, then there is convergence between the identified features with and without equilibrium only when the level of rationality tends to infinity. When there is multiple equilibria, the identified features of the game under rationalizability 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 when imposing rationalizability (no need to solve for fixed points). In the auction game 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 the latent distribution of valuation. Another strategic assumptions in auctions resulting in tighter bounds is the concept of $P$-dominance studied in 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 where NE strategy profiles are unreasonable while on the other hand, one can also find reasonable strategy profiles that are not Nash. Restrictions once Nash behavior is dropped are based typically on a set of “rationality” criteria. These criteria are enumerated in different papers and under different strategic scenarios. This paper studies 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 see its effect on parameter inference. Thus, depending on the application, identification of parameters of interest can certainly be studied under strategic assumptions other than rationalizability. We provide such an example in this paper.
Since 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 notes, “this indeterminacy is an accurate reflection of the difficult situation faced by players in a game” since logical guidelines and the rules of the game are not sufficient for uniqueness of predicted behavior. Hence, 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 that is 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 strategic environment, we implicitly assume that all players in that game are abiding exactly by these assumptions and playing exactly that game\(^1\).

Using equilibrium as a restriction to gain identifying power is well known in economics\(^2\). The objective of the paper is to 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 – of strategies and beliefs. 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 which may include, for example, the payoff parameters in a reduced-form game, or the distribution of valuations in an auction, as well as on the properties of higher order beliefs by the agents, which are incorporated into the econometric analysis. The features of this hierarchy of beliefs will characterize what we refer to as 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 has also been studied in 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 (2007a). 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 papers assume the existence of a small number of pre-specified 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-respond to a uniform distribution over the surviving actions. Using carefully designed experiments, these researchers have sought to explain which type fits the observed choices the best. This paper differs from the aforementioned work because we focus 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 where the researcher ignores how “rational” players are and where other primitives of the game are also 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 are typically 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 in 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 Nash equilibrium (see Bjorn and Vuong (1985), Bresnahan and Reiss (1991), Berry (1994), Tamer (2003), Andrews, Berry, and Jia (2003), Ciliberto and Tamer (2003), and Bajari, Hong, and Ryan (2005) among others). Here, we find that in the \( 2 \times 2 \) game with level-2 rationality, the outcomes of the game coincide with Nash, and hence econometric restrictions are the same. Section 4 considers static games with incomplete information. Empirical frameworks for these games are studied in Aradillas-Lopez (2005), Aguirregabiria and Mira (2004), Seim (2002), Pakes, Porter, Ho, and Ishii (2005), Berry and Tamer (2006) among 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 in 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 $2 \times 2$ 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 which map into stronger restrictions that can be used for identification. If the game admits a unique equilibrium, the restrictions of the model converge towards 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 since one does not need to solve for fixed points, but to simply iterate the beliefs towards 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 Siniscalchi (2003). Here, for any order $k$, we are only able to bound the unobserved valuation from above. Finally, section 6 concludes.

2 Nash Equilibrium and Rationality

In noncooperative strategic environment, 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 their 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 paper, we examine the effect of assuming Nash behavior on identification by comparing restrictions under Nash with ones obtained under rationality in the sense of Bernheim and Pearce. Below, 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 are given in 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 make use of the following notations. Let $R^i(0)$ be the set of all (possibly mixed) strategies that player $i$ can play and $R^{-i}(0)$ is the set of all strategies for players other than $i$. Then, heuristically:

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

• Level-2 rational strategies for player $i$ are strategy profiles $s^i \in R^i(0)$ such that there exists a strategy profile for other players in $R^{-i}(1)$ for which $s^i$ is a best response.

• Level-$t$ rational strategies: Defined recursively from level 1.

Notice that by construction, $R^i(1) \subseteq R^i(0)$ and $R^i(t) \subseteq R^i(t - 1)$. Finally, rationalizable strategies are ones that lie in the intersection of the $R$’s as $t$ increases to infinity. In the complete information game of Section 3 we will see that there exists a finite $k$ such that for $R^i(t) = R^i(k)$ for all $t \geq k$. In the incomplete information models of Sections 4 and 5, we will see that we can have $R^i(t) \subset R^i(k)$ for all $t > k$. In all these settings, a strategy is level-$k$ rational for a player if it is a best response to some strategy profile in $R^i(k - 1)$ by his opponents. If we iterate 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 this 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. We have parameters $\alpha_1$ and $\alpha_2$ that are of interest. The econometrician does not observe $t_1$ or $t_2$ and is interested in learning about the $\alpha$’s and the joint distribution of $(t_1, t_2)$. Assume also, as in entry games, that the $\alpha$’s are negative. In this example and the next, we assume that one has access to a random sample of observations
To learn about the parameters, we map the observed distribution of the data (the choice probabilities) to the distribution predicted by the model. Since this is a game of complete information, players observe all the payoff relevant information. In particular, in the first round of rationality, player 1 will play 1 if \( t_1 + \alpha_1 \geq 0 \) since this will be a dominant strategy. In addition, if \( t_1 \) is negative, player 1 will play 0. However, when \( t_1 + \alpha_1 \leq 0 \leq t_1 \), both actions 1 and 0 are level-1 rational: action 1 is rational since it can be a best response to player 2 playing 0, while action 0 is a best response to player 2 playing 1. The set \( \mathcal{R}(1) \) is summarized in Figure 1. For example, consider the upper right hand corner. For values of \( t_1 \) and \( t_2 \) lying there, playing 0 is not a best response for either player. Hence, \((1,1)\) is the unique level-1 rationalizable strategy (which is also the unique NE). Consider now the middle region on the right hand side, i.e., \((t_1,t_2) \in [-\alpha_1,\infty) \times [0,-\alpha_2] \). In level-1 rationality, 0 is not a best reply for player 1, but 2 can play either 1 or 0: 1 is a best reply when 1 plays 0, and 0 is a best reply for player 2 when 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. Hence \( R(1) = \{\{1\}, \{0, 1\}\} \) while the rationalizable set reduces to the outcome \((1, 0)\). Here, \( R(k) = R(2) = \{\{1\}, \{0\}\} \) for all \( k \geq 2 \). In the middle square, we see that the game provides no observable restrictions: any outcome can be potentially observable since both strategies are rational at any level of rationality. Notice 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.

### 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 \( \theta = (\alpha_1, \alpha_2, F(\ldots)) \) where \( F(\ldots) \) is the joint distribution of \((t_1, t_2)\). One interesting approach to conduct inference on the sharp set 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 \( \sum_{i,j} p_{ij} = 1 \). Hence, we make inference on the set of probabilities \((p_{ij}, i, j \leq k)\) and \((\alpha_1, \alpha_2)\). We highlight this below for level 2 rationality. In particular, we say that

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

if and only if:

\[
P_{11} = \sum_{i,j} p_{ij} \left( 1[\alpha \leq -\alpha_1; \gamma \geq -\alpha_2] + l_{ij}^{(1,1)} 1[\alpha \leq -\alpha_1; \gamma \leq -\alpha_2] \right)
\]

\[
P_{00} = \sum_{i,j} p_{ij} \left( 1[\gamma \leq 0; \alpha \leq 0] + l_{ij}^{(0,0)} 1[\gamma \leq -\alpha_1; \alpha \leq -\alpha_2] \right)
\]

\[
P_{10} = \sum_{i,j} p_{ij} \left( 1[\alpha \geq 0; \gamma \leq 0] + 1[\gamma \leq -\alpha_1; \alpha \leq -\alpha_2] + l_{ij}^{(1,0)} 1[\gamma \leq -\alpha_1; \alpha \leq -\alpha_2] \right)
\]

\[
P_{01} = \sum_{i,j} p_{ij} \left( 1[\gamma \leq 0; \alpha \geq 0] + 1[\alpha \leq -\alpha_1; \gamma \geq -\alpha_2] + l_{ij}^{(0,1)} 1[\alpha \leq -\alpha_1; \gamma \leq -\alpha_2] \right)
\]

for some \((l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)}) \geq 0 \) and \( l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1 \) for all \( i, j \leq k \). One can think of the \( l \)'s as the "selection mechanisms" that pick an outcome in the region where the model predicts multiple outcomes. We are treating the support points as known, but this is without loss of generality since those too can be made part of \( \theta \). The above equalities (and inequalities), for a given \( \theta \), are similar to first order conditions from a linear programming problem and hence can be solved fast using linear programming algorithms. In particular, Consider the objective function in (3.1) below. Note first that \( Q(\theta) \leq 0 \) for all \( \theta \) in the
parameter space. And,

$$\theta \in \Theta_f$$

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

\[
Q(\theta) = \max_{v_1, \ldots, v_8, \theta} \left( -v_1 - \ldots - v_8 \right) \quad \text{s.t.}
\]

\[
P_{i1} = \sum_{i,j} p_{ij} \left( 1[s_i \geq -\alpha_1; s_j \geq -\alpha_2] + l_{ij}^{(1,1)}1[0 \leq s_i \leq -\alpha_1; 0 \leq s_j \leq -\alpha_2] \right) = v_1 - v_2
\]

\[
P_{00} = \sum_{i,j} p_{ij} \left( 1[s_i \leq 0; s_j \leq 0] + l_{ij}^{(0,0)}1[0 \leq s_i \leq -\alpha_1; 0 \leq s_j \leq -\alpha_2] \right) = v_3 - v_4
\]

\[
P_{00} = \sum_{i,j} p_{ij} \left( 1[s_i \geq 0; s_j \geq 0] + 1[s_i \geq -\alpha_1; 0 \leq s_i \leq -\alpha_2] + l_{ij}^{(0,1)}1[0 \leq s_i \leq -\alpha_1; 0 \leq s_j \leq -\alpha_2] \right) = v_5 - v_6
\]

\[
P_{01} = \sum_{i,j} p_{ij} \left( 1[s_i \leq 0; s_j \geq 0] + 1[0 \leq s_i \leq -\alpha_1; s_j \geq -\alpha_2] + l_{ij}^{(0,1)}1[0 \leq s_i \leq -\alpha_1; 0 \leq s_j \leq -\alpha_2] \right) = v_7 - v_8
\]

\[
v_i \geq 0; \quad \left( l_{ij}^{(1,1)}, l_{ij}^{(0,0)}, l_{ij}^{(0,1)}, l_{ij}^{(1,0)} \right) \geq 0; \quad l_{ij}^{(1,1)} + l_{ij}^{(0,0)} + l_{ij}^{(0,1)} + l_{ij}^{(1,0)} = 1 \quad \text{for all} \quad 1 \leq i, j \leq k
\]

(3.1)

First, note that for any \( \theta \), 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 \( v_1 = P_{11} - \sum_{i,j} p_{ij}1[s_i \geq -\alpha_1; s_j \geq -\alpha_2] \) and \( v_2 = 0 \) if \( P_{11} - \sum_{i,j} p_{ij}1[s_i \geq -\alpha_1; s_j \geq -\alpha_2] \geq 0 \), otherwise set \( v_2 = -(P_{11} - \sum_{i,j} p_{ij}1[s_i \geq -\alpha_1; s_j \geq -\alpha_2]) \) and \( v_1 = 0 \) and similarly for the rest. Moreover, \( \theta \in \Theta_f \) if and only if \( Q(\theta) = 0 \). One can collect all the parameter values for which the above objective function is equal to zero (or approximately equal to zero) A similar linear programming procedure was used in Honoré and Tamer (2005). The sampling variation comes from having to replace the choice probabilities \( P_{11}, P_{12}, P_{21}, P_{22} \) with their sample analogs which would result in a sample objective function \( Q_n(.) \) that can be used to conduct inference.

More generally, and without making support assumptions, a practical way to conduct inference if assumes 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 above inequalities do not exploit all the information and hence the identified set based on these inequalities is not sharp\(^3\). However, these inequalities based moment conditions are simple to use and can be generalized to large games. Heuristically then, the model
identifies, by definition, the set of parameters $\Theta_I$ such that the above inequalities are satisfied. Moreover, we say that the model point identifies a unique $\theta$ if the set $\Theta_I$ is a singleton.

In Figure 2 we provide 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 we allow for mixed strategies. In particular, without allowing for mixed strategies, in the middle square of Figure 2, the only observable implication is (1, 0) and (0, 1). However, it reverts to all outcomes once one consider the mixed strategy equilibrium. To get an idea of the identification gains when we assume rationality vs equilibrium, we simulated a stylized version of the above game in the case where $t_p$ is standard normal for $p = 1, 2$ and the only object of interest is the vector $(\alpha_1, \alpha_2)$. We compare the identified set of the above game under $k = 1$ rationality and NE when we only consider pure strategies.

We see from Figure 3 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 while under rationality, the model only provides upper bounds on the alpha’s. But, if we add exogenous variations in the profits ($X$’s), the identified region under rationality will shrink. The next section examines 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 above but under the assumptions that player 1 (2) does not observe $t_2$ ($t_1$) or that the signals are private information. We will
Above we see the identified regions for \((\alpha_1, \alpha_2)\) under \(k = 1\)-rationality (left display) and Nash (right display). 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.) Notice that on the left, the model only places upper bounds on the alphas. Under Nash on the other hand, \((\alpha_1, \alpha_2)\) are constrained to lie a much smaller set (the inner “circle”).

denote player \(p \in \{1, 2\}\)’s opponent to be \(-p\). We will let \(I_p\) denote the signals used by player \(p\) to obtain information about \(t_{-p}\), where \(t_p \in I_p\) could be a special case. Player \(p\) holds beliefs about his opponent’s type conditional on \(I_p\), and these beliefs can be summarized by a subjective distribution function. Let \(\pi_2(I_1)\) denote player 1’s subjective probability of entry for player 2, and define \(\pi_1(I_2)\) analogously for player 2. Given his beliefs, the expected utility function of player 1 is

\[
U(a_1, t_1, I_1) = \begin{cases} 
  t_1 + \alpha_1 \pi_2(I_1) & \text{if } a_1 = 1 \\
  0 & \text{otherwise}
\end{cases}
\]

Similarly for player 2 we have

\[
U(a_2, t_2, I_2) = \begin{cases} 
  t_2 + \alpha_1 \pi_1(I_2) & \text{if } a_2 = 1 \\
  0 & \text{otherwise}
\end{cases}
\]

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

\[
Y_1 = \mathbb{1}\{U(1, t_1, I_1) \geq 0\}, \quad Y_2 = \mathbb{1}\{U(1, t_2, I_2) \geq 0\}.
\]

(4.2)
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, since we focus on the case where $t_p$ is continuously distributed, the event $U(1, t_p, I_p) = 0$ occurs with probability zero. Our assumptions differ from Nash equilibrium because we do not impose the restriction that subjective beliefs are consistent with players’ actual behavior. Again, we assume here both $\alpha_1$ and $\alpha_2$ are negative.

4.1 Implications of level-1 rationality

We maintain the expected utility maximization assumption and the resulting decision rules (4.2). In the first round of rationality, we know that for any belief function, or without making any common prior assumptions, the following holds:

\[
\begin{align*}
    t_1 + \alpha_1 &\geq 0 \implies U(1, t_1, I_1) = t_1 + \alpha_1 \pi_2(I_1) \geq 0 \quad \forall \pi_2(I_1) \in [0, 1] \\
t_1 &\leq 0 \implies U(1, t_1, I_1) = t_1 + \alpha_1 \pi_2(I_1) \leq 0 \quad \forall \pi_2(I_1) \in [0, 1]
\end{align*}
\]

(4.3)

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, he must satisfy

\[
\begin{align*}
    t_1 + \alpha_1 &\geq 0 \implies a_1 = 1 \\
t_1 &\leq 0 \implies a_1 = 0
\end{align*}
\]

Now, let $0 \leq t_1 \leq -\alpha_1$. For a player that is rational of order one, there exists well defined beliefs that rationalizes either 1 or 0. Hence, 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. Notice here that the $(t_1, t_2)$ space is divided into 9 regions: 4 regions where the outcome is unique, 4 regions with 2 potentially observable outcomes, and the middle square where any outcome is potentially observed. To make inference based on this model, one needs to map these regions into predicted choice probabilities. To obtain the sharp set of parameters that is identified by the model, one can supplement this model with consistent “selection rules” that specifies in regions of multiplicity, the probability of observing the various outcomes (this would be a function of of both $(t_1, I_1)$ and $(t_2, I_2)$).

**Result 1** For the game with incomplete information, let the players be rational with order 1 (level-1 rational) and denote $W_p \equiv t_p \cup I_p$. Then the choice probabilities predicted by the
model are:

\[
P(1,1) = \int_{III} dF + \int_{II} S^{II}_{(1,1)}(W_1,W_2)dF + \int_{VI} S^{VI}_{(1,1)}(W_1,W_2)dF + \int_{V} S^{V}_{(1,1)}(W_1,W_2)dF
\]
\[
P(0,0) = \int_{VII} dF + \int_{VIII} S^{VIII}_{(0,0)}(W_1,W_2)dF + \int_{IV} S^{IV}_{(0,0)}(W_1,W_2)dF + \int_{V} S^{V}_{(0,0)}(W_1,W_2)dF
\]
\[
P(0,1) = \int_{I} dF + \int_{II} S^{II}_{(0,1)}(W_1,W_2)dF + \int_{IV} S^{IV}_{(0,1)}(W_1,W_2)dF + \int_{V} S^{V}_{(0,1)}(W_1,W_2)dF
\]

(4.4)

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

The functions \( S \) above are unknown and represent “selection” functions that represent the probabilities of selecting a particular outcome in a region of multiplicity. Suppose for simplicity that \( T_p = t_p \), so players condition their beliefs only on the realization of their own type. The restrictions in (4.4) can be exploited, for example, by discretizing the joint distribution of \((t_1, t_2)\) for example as we discussed in the complete-information case, to construct the identified set. The latter, \( \Theta_I \), is the set of parameters for which the above equalities in (4.4) are satisfied. Implications of the set of equalities above is a set of moment inequalities that are constructed by exploiting the fact that the \( S \) functions are probabilities and hence are positive. So, for example, an implication of Result 1 above is that
\[
\int_{III} dF \leq P(1,1) \leq \int_{III} dF + \int_{II\cup III\cup UV} dF
\]
where the bounds of this inequalities do not involve the unknown functions \(S\).

Next, we will analyze the behavior of players who assume that their opponents are (at least) level-k rational for \(k \geq 1\). Level-2 rational players will be those whose second order beliefs for their opponent are compatible with the bounds implied by (4.3). As we will see, by eliminating beliefs that violate (4.3), we will be able to reduce the set of level-2 rational beliefs, from the entire \([0,1]\) interval to a segment of it. Further rounds of iterated thinking will refine these bounds even more. Unlike the Bayesian-Nash equilibrium case, we will not impose the requirement that beliefs are correct. We will only rule 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 will 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 for example that level-2 rational players are all those whose beliefs are consistent with the bounds implied by (4.3). As we will see, by eliminating beliefs that violate (4.3), 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 will be 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 \(I_p\), with \(t_p \in I_p\) being a special case. The true distribution of \((t_1 \cup t_2 \cup I_1 \cup I_2)\) is common knowledge to both players. This is the common prior assumption. Even though it played no role in the analysis of level-1 rational behavior, the common prior assumption will be important for higher levels of rationality. We will consider strategies (decision rules) for player \(p\) that are threshold functions of \(t_p\):

\[
Y_p = \mathbb{1}\{t_p \geq \mu_p\} \quad \text{for } p = 1, 2,
\]

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 as well as those who play a Bayesian Nash equilibrium (BNE) are two special cases. In the construction of his expected utility, player \( p \) forms \textit{subjective} beliefs about \( \mu_{-p} \) that can be summarized by a probability distribution for \( \mu_{-p} \) given \( I_p \). These beliefs are derived as part of a solution concept. For example they include BNE beliefs as a special case (in which case all players know those equilibrium beliefs to be correct). Here, let \( \hat{G}_1(\mu_2|I_1) \) denote Player 1’s subjective distribution function for \( \mu_2 \) given \( I_1 \), and define \( \hat{G}_2(\mu_1|I_2) \) analogously for player 2. A strategy by player \( p \) is \textit{rationalizable} if it is the best response (in the expected-utility sense) given some beliefs \( \hat{G}_p(\mu_{-p}|I_p) \) that assign zero probability mass to strictly dominated strategies by player \(-p\). A rationalizable strategy by player \( p \) is described by

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

where the support \( S(\hat{G}_p) \) excludes values of \( \mu \) that result in dominated strategies within the class (4.5) (see footnote 5). Note that the subset of rationalizable strategies within the class (4.5) are of the form \( \mu_p = -\alpha_p \int_{S(\hat{G}_p)} E\left[ \mathbb{1}\{t_{-p} \geq \mu\} | I_p, \mu \right] d\hat{G}_p(\mu|I_p) \). In this setting, rationalizability requires expected utility maximization for a given set of beliefs, but it does not require those beliefs to be correct. It only imposes the condition that \( S(\hat{G}_p) \) exclude values of \( \mu_{-p} \) that are dominated. We eliminate such values by iterated deletion of dominated strategies.

We describe now the iterative procedure that restrict \( S(\hat{G}_p) \) by iterated dominance. As before, we maintain that the signs of the strategic-interaction parameters \( (\alpha_1, \alpha_2) \) are known. Specifically, suppose \( \alpha_p \leq 0 \). Then, repeating arguments from the previous section on \( k = 1 \)-rationalizable outcomes, we see looking at (4.6) that we must have (event-wise comparisons):

\[
\mathbb{1}\{t_p + \alpha_p \geq 0\} \leq \mathbb{1}\{Y_p = 1\} \quad \text{and} \quad \mathbb{1}\{t_p < 0\} \leq \mathbb{1}\{Y_p = 0\}
\]

\textit{Decision rules that do not satisfy these conditions are strictly dominated for all possible beliefs.} Therefore, the subset of strategies within the class (4.5) 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 \( \mu_p \) correspond to dominated strategies. We refer, in this set-up, to the subset of strategies that satisfy \( \mu_p \in [0, -\alpha_p] \) as \textit{level-1 rationalizable strategies}. Note, as before, that these \( \mu \)'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 (4.7). 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,
$$\mu_p = -\alpha_p \int_0^{-\alpha_p} E[\mathbb{1}\{t_{-p} \geq \mu\}|\mathcal{I}_p, \mu] d\hat{G}_p(\mu|\mathcal{I}_p)$$
where player $p$’s beliefs $\hat{G}_p(\cdot|\mathcal{I}_p)$ satisfy $\hat{G}_p(0|\mathcal{I}_p) = 0$ and $\hat{G}_p(-\alpha_p|\mathcal{I}_p) = 1$, i.e., 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. Hence, exploiting this monotonicity, it is easy to see that for an outside
observer, the subset of level-2 rationalizable strategies must satisfy
$$\mu_1 \in \left[-\alpha_1 E[\mathbb{1}\{t_2 \geq -\alpha_2\}|\mathcal{I}_1], -\alpha_1 E[\mathbb{1}\{t_2 \geq 0\}|\mathcal{I}_1]\right],$$
$$\mu_2 \in \left[-\alpha_2 E[\mathbb{1}\{t_1 \geq -\alpha_1\}|\mathcal{I}_2], -\alpha_2 E[\mathbb{1}\{t_1 \geq 0\}|\mathcal{I}_2]\right].$$
Level-k rational players are those whose beliefs are consistent with assuming that their oppo-
nents are level-(k-1) rational. Notice that this definition is a statement about a player’s
higher-order beliefs up to order $k - 1$. Specifically, any player who believes his opponent
undertakes (at least) $k - 1$ rounds of iterated deletion of dominated strategies in the con-
struction of his expected utility will be a level-k rational player. By induction, it is easy to
prove the following claim.

**Claim 1** If $\alpha_p \leq 0$, a strategy of the type $Y_p = \mathbb{1}\{t_p \geq \mu_p\}$ is level-k rationalizable if and
only if $\mu_1$ and $\mu_2$ satisfy
$$\mu_p \in [0, -\alpha_p] \equiv [\mu^L_{p,1}, \mu^U_{p,1}], \text{ for } k = 1 \text{ and } p \in \{1, 2\},$$
$$\mu_1 \in \left[-\alpha_1 E[\mathbb{1}\{t_2 \geq \mu^U_{2,k-1}\}|\mathcal{I}_1], -\alpha_1 E[\mathbb{1}\{t_2 \geq \mu^L_{2,k-1}\}|\mathcal{I}_1]\right] \equiv [\mu^L_{1,k}, \mu^U_{1,k}], \text{ for } k > 1 \ (4.8)$$
$$\mu_2 \in \left[-\alpha_2 E[\mathbb{1}\{t_1 \geq \mu^U_{1,k-1}\}|\mathcal{I}_2], -\alpha_2 E[\mathbb{1}\{t_1 \geq \mu^L_{1,k-1}\}|\mathcal{I}_2]\right] \equiv [\mu^L_{2,k}, \mu^U_{2,k}], \text{ for } k > 1.$$  

The bounds described in (4.8) contain any set of beliefs that can be rationalized after $k - 1$
rounds of iterated deletion of dominated strategies. We will present identification results
based on this entire range without additional restrictions about how level-k players actually
choose their beliefs from within this space of rationalizable beliefs.

**Remark 1** Any level-k rational player is also level-$k'$ rational for any $1 \leq k' \leq k - 1$. Also,
for $p \in \{1, 2\}$, with probability one, we have $[\mu^L_{p,k}, \mu^U_{p,k}] \subseteq [\mu^L_{p,k-1}, \mu^U_{p,k-1}]$ for any $k > 1,$
with strict inclusion if \( \alpha_p \neq 0 \) and if \( t_{-p} \) 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 player \( p \) conditions his beliefs on.

The two statements in Remark 1 follow because conditional on \( \mathcal{I}_p \), the support \( \mathcal{S}(\hat{G}_p) \) 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 \( \mathcal{S}(\hat{G}_p) \) would collapse to the singleton given by BNE beliefs as \( k \to \infty \). Whenever it is warranted, we will 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 (4.4) above to map the observed choice probabilities to the predicted ones. In particular, we can use the thresholds from Claim 1 above to construct a map between the model and the observable outcomes using (4.5) above. This is illustrated in Figure 5 for the case \( \mathcal{I}_p = t_p \) (players condition their beliefs exclusively on the realization of their own type). There, \( P_{t_1}(\cdot) \) abbreviates the conditional distribution of \( t_2|t_1 \), with \( P_{t_2}(\cdot) \) defined analogously. We see that as one moves from level 1 to level 2, the middle square shrinks. As we will see below, 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.](image)
Result 2 Let

$$\pi_p^L(1; \mathcal{I}_p) = 0, \quad \pi_p^U(1; \mathcal{I}_p) = 1 \quad \text{for } p = 1, 2,$$

and let

$$\pi^L_t(k; \mathcal{I}_1) = E[1\{t_2 + \alpha_2 \pi^U_2(k - 1; \mathcal{I}_2) \geq 0\}|\mathcal{I}_1], \quad \pi^U_t(k; \mathcal{I}_1) = E[1\{t_2 + \alpha_2 \pi^U_2(k - 1; \mathcal{I}_2) \geq 0\}|\mathcal{I}_1],$$

$$\pi^L_t(k; \mathcal{I}_2) = E[1\{t_1 + \alpha_1 \pi^U_1(k - 1; \mathcal{I}_1) \geq 0\}|\mathcal{I}_2], \quad \pi^U_t(k; \mathcal{I}_2) = E[1\{t_1 + \alpha_1 \pi^U_1(k - 1; \mathcal{I}_1) \geq 0\}|\mathcal{I}_2].$$

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

$$\mathcal{R}^p(k) = \left\{ Y_p = 1\{t_p + \alpha_p \pi_{-p}(\mathcal{I}_p) \geq 0\} : \pi_{-p} \in \left[\pi^L_{-p}(k; \cdot), \pi^U_{-p}(k; \cdot)\right]\right\} \text{ for } p = 1, 2.$$

In the next section, we parametrize 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 now on, we will express $t_p = X'_p/\beta_p - \varepsilon_p$, where $X_p$ is observable to the econometrician, $\varepsilon_p$ is not, and $\beta_p$ must be estimated (along with $\alpha_p$, the strategic-interaction parameter for player $p$). Player $p$ observes the realization of his own $X_p$ and $\varepsilon_p$, 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, we denote the vector of signals used by player $p$ to condition his beliefs by $\mathcal{I}_p$. Throughout, we will assume $(\varepsilon_1, \varepsilon_2)$ to be continuously distributed, with scale normalized to one and unbounded support. For simplicity, we will assume that $\varepsilon_1$ is independent of $\varepsilon_2$ and we will denote their CDF as $H_p(\cdot)$ for $p = 1, 2$.

We can extend the results that follow and obtain constructive identification results for the case where $\varepsilon_1$ and $\varepsilon_2$ are statistically interdependent as long as their joint distribution is assumed to be known possibly up to a finite-dimensional parameter, which would measure their interdependence. For simplicity, we limit ourselves here to the case where beliefs are conditioned on observables to the researcher, that is, $\mathcal{I}_p$ is observable. Without ad-hoc assumptions (e.g., that players somehow ignore the interdependence between $\varepsilon_1$ and $\varepsilon_2$ in the construction of their beliefs), observable signals $\mathcal{I}_p$ requires independence between $\varepsilon_1$ and $\varepsilon_2$. We will define the identified set of parameters and then we will provide an objective function that can be used to construct the identified set. This function will depend on the level $k$ of rationality that the econometrician assumes ex-ante. We then discuss the identification of $k$.

After that, we provide a set of sufficient conditions that will guarantee point identification under some assumptions. These point identification results provide insights into the kind of “variation” that is 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 to be 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 will assume that \( H_1 \) and \( H_2 \) are common knowledge among the players, and we will also assume that the econometrician knows these common prior distributions. We will assume that players use the true distributions as their priors for payoff covariates \( X_p \) and signals \( I_p \), both of which are observed by the econometrician. Implicitly, we also assume that the true values of \( \beta_p \), \( \alpha_p \) 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 will constitute the foundation for our identification results. As we mentioned previously, the results that follow can be extended to the case where \( \varepsilon_1 \) and \( \varepsilon_2 \) are statistically interdependent as long as their joint distribution is assumed to be known to the econometrician possibly up to a finite-dimensional parameter. For simplicity we focus on the case where \( \varepsilon_1 \) and \( \varepsilon_2 \) are independent.

Iterated dominance and bounds for beliefs

For ease of exposition we assume that both players condition on the same vector of signals which we will denote by \( \mathcal{I} \). This would include the case where the only source of private information for player \( p \) is \( \varepsilon_p \) and \( \mathcal{I} = X_1 \cup X_2 \). We will go back 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

\[
\pi^L_p(\theta|k = 1, \mathcal{I}) = 0, \quad \pi^U_p(\theta|k = 1, \mathcal{I}) = 1, \quad \text{and for } k \geq 2 \text{ let}
\]

\[
\pi^L_1(\theta|k, \mathcal{I}) = E \left[ H_1(X'_1\beta_1 + \alpha_1\pi^U_2(\theta|k-1, \mathcal{I})) | \mathcal{I} \right] ; \quad \pi^U_1(\theta|k, \mathcal{I}) = E \left[ H_1(X'_1\beta_1 + \alpha_1\pi^L_2(\theta|k-1, \mathcal{I})) | \mathcal{I} \right] \\
\pi^L_2(\theta|k, \mathcal{I}) = E \left[ H_2(X'_2\beta_2 + \alpha_2\pi^U_1(\theta|k-1, \mathcal{I})) | \mathcal{I} \right] ; \quad \pi^U_2(\theta|k, \mathcal{I}) = E \left[ H_2(X'_2\beta_2 + \alpha_2\pi^L_1(\theta|k-1, \mathcal{I})) | \mathcal{I} \right]
\]

\[
(4.9)
\]

\( \pi^L_p(\theta|k, \mathcal{I}) \) 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 assumptions above, these bounds are identified for any \( \theta \) and \( k \). In the case where we want to allow for correlation in \( \varepsilon_1 \) and \( \varepsilon_2 \), the belief function for player \( p \) will depend on \( \varepsilon_p \) which would be part of a player specific information set and
would be the conditional CDF of $\varepsilon_p|\varepsilon_{-p}$. By induction, it is easy to show that

$$[\pi_{-p}^L(\theta|k; \mathcal{I}), \pi_{-p}^U(\theta|k; \mathcal{I})] \subseteq [\pi_{-p}^L(\theta|k-1; \mathcal{I}), \pi_{-p}^U(\theta|k-1; \mathcal{I})]$$

w.p.1 in $\mathcal{S}(\mathcal{I})$. (4.10)

This monotonic feature holds even if players condition on different information sets. Moreover, the inclusion in (4.10) is strict if the strategic-interaction coefficients are nonzero, and if $\varepsilon_p$ 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 $\theta$ and $k \in \{2, 3, 4, 5\}$. The left display in Figure 6 provides the belief iterations with a unique BNE while the right display shows the iterations with multiple BNE.

Figure 6: Rationalizable Beliefs for $k = 2, 3, 4$ and $5$

---

Bounds for level-$k$ rationalizable beliefs when $\mathcal{I}_1 = \mathcal{I}_2 \equiv \mathcal{I}$ (players condition on the same set of signals). Vertical axis shows level-$k$ rationalizable bounds for player 1’s beliefs about $\Pr(Y_2 = 1|\mathcal{I})$. Horizontal axis shows the equivalent objects for player 2. The graphs correspond to a particular realization $\mathcal{I}$ and a given parameter value $\theta$.

**Identified set for $\theta$ based on level-$k$ rationality**

Let $W_p = X_p \cup \mathcal{I}$. It follows from the discussion in Subsection 4.2 (see Result 2) that the identified set for $\theta$ under the assumption that players are level-$k$ rational is given by

$$\Theta_I(k) = \left\{ \theta \in \Theta : \exists \pi_1(\cdot), \pi_2(\cdot) \in [\pi_1^L(\theta|k; \cdot), \pi_1^U(\theta|k; \cdot)] \times [\pi_2^L(\theta|k; \cdot), \pi_2^U(\theta|k; \cdot)] \text{ such that } E[Y_p|W_p] = H_p(X_p\beta_p + \alpha_p\pi_{-p}(\mathcal{I})) \text{ w.p.1. for } p = 1, 2 \right\}.$$
We will exploit the fact that under our assumptions the bounds for level-k rational beliefs are identified in order to characterize a set \( \Theta(k) \) that includes \( \Theta_I(k) \). Our characterization will be constructive and will be based on conditional moment inequalities. To proceed, note that player \( p \) is level-k rational iff

\[
\mathbb{I}\{ X_p'\beta_p + \alpha_p \pi_{-p}^U(\theta|k; \mathcal{I}) \geq \varepsilon_p \} \leq \mathbb{I}\{ Y_p = 1 \} \leq \mathbb{I}\{ X_p'\beta_p + \alpha_p \pi_{-p}^L(\theta|k; \mathcal{I}) \geq \varepsilon_p \} \text{ w.p.1.}
\]

These inequalities must hold with probability one for all realizations of \((X_p, \varepsilon_p, \mathcal{I})\). It follows that level-k rational players must satisfy

\[
H_p(X_p'\beta_p + \alpha_p \pi_{-p}^U(\theta|k; \mathcal{I})) \leq E[Y_p|W_p] \leq H_p(X_p'\beta_p + \alpha_p \pi_{-p}^L(\theta|k; \mathcal{I})) \text{ w.p.1.}
\]

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

\[
\Theta(k) = \left\{ \theta \in \Theta : H_p(X_p'\beta_p + \alpha_p \pi_{-p}^U(\theta|k; \mathcal{I})) \leq E[Y_p|W_p] \leq H_p(X_p'\beta_p + \alpha_p \pi_{-p}^L(\theta|k; \mathcal{I})) \text{ w.p.1, } p = 1, 2 \right\}.
\]

(4.12)

Clearly, if players are level-k rational, we have \( \Theta_I(k) \subseteq \Theta(k) \). If \( X_p \in \mathcal{I} \) for both players, it is easy to show that \( \Theta_I(k) = \Theta(k) \). This follows because the set \([\pi_I^U(\theta|k; \cdot), \pi_I^L(\theta|k; \cdot)] \times [\pi_2^U(\theta|k; \cdot), \pi_2^L(\theta|k; \cdot)]\) is connected, and the distributions \( H_1, H_2 \) are continuous. The characterization \( \Theta(k) \) is constructive and it will be the one we will use. To allow for the case where \( W_p \) has continuous support, we will 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

\[
\Lambda_p(\theta|a, b; k) = E\left[\left(1 - \mathbb{I}\{ H_p(X_p'\beta_p + \alpha_p \pi_{-p}^U(\theta|k; \mathcal{I})) \leq Pr(Y_p = 1|W_p) \leq H_p(X_p'\beta_p + \alpha_p \pi_{-p}^L(\theta|k; \mathcal{I})) \} \right) \right]
\times \mathbb{I}\{a \leq W_p \leq b\};
\]

\[
\Gamma_p(\theta|k) = \int \int \Lambda_p(\theta|a, b; k)dF_{W_p}(a)dF_{W_p}(b); \quad \Gamma(\theta|k) = \left(\Gamma_1(\theta|k), \Gamma_2(\theta|k)\right).
\]

(4.13)

where the inequality \( a \leq W_p \leq b \) is element-wise and \( W_p \sim F_{W_p}(\cdot) \). Take any \( 2 \times 2 \) positive definite matrix \( \Omega \). The set in (4.12) can be expressed as

\[
\Theta(k) = \left\{ \theta \in \Theta : \Gamma(\theta|k)'\Omega\Gamma(\theta|k) = 0 \right\}.
\]

(4.14)
This will be the definition of identified set for $\theta$ that we will use under the assumption that all players in the game are level-k rational. More precisely, going back 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 $k$. Notice also that as compared with the Bayesian Nash solution, here one does 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 below 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 $k = 1$, one does not need to specify the common prior assumption since beliefs here play no role. Hence, results will be robust to this assumption. However, depending on the magnitude of the $\alpha_p$’s, the bounds on choice probabilities predicted by such a model (where $k = 1$) can be wide.

Under certain conditions, the identified set in (4.14) would consist only of $\theta_0$, the true parameter value. An example would be a case in which there exist realizations of the vector of signals $I$ where the players are “forced” to take one of their actions with probability one regardless of their beliefs. To be concrete, suppose the linear index $X_p^t\beta_p$ has unbounded support for both players and suppose 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 $I$ is such that there exist regions of $S(I)$ such that $S(X_p^t\beta_p|I)$ is concentrated around arbitrarily large positive or arbitrarily large negative values, the identified set $\Theta(k = 2)$ defined in (4.14) would collapse to a singleton $\theta_0$, the true parameter value. We refer to this as a case of “informative signals” and formalize this point-identification result in the following subsection.

4.4 Sufficient point identification conditions

In this section, we study the problem of point identification of the parameter of interests in the game above. In particular, we provide sufficient point identification conditions for level-1 rational play and for levels $k > 1$. These conditions can provide insights about what is required to shrink the identified set to a point (or a vector). Here, we allow for the information sets to be different, i.e., that player $p$ conditions on $I_p$ when making decisions
and allow for exclusion restrictions where $I_1 \neq I_2$. 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 $\theta = (\theta_1, \theta_2)$, we have the following identification result.

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

**A1.i** 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 $q \in \mathbb{R}^{\text{dim}(X_{\ell,p})}$, there exists $C_{b,q,m} > 0$ such that

$$
\Pr(\epsilon_p \leq bX_{\ell,p} + q'X_{\ell,p}|X) > m \forall X_{\ell,p} : \text{sign}(b) \cdot X_{\ell,p} > C_{b,q,m}.
$$

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

$$
\Pr\left(|X_{d,p}'(\beta_{d,p} - \tilde{\beta}_{d,p})| > |\alpha_p|\left|X \setminus X_{d,p}\right) > 0.
$$

If all we know is that players are level-1 rational:

(a) If (A1.i) holds, the coefficients $\beta_{\ell,p}$ are identified.

(b) If (A1.ii) holds, 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 $X_\Delta \in \mathcal{S}(X_p)$ such that $\Pr(Y_p = 1|X) < \Pr(\epsilon_p \leq X_p'\beta_{q_0} + \alpha_{p_0}|X) + \Delta$ whenever $X_p \in X_\Delta$. If (A1.i-ii) holds and player $p$ is pessimistic with positive probability, the identified set for $\alpha_p$ is $\{\alpha_p \in \Theta : \alpha_p \leq \alpha_{p_0}\}$.

The results in Theorem 1 imposed no restrictions on $I_p$. In particular, players can condition their beliefs on unobservables to the econometrician. A special case of condition (A1.i) is when $\epsilon_p$ is independent of $X$. The condition in (A1.ii) 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 cases $\alpha_p \geq 0$ and $\alpha_1\alpha_2 \leq 0$. The proof of the above theorem 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 if a higher degree of rationality will help in the task of point identifying $\alpha_p$. To simplify the analysis, we will assume from now on that $\varepsilon_p$ is independent of $X$ and of $I \equiv (I_1, I_2)$. This assumption could be replaced with one along the lines of (A1) in Theorem 1. We make the assumption that $I$ is observed by the econometrician. We will relax this assumption in a later section. Again, let the common prior assumption be denoted by $H_p(.)$. The beliefs of the players for any level $k$ rationality can be constructed as we did in the previous section. Our point identification sufficient conditions are summarized in theorem 2 below.

**Theorem 2** Suppose there exists a subset $X_1^* \subseteq S(X_1)$ where $X_1$ has full-column rank such that for any $X_1 \in X_1^*$, $\varepsilon > 0$ and $\theta_2 \in \Theta$, there exist $\mathcal{S}_{1}^* \subseteq S(I_1|X_1)$ and $\mathcal{S}_{1}^{**} \subseteq S(I_1|X_1)$ such that

For all $I_1 \in \mathcal{S}_{1}^*$,

$$\max \left\{ 1 - E[H_2(X_2^2\beta_2 + \Delta_2)|I_1] , E[H_2(X_2^2\beta_2 + \Delta_2)|I_1] - E[H_2(X_2^2\beta_2 + \Delta_2 + \alpha_2)|I_1] \right\} < \varepsilon,$$

For all $I_1 \in \mathcal{S}_{1}^{**}$,

$$\max \left\{ E[H_2(X_2^2\beta_2 + \Delta_2 + \alpha_2)|I_1] , E[H_2(X_2^2\beta_2 + \Delta_2)|I_1] - E[H_2(X_2^2\beta_2 + \Delta_2 + \alpha_2)|I_1] \right\} < \varepsilon.$$

(4.17)

A special case in which (4.17) holds is when there exists $X_2_0 \in (X_2 \cap W_1)$ with nonzero coefficient in $\Theta$ such that $X_2_0$ has unbounded support conditional on $(X_2 \cup W_1) \setminus X_2_0$. We could refer to (4.17) as an “informative signal” condition. Note that implicit in (4.17) is an exclusion restriction in the parameter space that precludes $\beta_2 = 0$ for any $\theta_2 \in \Theta$. If (4.17) holds, then for any $\theta \in \Theta$ such that $\theta_1 \neq \theta_{10}$, there exists either $W_1^* \subseteq S(W_1)$ or $W_1^{**} \subseteq S(W_1)$ such that

$$H_1(X_1^1\beta_1 + \Delta_1 + \alpha_1\pi_2^L(\theta|k; I_1)) < H_1(X_1^1\beta_{10} + \Delta_{10} + \alpha_{10}\pi_2^L(\theta_0|k; I_1)) \forall W_1 \in W_1^*, \ k \geq 2,$$

$$H_1(X_1^1\beta_1 + \Delta_1 + \alpha_1\pi_2^U(\theta|k; I_1)) > H_1(X_1^1\beta_{10} + \Delta_{10} + \alpha_{10}\pi_2^U(\theta_0|k; I_1)) \forall W_1 \in W_1^{**}, \ k \geq 2.$$

(4.18)

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 $\theta$ will be disjoint with those of $\theta_0$ with positive probability. As a consequence, if (4.17) holds and the population of Players 1 are at least level-2 rational, $\theta_{10}$ is identified. By symmetry, $\theta_{20}$ will be point-identified if the above 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 $\varepsilon_p$, Figures 7 and 8 illustrate four graphical examples of how the "informative signals" condition (4.17) 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 $\Delta_1$ is subsumed in $X'_1\beta_1$ in the labels of these figures.

4.5 On the 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 $\varepsilon_p$ relative to that of $X'_p\beta_p$, it is not possible to reject a value of $k$ on the basis of observed choices. Our setup however 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 one, 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 $\theta$. If all players are at least level-2 rational and the conditions for point-identification of $\theta$ described in Theorem 2 hold, $k_0$ would be point-identified because $Q(\theta_0|k) = 0$ if and only if $k \leq k_0$, where $Q(\theta|k)$ is defined in Equation (4.14). To see why this is not true when $\theta$ is set-identified, go back to parts (i)-(ii) following Equation (4.19). Otherwise, if the conditions for Theorem 2 do not hold, suppose we maintain the assumption $k_0 \geq 1$ (the only interesting case). One can start with $k = 1$ and construct $\Theta(1)$ (as defined in 4.14). Next, for any $k \geq 2$ define

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

(4.19)

where $Q(\theta|k)$ is as defined in (4.14). Then,

(i) $Q(k) = 0$ for all $k \leq k_0$. However, $Q(k) = 0$ does not imply $k \leq k_0$.

(ii) $Q(k) > 0$ implies $k > k_0$.

Suppose that different observations in the data set correspond to a game with a different level of rationality, then if $Q(k) > 0$ and $Q(k-1) = 0$, one would reject the hypothesis that all the population is at least level $k$ rational. If we assume ex-ante that $k_0 \geq k > 1$ we could simply replace $\Theta(1)$ with $\Theta(k)$ in the definition of $Q(k)$ in Equation (4.19). Alternatively, in settings where at least a subset of the structural parameter $\theta$ is known (e.g., experiments), we could evaluate if 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 $I_p$ be the signal player $p$ uses to condition his beliefs about his opponent’s expected choice, and let $I = (I_1, I_2)$. The set of Bayesian-Nash equilibria (BNE) is defined as any pair $(\pi_1^*(I_2), \pi_2^*(I_1))$ that satisfies

$$\pi_1^*(I_2) = E[H_1(X'_1\beta_1 + \alpha_1\pi_2^*(I_1))|I_2]$$
$$\pi_2^*(I_1) = E[H_2(X'_2\beta_2 + \alpha_2\pi_1^*(I_2))|I_1]$$

(4.20)

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

**Proposition 1** Let

$$R(I; k) = \left[\pi^k_1(\theta | I_2), \pi^U_1(\theta | I_2)\right] \times \left[\pi^L_2(\theta | I_1), \pi^U_2(\theta | I_1)\right]$$

denote the set of level-$k$ rationalizable beliefs. Then, with probability one, the BNE set described in (4.20) is contained in $R(I; k)$ for any $k \geq 1$.

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

**Claim 2** Let $\pi^*(I) \equiv (\pi_1^*(I_2), \pi_2^*(I_1))$ be any BNE. Then, for any $k \geq 1$, with probability one, we have: $\pi^*(I) \in R(I; k)$ implies $\pi^*(I) \in R(I; k + 1)$ w.p.1.

**Proof of Claim 2:** If $\alpha_p = 0$ for $p = 1$ or $p = 2$, the result follows trivially. Suppose $\alpha_1 = 0$, then $\pi^U_1(\theta | I_2) = \pi^L_1(\theta | I_2) = \pi^U_1(I_2) = \pi^L_1(I_2) = E[H_1(X'_1\beta_1) | I_2]$ and $\pi^L_2(\theta | I_1) = \pi^U_2(\theta | k; I_1) = \pi^U_1(I_1) = E[H_2(X'_2\beta_2 + \alpha_2\pi^U_1(I_2)) | I_1]$ for all $k \geq 1$. We focus on the case $\alpha_p < 0$ for $p = 1, 2$. Now, suppose $\pi^*(I) \in R(I; k)$, but $\pi^*(I) \notin R(I; k)$. Suppose for example that $\pi^U_1(\theta | k + 1; I_2) > \pi^U_1(I_2)$. Since $\alpha_1 < 0$, this can be true if and only if

$$E[H_1(X'_1\beta_1 + \alpha_1\pi^U_2(\theta | k; I_2)) | I_2] > E[H_1(X'_1\beta_1 + \alpha_1\pi^U_2(I_1)) | I_2]$$

For this inequality to be satisfied, it cannot be the case that $\pi^U_2(I_1) \leq \pi^U_2(\theta | k; I_1)$. But this violates the assumption that $\pi^*(I) \in R(I; k)$. Therefore, we must have $\pi^U_1(\theta | k + 1; I_2) \leq \pi^U_1(I_2)$. Suppose now that $\pi^U_1(\theta | k; I_2) < \pi^U_1(I_2)$. This can be true if and only if

$$E[H_1(X'_1\beta_1 + \alpha_1\pi^U_2(\theta | k; I_2)) | I_2] < E[H_1(X'_1\beta_1 + \alpha_1\pi^U_2(I_1)) | I_2]$$

27
For this inequality to be satisfied, it cannot be the case that \( \pi^*(\mathcal{I}_1) \geq \pi^*_2(\theta|k; \mathcal{I}_1) \). Once again, this violates the assumption \( \pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k) \). Therefore, we must have \( \pi^*_1(\theta|k + 1; \mathcal{I}_2) \geq \pi^*_1(\mathcal{I}_2) \). These results imply that we must have \( \pi^*_1(\theta|k + 1; \mathcal{I}_2) \leq \pi^*_1(\mathcal{I}_2) \leq \pi^*_1(\theta|k + 1; \mathcal{I}_2) \).

Following the same steps we can establish that we must have \( \pi^*_2(\theta|k + 1; \mathcal{I}_1) \leq \pi^*_2(\mathcal{I}_1) \leq \pi^*_2(\theta|k + 1; \mathcal{I}_1) \). Combined, these yield \( \pi^*(\mathcal{I}) \in \mathcal{R}(\mathcal{I}; k + 1) \) as claimed.

\[ \square \]

**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 \( k \geq 1 \).

\[ \square \]

**BNE vs Rationalizability: Identification**

Naturally, it is always guaranteed that one gets a weakly smaller identified set with BNE assumptions since 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 the largest equilibria. In the case of a unique equilibrium, one can see that in the above 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 a consequence of the normal-form payoff parameterization of the game. In addition, in the simple example above, predicted outcomes based on rationality of order \( k \), for any \( k \), are a lot easier to solve for since they do not require solutions to fixed point problems especially in cases of multiple equilibria.

### 5 Identification in First Price IPV 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 it 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 Bayesian-Nash equilibrium (BNE), nonparametric point identification for this distribution has been established for example by Guerre, Perrigne, and...
Vuon (1999). Hence, 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), abbreviated henceforth as BS. Other strategic assumptions that can be used and that deliver qualitatively different results than BS’s interim rationalizability is the $P$ dominance concept introduced for auctions setups by Dekel and Wolinsky (2003) and more recently Crawford and Iriberri (2007b). Here, we just highlight what can be learned with the BS setup and compare those to Bayesian Nash equilibrium. 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 will see, 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 zero) and allow, in principle, the case $\omega = +\infty$. Assume for the moment that the seller’s reservation price $p_0$ is equal to zero. We will explicitly introduce a strictly positive reservation price below.

**Assumptions about bidders’ beliefs**

Following BS, we assume that bidders expect all positive bids to win with strictly positive probability and this is common knowledge. This condition will ensure that it is common knowledge that no bidder will bid beyond his/her valuation irrespective of his beliefs. It also implies that every bidder with nonzero private value will submit a strictly positive bid. Therefore, with probability one the number of potential bidders $N$ is equal to the number of actual bidders (only a bidder with valuation equal to zero 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

$$\mathcal{B} = \left\{ b : [0, \omega) \rightarrow \mathbb{R}_+ : b(v) \leq v, \text{ and } v > v' \Rightarrow b(v) > b(v') \right\}. \quad (5.21)$$

We will let $\mathcal{N}$ denote the number of potential bidders in the population and denote $\mathcal{B}_{-i} = \mathcal{B}^{N-1}$. 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. A conjecture by bidder $i$ is a degenerate belief that assigns probability mass one to a singleton $\{b_j\}_{j \neq i} \in \mathcal{B}_{-i}$. 

29
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 will see, restricting attention to beliefs in $\mathcal{B}$ will yield rationalizable upper bounds for bids which also belong in $\mathcal{B}$. It also simplifies the analysis, for example, by ruling out ties in the characterization of players’ expected utility. Finally, as we will argue below (and is formally shown in BS), restricting attention to beliefs in $\mathcal{B}$ will imply that Bayesian-Nash equilibrium (BNE) optimal bids are always rationalizable.

5.1 Implications of level-k rationality in bids

Here, we follow the setup in BS, and our notation will differ from that of previous sections. We have a population of $\mathcal{N}$ risk-neutral potential buyers, bidding simultaneously for a single object. With a zero reservation price, 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 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 we focus on assigns probability zero to ties. Therefore, the decision problem for bidder $i$ can be expressed as

$$\max_{b \geq 0} (v_i - b) \hat{\Pr}_i \left[ \max_{j \neq i} b(v_j) \leq b \right], \tag{5.22}$$

where $\hat{\Pr}_i(\cdot)$ denotes bidder $i$’s subjective probability, derived from his beliefs and knowledge of $F_0(\cdot)$. For level-1 rational bidding, any bidder $i$ whose bids satisfy

$$b \leq v_i \equiv \overline{B}_1(v_i; \mathcal{N}) \text{ w.p.1} \tag{5.23}$$

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 $\mathcal{B}_{-i}$. Hence we have

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

This was proved in BS where they also show that the bound is sharp, i.e., 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 since in this setup, one cannot bound the bids from below. This 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 will derive strategies for identification of $F(\cdot)$ based on the BS results, but these strategies can be easily adapted to other strategic setups like ones 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 $B_{-i}$. The most pessimistic assessment in $B_{-i}$ is given by the conjecture $b(v_j) = B_1(v_j; N) = v_j$ for all $j \neq i$ (the upper bound for bids for level-1 rational bidders). Since bidder $i$ knows $F_0$, his optimal expected utility for this assessment is

$$
\max_{b \geq 0} (v_i - b) \Pr \left[ \max_{j \neq i} B_1(v_j; N) \leq b \right] = \max_{b \geq 0} (v_i - b) \Pr \left[ \max_{j \neq i} v_j \leq b \right] = \max_{b \geq 0} (v_i - b) F_0(b)^{N-1} = \pi^*_2(v_i; N)
$$

where $\pi^*_2(v_i; N)$ is the lower bound for optimal expected utility (5.22) for all beliefs in $B_{-i}$. 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 one). Any bid submitted by a rational (i.e., expected-utility maximizer) bidder with beliefs in $B_{-i}$ must satisfy

$$
v_i - b \geq \pi^*_2(v_i; N) \Rightarrow b \leq v_i - \pi^*_2(v_i; N) = B_2(v_i; N) \quad \text{w.p.1} \quad (5.25)
$$

We refer to bidders who satisfy (5.25) as level-2 rational bidders. Given our assumptions, $B_2(v_i; N)$ is increasing, concave and satisfies $B_2(v_i; N) \leq B_1(v_i; N) = v_i$, with strict inequality for all $v_i > 0$. Therefore, $B_2 \in B$. Let $\overline{S}_2(\cdot; N)$ denote the inverse of $B_2(\cdot; N)$. We refer to level-3 rational bidders as those whose beliefs incorporate the level-2 upper bound (5.25). The most pessimistic assessment for level-3 rational bidders is the conjecture $b(v_j) = B_2(v_j; N)$ for all $j \neq i$. The optimal expected utility for this pessimistic assessment is

$$
\max_{b \geq 0} (v_i - b) \Pr \left[ \max_{j \neq i} B_2(v_j; N) \leq b \right] = \max_{b \geq 0} (v_i - b) F_0(\overline{S}_2(b; N))^{N-1} = \pi^*_3(v_i; N) \quad (5.26)
$$

Using the same logic that led to (5.25), the set of rationalizable bids for level-3 rational bidders must satisfy

$$
v_i - b \geq \pi^*_3(v_i; N) \Rightarrow b \leq v_i - \pi^*_3(v_i; N) = B_3(v_i; N) \quad \text{w.p.1.} \quad (5.27)
$$
The level-3 upper bound for rationalizable bids, $\overline{B}_3(\cdot;\mathcal{N})$ is increasing, 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 $\overline{B}_2(v_i;\mathcal{N}) = v_i - \pi^*_i(v_i;\mathcal{N}) \equiv \overline{B}_1(v_i;\mathcal{N}) - \pi^*_i(v_i;\mathcal{N})$. Therefore, for any $b$ we have $\Pr[\max_j \overline{B}_2(v_j;\mathcal{N}) \leq b] \geq \Pr[\max_j \overline{B}_1(v_j;\mathcal{N}) \leq b]$. Immediately, this implies $\pi^*_i(\cdot;\mathcal{N}) \geq \pi^*_i(\cdot;\mathcal{N})$ and therefore $\overline{B}_3(\cdot;\mathcal{N}) \leq \overline{B}_2(\cdot;\mathcal{N})$. Since $F_0(\cdot)$ is not assumed to have point masses, all the above inequalities are strict for any $v_i > 0$. Proceeding iteratively, the level-k bound for rationalizable bids is given by

$$b_i \leq v_i - \pi^*_i(v_i;\mathcal{N}) \equiv \overline{B}_k(v_i;\mathcal{N}) \text{ w.p.1.,}$$

where

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

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, 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 have shown 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 will be consistent with this type of behavior.

**Example.** Suppose private values are exponentially distributed, with $F_0(v) = 1 - \exp\{-\theta v\}$ and $\theta > 0$. We have $F_0(v)/f_0(v) = \frac{1 - \exp\{-\theta v\}}{\theta \exp\{-\theta v\}} = \frac{1}{\theta} \exp\{\theta v\} - \frac{1}{\theta}$, 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 $\theta = -0.25$, $\mathcal{N} = 2$ (two bidders) and $k = 1, 2, 3, 4$. This graphical example illustrate the features described above for these bounds. Namely, $\overline{B}_k(\cdot;\mathcal{N})$, is continuous, increasing, concave, 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 $\overline{B}_4(v;\mathcal{N})$.

### 5.2 Identification with level-$k$ rationality in a parametric model

This section exploits the above bounds to learn about the distribution of valuation given a random sample of bids. We focus first on the hypothetical case where there is no reserve price set by the seller and, for each auction, we observe all bids submitted and we also know
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

$$F_{v}^{\Theta} = \left\{ F(\cdot; \theta) : \theta \in \Theta, \text{ and } F_0(\cdot) = F(\cdot; \theta_0) \text{ for some } \theta_0 \in \Theta \right\} \quad (5.29)$$

Here, one can also think of $\Theta$ as a set of functions and hence the above definition accommodates nonparametric analysis. Denote the level-k upper bound that corresponds to $F(\cdot; \theta)$ by $B_k(\cdot; N|\theta)$.

**Level-1 rationality:** For rationality of level 1, the game predicts that

$$0 \leq b_l^i \leq v_l^i \text{ for all } i = 1, \ldots, N \quad l = 1, \ldots, L$$

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, one can introduce auction heterogeneity that is observed). This implies that

$$F_0(t; \theta) \equiv P(v \leq t) \leq P(b \leq t) \equiv G_0(t)$$
So, with the first level of rationality, we can bound the valuation distribution above by the observed distribution of the bids. Inference here will be handled below and is based on replacing the observed bids distribution with its consistent empirical analog.

**Level-k rationality:** Similarly to above, for level-\(k\) and any \(\theta \in \Theta\), we have

\[
0 \leq b_i \leq v_i - \pi^*(v_i; N|\theta) \equiv B_k(v_i; N)
\]

for all \(i = 1, \ldots, N\) \(l = 1, \ldots, L\).

Hence, this means that if bidders are level-k rational,

\[
F(S_k(t; N|\theta_0); \theta_0) \leq P(b \leq t) \equiv G_b(t)
\]

where, as before, \(S_k\) denotes the inverse function of \(B_k\). Here, the bound is a bit more complicated since the function \(S\) depends also on \(F_0\). Using the notation in Section 2, the space of strategies (bidding functions) for level-k rational players is

\[
R^i(k) = \{b \in B : b(\cdot) \leq B_k(\cdot; N|\theta)\}.
\]

As we did in Subsection 4.3, will characterize the identified set for \(\theta\) based on level-k rationality in terms of an objective function.

**Proposition 2** Suppose \(F_0\) belongs to a space of distribution functions as described in (5.29). Moreover, suppose we have a random sample of size \(L\) of auctions each of which has \(N\) bidders and where we observe all bids. Take \(k \in \mathbb{N}^+\), and let

\[
\Lambda(\theta|a, c; k) = \int \left(1 - \mathbf{1}\{F_b(b) \geq F(S_k(b; N|\theta); \theta)\}\right) \mathbf{1}\{a \leq b \leq c\} dF_b(b),
\]

\[
\Gamma(\theta|k) = \int \int \Lambda(\theta|a, c; k) dF_b(a) dF_b(c) \quad (5.30)
\]

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

\[
\Theta(k) = \left\{\theta \in \Theta : \Gamma(\theta|k)^2 = 0\right\}.
\]

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

**Assumption B1** Suppose 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

\[
\Pr\left(b_i \leq B_{k_0}(v_i; N|\theta_0)\right) = 1, \quad \text{and} \quad \Pr\left(b_i = B_{k_0}(v_i; N|\theta_0)\right) > 0.
\]
Proposition 3 Suppose Assumption B1 holds and let $\Theta(k_0)$ be as defined in Proposition 2. For $\theta \in \Theta$ let

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

(5.31)

Then, the identified set is

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

(5.32)

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

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

(5.33)

then $\Theta_0^* = \{\overline{\theta}\}$ and consequently, $\theta_0 = \overline{\theta}$. Under Assumption B1, $\theta \notin \Theta(k_0)$ implies $\theta \neq \theta_0$ and $\theta \in \Theta(k_0)$ holds only if $\theta \notin F^c(\theta_0)$. Suppose we have $\theta, \theta' \in \Theta(k_0)$ such that $\theta' \in F^c(\theta)$. Then, it cannot be the case that $\theta = \theta_0$ because $\theta' \in F^c(\theta_0)$ would imply $\theta' \notin \Theta(k_0)$. Thus, any such $\theta$ can be discarded as the true $\theta_0$. To see how this result is constructive, suppose $F^\Theta$ is a space of exponentially distributed valuations, Assumption B1 holds and we find that the largest value of $\Theta(k_0)$ is $\overline{\theta} < \infty$. This would immediately imply $\theta_0 = \overline{\theta}$. Figure 10 illustrates this result for the exponential distribution. As we can see there, if Assumption B1 holds with $k_0 = 2$ and if we knew that $\{0.25, 0.50, 0.75, 1.00, 1.25\} \subset \Theta(k_0)$, it would follow immediately that $\theta_0 \geq 1.25$. More generally, as in previous sections, the characterization of the identified set $\Theta(k)$ is amenable to recent set inference methods.

Remark 3 Let

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

(5.34)

Given our assumptions, the results in (BS) can be used to show that $\overline{B}_\infty(\cdot;N|\theta)$ exists, and is a continuous, increasing, concave and invertible mapping that satisfies $\overline{B}_\infty(\cdot;N|\theta) \geq b^\text{BNE}(\cdot;N|\theta)$. Notice that, unlike the incomplete information game in Section ??, rationalizable behavior here 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, we must have $b_i \leq \overline{B}_\infty(v_i;N|\theta)$ w.p.1. The results in Proposition 3 would follow if the conditions stated there hold for the mapping $\overline{B}_\infty(\cdot;N|\theta)$.
Figure 10: If Assumption B1 holds with $k_0 = 2$ and if we knew that \( \{0.25, 0.50, 0.75, 1.00, 1.25\} \subset \Theta(k_0) \), it would follow immediately that $\theta_0 \geq 1.25$.

5.2.1 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 until the next section. In particular, we observe

$$b^* = \max_{i=1, \ldots, N} b_i. \quad (5.35)$$

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

$$b^* = \max_{i=1, \ldots, N} b_i \leq \max_{i=1, \ldots, N} \overline{B}_k(v_i; N|\theta_0) = \overline{B}_k(\max_{i=1, \ldots, N} v_i; N|\theta_0) \equiv \overline{B}_k(v^*; N|\theta_0). \quad (5.36)$$

Then, we must have

$$\Pr(b^* \leq b) \geq \Pr(\overline{B}_k(v^*; N|\theta_0) \leq b) \quad \forall b \in \mathbb{R}. \quad (5.37)$$

Since private values are iid, it follows that $v^* \sim F(\cdot; \theta_0)^N$. Let $F_{b^*}(\cdot)$ denote the distribution function of $b^*$, the highest bid. Equation (5.37) becomes

$$F_{b^*}(b) \geq F(\overline{B}_k(b; N|\theta_0); \theta_0)^N \quad \forall b \in \mathbb{R}. \quad (5.38)$$
where, as before, \( \overline{S}_k(\cdot; N|\theta) \) denotes the inverse function of the upper bound \( \overline{B}_k(\cdot; N|\theta) \). Clearly, by the nondecreasing properties of distribution functions, (5.38) holds for all \( b \in \mathbb{R} \) if and only if it holds for all \( b \in S(b^*) \) (the support of \( b^* \)). We conclude that this implies

\[
F_{b^*}(b) \geq F(\overline{S}_k(b; N|\theta_0); \theta_0)^N \quad \forall b \in S(b^*). \tag{5.39}
\]

Equation (5.39) can be used, as in the previous to conduct inference on the set of consistent models. To do that, a similar objective function as the one in Proposition 2 above can be used. The results in Proposition 3 would also follow if Assumption B1 holds for \( b^* \).

5.3 Introducing a binding reserve price

Suppose 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 \( B(p_0) \) denote the space of all Borel-measurable functions of the form

\[
\left\{ b: [0, \omega) \to \mathbb{R}_+ : b(v) < p_0 \ \forall \ v < p_0; \ b(p_0) = p_0, \ \text{and for all} \ v > p_0: b(v) \leq v, \ \text{and} \ v > v' \Rightarrow b(v) > b(v') \right\}.
\tag{5.40}
\]

We will let \( N \) denote the number of potential bidders in the population and denote \( B_{-i}(p_0) = B(p_0)^{N-1} \). Beliefs for bidder \( i \) are probability distributions defined over a sigma-algebra \( \Delta_{B_{-i}}(p_0) \), where this sigma-algebra is such that singletons in \( B_{-i} \) are measurable (see footnote 12). As before, conjectures are defined as degenerate beliefs that assigns probability mass one to a singleton \( \{b_j\}_{j \neq i} \in B_{-i} \). We maintain the assumption that \( F_0(\cdot) \) and \( N \) are common knowledge among potential bidders.

A consequence of a binding reserve price is that the number of potential bidders \( N \) may no longer 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 \(b_j < p_0\) w.p.1 for all potential bidders. As in the case of zero reservation price, restricting attention to beliefs in \(B(p_0)\) will yield rationalizable upper bounds which also belong in \(B(p_0)\). 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 \(B(p_0)\) will imply that Bayesian-Nash equilibrium (BNE) optimal bids are always rationalizable.

**Level-k rationalizable bids with a nonzero reserve price**

The construction of rationalizable upper bounds will follow the same interim-rationalizability steps as in Subsection 5.1. Any bidder \(i\) with \(v_i \geq p_0\) whose bids satisfy

\[
b \leq v_i \quad \text{w.p.1,}
\]

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

\[
\max_{b \geq p_0} (v_i - b) \hat{\Pr}_i \left[ \max \left\{ p_0, \max_{j \neq i} b(v_j) \right\} \leq b \right],
\]

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

\[
\max_{b \geq p_0} (v_i - b) F_0(b)^{N-1} = \pi^*_2(v_i; N, p_0),
\]

which follows because \(\hat{\Pr}_i \left[ \max \left\{ p_0, \max_{j \neq i} v_j \right\} \leq b \right] = F_0(b)^{N-1} \mathbb{1}\{b \geq p_0\}\) (recall that \(F_0, N\) and \(p_0\) are common knowledge among bidders). Using the same arguments that followed Equation (5.24), level-2 rational bidders with \(v_i \geq p_0\) must satisfy

\[
p_0 \leq b \leq v_i - \pi^*_2(v_i; N, p_0) \equiv \overline{B}_2(v_i; N, p_0).
\]

\(\overline{B}_2(v_i; N, p_0)\) 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 \(\overline{B}_2(p_0; N, p_0) = p_0\). In particular, the inverse function of \(\overline{B}_2(\cdot; N, p_0)\) is well-defined for all values and bids \(\geq p_0\). As before, we will denote this inverse function by \(\overline{S}_2(\cdot; N, p_0)\). Note that, in general, (5.43) has corner solutions. That is, there exists a range of valuations \(v_i > p_0\) such that \(\pi^*_2(v_i; N, p_0) = (v_i -

38
This, of course, will not impact the continuity, monotonicity and invertibility properties of the upper bound \( \overline{B}_2(\cdot; N, p_0) \) 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 range of valuations are irrelevant for the optimal decision process of bidders. Proceeding iteratively, the level-k bound for rationalizable bids is given by

\[
b_i \leq v_i - \pi^*_k(v_i; N, p_0) = \overline{B}_k(v_i; N, p_0),
\]

where \( \pi^*_k(v_i; N) = \max_{b \geq p_0} \left( v_i - b \right) F_0(\overline{S}_{k-1}(b; N, p_0))^{N-1} \),

(5.45)

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

5.3.1 Identification with level-k rationality when only winning bids are observed

If we replace Assumption B0 with B0’, all the results in Subsection 5.2.1 hold with a binding reserve price for all \( v_i \geq p_0 \) and \( b_i \geq p_0 \). Consider a semiparametric setting as the one described in (5.29), where the distribution of valuations is allowed to depend on the publicly observed reserve price \( p_0 \)

\[
F^{\Theta, p_0}_{\theta} = \left\{ F(\cdot; \theta, p_0) : \theta \in \Theta, \text{ and } F_0(\cdot; p_0) = F(\cdot; \theta_0, p_0) \text{ for some } \theta_0 \in \Theta \right\}
\]

(5.46)

Let \( \overline{B}_k(\cdot; N|\theta, p_0) \) denote the k-level upper bound for rationalizable bids that would be induced by a given distribution \( F(\cdot; \theta, p_0) \in F^{\Theta, p_0}_\theta \), let \( \overline{B}_k(\cdot; N|\theta, p_0) \) denote its inverse function. Let \( b^* \) denote the winning bid, and \( F_{b^*}(\cdot; p_0) \) denote its distribution function (given \( p_0 \)). Note that \( b^* = \max_{i=1, \ldots, N} b_i \), truncated from below at \( p_0 \). This automatic truncation ensures that the bounds in (5.45) are satisfied. As we mentioned previously, bids below \( p_0 \) may not satisfy these bounds. If bidders are level-k rational, for any reserve price \( p_0 \) we must have

\[
F_{b^*}(b; p_0) \geq F(\overline{S}_k(b; N|\theta_0, p_0); \theta_0, p_0)^N \quad \forall b \in S(b^*|p_0),
\]

(5.47)

where \( S(b^*|p_0) \) is the support of \( b^* \) given \( p_0 \). This result is the equivalent to Equation (5.39).

**Proposition 4** Suppose \( F_0 \) belongs to a space of distribution functions as described in (5.46). Moreover, suppose we have a random sample of size \( L \) of auctions each of which has \( N \) bidders and where we only observe the winning bid in every auction. Let the reservation price \( p_0 \) be known. Define

\[
\Lambda(\theta|a, c; k, p_0) = \int \left( 1 - \mathbf{1}\left\{ F_{b^*}(b; p_0) \geq F(\overline{S}_k(b; N|\theta, p_0); \theta, p_0)^N \right\} \right) \mathbf{1}\{a \leq b \leq c\} dF_{b^*}(b; p_0),
\]

(5.48)

\[
\Gamma(\theta|k, p_0) = \int \int \Lambda(\theta|a, c; k, p_0) dF_{b^*}(a; p_0) dF_{b^*}(c; p_0).
\]

39
Then, under the sole assumption that all bidders are level-$k$ rational, the identified set is:

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

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

$$
F_c(\theta, p_0) = \left\{ \theta \in \Theta : \exists \theta' \in \Theta(k_0, p_0) \text{ such that } \theta' \in \mathcal{F}(\theta, p_0) \right\},
$$

Then, the identified set is

$$
\Theta^*_0(p_0) = \left\{ \theta \in \Theta : \theta \in \Theta^*_0(p_0) \text{ w.p.1 (with respect to } p_0) \right\}.
$$

Note that the identification result in (5.50) requires that we explicitly assume that Assumption B1 holds for winning bids, see footnote 15. With a nonzero reserve price, the number of actual bidders in a given auction may differ from $N$. The characterization of the identified set in Proposition 4 can still be constructive in this case if we assume that $N$ is the same across all auctions in the population, and if the number of actual bidders is observed. For the $\ell$th auction, denote the latter by $I_\ell$. Therefore, $I_\ell = \sum_{i=1}^{N} \mathbb{1}\{v_i \geq p_0\}$ and

$$
E[I_\ell] = NE_{p_0}[F(p_0; \theta_0, p_0)] = \frac{E[I_\ell]}{E_{p_0}[F(p_0; \theta_0, p_0)]},
$$

where $E_{p_0}[\cdot]$ denotes the expectation taken with respect to the reserve price, which is assumed to be observed for any given auction. The above result is the basis for identifying $N$. It follows then that Proposition 4 is a constructive identification result.

### 5.3.2 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, one would be interested in having an identification result for both $\theta$ and $k_0$ simultaneously.

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

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

Then, if Assumption B1 is satisfied with $k_0 \geq 2$, the following results hold
(i) \( \Gamma(k) = 0 \) for all \( k \leq k_0 \). However, \( \Gamma(k) = 0 \) does not imply \( k \leq k_0 \).

(ii) \( \Gamma(k) > 0 \) implies \( k > k_0 \).

It follows from Proposition 5 that any \( k' \) such that \( \Gamma(k') > 0 \) 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 \( \{ k \in \mathbb{N} : \Gamma(k) = 0 \} \) includes all \( k \leq k_0 \), but it 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 are usually unchallenged in identification analysis, and hence econometricians focus on robustness of estimation results to those auxiliary assumptions (that are not implied by theory—such as functional forms and distributional assumptions). This paper studies the identifying role that some behavioral assumptions play. Mainly, we examined the identifying power of equilibrium in three simple games. We replaced equilibrium with a form of rationality (interim rationalizability) which includes equilibrium as a special case, and compared the identified features of the game under rationality and under equilibrium. The games we studied are stylized versions of empirical models considered and applied in the literature and hence insights provided here can be carried over to those empirical frameworks. However, we do not advocate dropping the equilibrium assumptions from empirical work. But rather, the paper simply examines the question of what is the identifying power of equilibrium in these simple setups. For example, it is not clear that one would want to drop equilibrium in a first price auction since 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 and a balancing act guided by the economics of the particular application at hand needs to be done. In addition, we do not advocate either, the use of rationalizability per se as the basis for strategic interaction. There are other frameworks that are in the literature, but, we do note that the form of rationalizability used here has received a lot of attention by game theorists (see for example Morris and Shin (2003), Dekel, Fudenberg, and Morris (2007) and the references
cited therein). Also, interim rationalizability allows us to incorporate the concept of higher order beliefs into the econometric analysis through what we defined here as rationality levels.

Some questions remain to be answered and we leave those for ongoing and future work. As far as the above results, the paper here is concerned with identification. A natural extension would be to study the statistical properties of estimators proposed above and apply those estimators in empirical examples. Another avenue of research is to extend some of the ideas above to dynamic setups. It is well known that inference in dynamic games is hard when one tries to account for the presence of multiple equilibria. The identification question is complicated due mostly to the complexity of the underlying economic model and say beliefs off the equilibrium, where no data is available. From a practical perspective, estimating dynamic games while allowing for different data points to be generated by a different equilibrium is a hard problem since it involves solving for multiple fixed points in a complicated nonlinear problem. So, relaxing equilibrium in this setting might lead to enormous computational advantages due to not having to solve for these fixed points. It might also 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.
References


7 Appendix

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 $S(X)$,

$$\Pr(\varepsilon_1 > X'_1\beta_1, \varepsilon_2 > X'_2\beta_2 | X) \leq \Pr(Y_1 = 0, Y_2 = 0 | X) \leq \Pr(\varepsilon_1 > X'_1\beta_1 + \alpha_1, \varepsilon_2 > X'_2\beta_2 + \alpha_2 | X)$$

$$\Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1, \varepsilon_2 > X'_2\beta_2 | X) \leq \Pr(Y_1 = 1, Y_2 = 0 | X) \leq \Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1, \varepsilon_2 \leq X'_2\beta_2 + \alpha_2 | X)$$

$$\Pr(\varepsilon_1 > X'_1\beta_1, \varepsilon_2 \leq X'_2\beta_2 + \alpha_2 | X) \leq \Pr(Y_1 = 0, Y_2 = 1 | X) \leq \Pr(\varepsilon_1 > X'_1\beta_1 + \alpha_1, \varepsilon_2 \leq X'_2\beta_2 | X)$$

$$\Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1, \varepsilon_2 \leq X'_2\beta_2 + \alpha_2 | X) \leq \Pr(Y_1 = 1, Y_2 = 1 | X) \leq \Pr(\varepsilon_1 \leq X'_1\beta_1, \varepsilon_2 \leq X'_2\beta_2 | X)$$

(7.53)

We denote the true parameter value by $\theta_0$. To prove part (a), take any $\tilde{\beta}_1 \neq \beta_{1o}$ such that $\beta_{l,1} \neq \beta_{l,1o}$. Given this and the support properties of $X_{l,1}$, for any scalar $d$ we can observe either of the following two events with positive probability: (i) $X'_1\beta_1 + d > X'_1\beta_{1o}$ or (ii) $X'_1\beta_1 < X'_1\beta_{1o} + \alpha_{1o}$. Take case (i) first, with $d = \alpha_1$ (arbitrary): if $\beta_{2o,2,2o} > 0$ we can make $\beta_2 X_2 \to +\infty$ and $\beta_{2o} X_2 \to +\infty$. By Assumption (A1), this yields $\Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1, \varepsilon_2 \leq X'_2\beta_2 + \alpha_2 | X) \to \Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1 | X)$, and $\Pr(\varepsilon_1 \leq X'_1\beta_{1o}, \varepsilon_2 \leq X'_2\beta_{2o} | X) \to \Pr(\varepsilon_1 \leq X'_1\beta_{1o} | X)$. Therefore, with positive probability as $X_2$ explodes, $\Pr(\varepsilon_1 \leq X'_1\beta_1 + \alpha_1, \varepsilon_2 \leq X'_2\beta_2 + \alpha_2 | X) > \Pr(\varepsilon_1 \leq X'_1\beta_{1o}, \varepsilon_2 \leq X'_2\beta_{2o} | X) > \Pr(Y_1 = 1, Y_2 = 1 | X)$, which violates (7.53). If $\beta_{2o,2,2o} < 0$, the result is easier to obtain by making $\beta_2 X_2 \to -\infty$ and $\beta_{2o} X_2 \to -\infty$. For case (ii), drive $\beta_2 X_2 \to -\infty$ and $\beta_{2o} X_2 \to -\infty$ if $\beta_2 > 0$, or $\beta_2 X_2 \to +\infty$ and $\beta_{2o} X_2 \to +\infty$ if $\beta_2 < 0$. In either case we eventually obtain $\Pr(\varepsilon_1 > X'_1\beta_1, \varepsilon_2 > X'_2\beta_2 | X) > \Pr(\varepsilon_1 > X'_1\beta_{1o} + \alpha_{1o}, \varepsilon_2 > X'_2\beta_{2o} + \alpha_{2o} | X) > \Pr(Y_1 = 0, Y_2 = 0 | X)$, which violates (7.53). This establishes identification of $\beta_{l,1}$, an analog proof shows that $\beta_{l,2}$ is identified, which proves part (a).

To establish part (b) focus on the worst-case scenario, and take $\tilde{\theta} \neq \theta_0$ where $\tilde{\beta}_{d,p} \neq \beta_{d,p}$, but $\tilde{\beta}_{l,p} = \beta_{l,p}$ for $p = 1, 2$ — the parameters of the unbounded-support shifters are fixed at their true values. Identification here must rely on the properties of $X_{d,p}$, the bounded-support shifters. The condition in the statement of the proposition ensures that (i) or (ii) (above) hold even if we fix $\tilde{\beta}_{l,p} = \beta_{l,p}$. To complete the proof of (b) we proceed as in the previous paragraph (note that we now have $\tilde{\beta}_{d,p}, \beta_{l,p} > 0$). The case $\tilde{\beta}_{d,p} \neq \beta_{d,p}$ and $\tilde{\beta}_{l,p} \neq \beta_{l,p}$ is straightforward along the same lines. Now, onto part (c). Consider $\tilde{\theta}$ that is equal to $\theta_0$ element-by-element except for $\tilde{\alpha}_1 \neq \alpha_{1o}$ and recall that the parameter space of interest has $\alpha_p \leq 0$. Clearly, none of the lower bounds in (7.53) evaluated at $\tilde{\theta}$ will ever be larger than the corresponding upper bounds evaluated at $\theta_0$, and none of the upper bounds evaluated at $\tilde{\theta}$ will ever be smaller than the corresponding lower bounds evaluated at $\theta_0$. Therefore, without further assumptions $\theta$ and $\theta_0$ are observationally equivalent and $\alpha_1$ is not identified. The only way we can proceed is by adding more structure on $\Pr(Y_1, Y_2 | X)$. We have $\Pr(\varepsilon_1 \leq X'_1\beta_{1o} + \alpha_{1o}) \leq \Pr(Y_1 = 1 | X) \leq \Pr(\varepsilon_1 \leq X'_1\beta_{1o})$, therefore $\Pr(\varepsilon_1 \leq X'_1\beta_{1o} + \tilde{\alpha}_1) > \Pr(Y_1 = 1 | X)$ only if $\tilde{\alpha}_1 > \alpha_{1o}$. Therefore, $\tilde{\theta}$ can violate (7.53) only if $\tilde{\alpha}_1 > \alpha_{1o}$. For any such $\tilde{\alpha}_1$, let $\Delta = \Pr(\varepsilon_1 \leq X'_1\beta_{1o} + \tilde{\alpha}_1) - \Pr(\varepsilon_1 \leq X'_1\beta_{1o} + \alpha_{1o}) > 0$. 

46
By the assumption in part (c), there exists a subset \( X_1 \in \mathbb{S}(X_1) \) such that \( \Pr(Y_1 = 1|X_1) < \Pr(\varepsilon_1 \leq X'_1|\beta_{1o} + \alpha_{1o} + \Delta = \Pr(\varepsilon_1 \leq X'_1|\beta_{1o} + \bar{\alpha}_1)) \). Make \( X'_2\beta_{2o} \to +\infty \) and the lower bound on the fourth inequality in (7.53) will be violated. This establishes part (c). Any \( \bar{\theta} \neq \theta_0 \) where \( \tilde{\alpha}_p \neq \alpha_{po} \) and either \( \tilde{\beta}_{e,p} \neq \beta_{e,po} \) or \( \tilde{\beta}_{d,p} \neq \beta_{d,po} \) can be shown not to be observationally equivalent to \( \theta_0 \) using the same arguments as in the previous paragraphs given the assumptions in parts (a) and (b). \( \square \)

**Proof of Theorem 2**

Suppose there exists a subset of realizations in \( \overline{X}_1 \subset X_1^* \) such that

\[
X'_1\beta_1 + \Delta_1 + \alpha_1 > X'_1\beta_{1o} + \Delta_{1o} + \alpha_{1o} \quad \forall X_1 \in \overline{X}_1. \tag{7.54}
\]

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

\[
H_1\left(X'_1\beta_1 + \Delta_1 + \alpha_1 p^L(X_1)\right) < H_1\left(X'_1\beta_{1o} + \Delta_{1o} + \alpha_{1o} p^L(X_1)\right). \tag{7.55}
\]

To see why \( p^L(X_1) \) and \( p^U(X_1) \) exist, fix \( p^U(X_1) = 1 \). By continuity, there exists a small enough \( \delta > 0 \) such that \( p^L(X_1) \geq 1 - \delta \) satisfies (7.55). If condition (4.17) in Theorem 2 holds, then there exists \( W_1^* \subset \mathbb{S}(W_1) \) such that

\[
\min\left\{ E\left[H_2(X'_2\beta_2 + \Delta_2 + \alpha_2)|I_1\right], E\left[H_2(X'_2\beta_{2o} + \Delta_{2o} + \alpha_{2o})|I_1\right]\right\} \geq p^L(X_1) \quad \forall W_1 \in W_1^*.
\]

\[
\max\left\{ E\left[H_2(X'_2\beta_2 + \Delta_2)|I_1\right], E\left[H_2(X'_2\beta_{2o} + \Delta_{2o})|I_1\right]\right\} \leq p^U(X_1) \quad \forall W_1 \in W_1^*. \tag{7.56}
\]

By definition, we have

\[
E\left[H_2(X'_2\beta_2 + \Delta_2 + \alpha_2)|I_1\right] = \pi^L_{2}\left(\theta|k = 2; I_1\right); \quad E\left[H_2(X'_2\beta_{2o} + \Delta_{2o})|I_1\right] = \pi^U_{2}\left(\theta_0|k = 2; I_1\right). \tag{7.57}
\]

Combining (7.56) and (7.57),

\[
\pi^L_{2}\left(\theta|k = 2; I_1\right) \geq p^L(X_1); \quad \pi^U_{2}\left(\theta_0|k = 2; I_1\right) \leq p^U(X_1) \quad \forall W_1 \in W_1^*. \tag{7.58}
\]

Combining (7.55) and (7.58),

\[
H_1\left(X'_1\beta_1 + \Delta_1 + \alpha_1 \pi^L_{2}\left(\theta|k = 2; I_1\right)\right) \leq H_1\left(X'_1\beta_1 + \Delta_1 + \alpha_1 p^L(X_1)\right) < H_1\left(X'_1\beta_{1o} + \Delta_{1o} + \alpha_{1o} p^L(X_1)\right) \leq H_1\left(X'_1\beta_{1o} + \Delta_{1o} + \alpha_{1o} \pi^U_{2}\left(\theta_0|k = 2; I_1\right)\right) \quad \forall W_1 \in W_1^*. \tag{7.59}
\]

This would correspond to the case described in the first line of Equation (4.18). Next, suppose (7.54) does not hold but there exists a subset of realizations \( \overline{X}_1^{**} \subset X_1^* \) such that

\[
X'_1\beta_{1o} + \Delta_{1o} + \alpha_{1o} > X'_1\beta_1 + \Delta_1 + \alpha_1 \quad \forall X_1 \in \overline{X}_1^{**}. \tag{7.60}
\]

47
Repeating the same arguments as above exchanging $\theta$ and $\theta_0$, we would arrive at the equivalent of (7.59), namely

$$H_1(\mathbf{X}_1' \beta_1 + \Delta_{10} + \alpha_1 \pi_{2}^{L}(\theta_0|k = 2; \mathcal{I}_1)) < H_1(\mathbf{X}_1' \beta_1 + \Delta_1 + \alpha_1 \pi_{2}^{L}(\theta|k = 2; \mathcal{I}_1)) \quad \forall \ W_1 \in \mathcal{W}_{1}^{**}.$$ \hspace{1cm} (7.61)

This would correspond to the case described in the second line of Equation (4.18). The last remaining possibility is that neither (4.18) nor (7.60) hold. In this case,

$$X_1' \beta_{10} + \Delta_{10} + \alpha_{10} = X_1' \beta_1 + \Delta_1 + \alpha_1 \quad \forall \ X_1 \in \mathcal{X}_1^*.$$ \hspace{1cm} (7.62)

Since $X_1$ has full column rank in $\mathcal{X}_1^*$, (7.62) implies $\beta_{10} = \beta_1$ and $\Delta_{10} + \alpha_{10} = \Delta_1 + \alpha_1$. Since $\theta_1 \neq \theta_{10}$, we must have either

$$\Delta_1 > \Delta_{10}, \quad \text{or} \quad \Delta_1 < \Delta_{10}.$$ \hspace{1cm} (7.63)

Suppose $\Delta_1 > \Delta_{10}$. This immediately yields $X_1' \beta_1 + \Delta_1 > X_1' \beta_{10} + \Delta_{10}$ for all $X_1 \in \mathcal{X}_1^*$. By continuity, we can find a pair $0 \leq p^L(X_1) < p^U(X_1) \leq 1$ such that

$$H_1(X_1' \beta_1 + \Delta_1 + \alpha_1 p^L(X_1)) > H_1(X_1' \beta_{10} + \Delta_{10} + \alpha_{10} p^L(X_1)).$$ \hspace{1cm} (7.64)

To see why $p^L(X_1)$ and $p^U(X_1)$ exist, fix $p^L(X_1) = 0$. By continuity, there exists a small enough $\delta > 0$ such that $p^U(X_1) \leq \delta$ satisfies (7.64). If condition (4.17) in Theorem 2 holds, then there exists $\mathcal{W}_1^{***} \subset S(W_1)$ such that

$$\begin{align*}
\text{Min} \left\{ E[H_2(X_2' \beta_2 + \Delta_2 + \alpha_2)|\mathcal{I}_1], \ E[H_2(X_2' \beta_2) + \Delta_2 + \alpha_2)|\mathcal{I}_1] \right\} &\geq p^L(X_1) \quad \forall \ W_1 \in \mathcal{W}_1^{***} \\
\text{Max} \left\{ E[H_2(X_2' \beta_2 + \Delta_2)|\mathcal{I}_1], \ E[H_2(X_2' \beta_2) + \Delta_2)|\mathcal{I}_1] \right\} &\leq p^U(X_1) \quad \forall \ W_1 \in \mathcal{W}_1^{***}.
\end{align*}$$ \hspace{1cm} (7.65)

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, Equation 7.57), we obtain

$$\pi_2^L(\theta|k = 2; \mathcal{I}_1) \leq p^U(X_1); \quad \pi_2^L(\theta_0|k = 2; \mathcal{I}_1) \geq p^L(X_1) \quad \forall \ W_1 \in \mathcal{W}_1^{***}.$$ \hspace{1cm} (7.66)

Using (7.64), this yields

$$H_1(X_1' \beta_1 + \Delta_1 + \alpha_1 \pi_2^U(\theta|k = 2; \mathcal{I}_1)) > H_1(X_1' \beta_{10} + \Delta_{10} + \alpha_{10} \pi_2^L(\theta_0|k = 2; \mathcal{I}_1)) \quad \forall \ W_1 \in \mathcal{W}_1^{***}.$$ \hspace{1cm} (7.67)

This corresponds to a case like the one described in the second line of Equation (4.18). If $\Delta_1 \leq \Delta_{10}$, the same arguments as above while exchanging $\theta$ with $\theta_0$ would lead us to conclude that there exists a set $\mathcal{W}_1^{4*} \subset S(W_1)$ such that

$$H_1(X_1' \beta_{10} + \Delta_{10} + \alpha_{10} \pi_2^L(\theta_0|k = 2; \mathcal{I}_1)) > H_1(X_1' \beta_1 + \Delta_1 + \alpha_1 \pi_2^U(\theta|k = 2; \mathcal{I}_1)) \quad \forall \ W_1 \in \mathcal{W}_1^{4*}.$$
We have now established Equation (4.18) in Theorem 2 for the case \( k = 2 \). The cases \( k > 2 \) follow immediately from here by recalling the monotonic property of rationalizable bounds which says that, with probability one,

\[
H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_L^k (\theta|k; \mathcal{I}_1)) \leq H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_U^k (\theta|k; \mathcal{I}_1)) \quad \forall k \geq 1
\]

\[
H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_U^k (\theta|k; \mathcal{I}_1)) \geq H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_L^k (\theta|k; \mathcal{I}_1)) \quad \forall k \geq 1.
\]

To see why this implies that the rationalizable bounds for Player 1’s conditional choice probabilities are disjoint with positive probability for all \( k \geq 2 \), recall that the level-2 bounds are given by

\[
\left[ H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_L^2 (\theta|k = 2; \mathcal{I}_1)), H_1(X'_1 \beta_1 + \Delta_1 + \alpha_1 \pi_U^2 (\theta|k = 2; \mathcal{I}_1)) \right] \quad \text{(for } \theta \text{)}
\]

\[
\left[ H_1(X'_1 \beta_{10} + \Delta_{10} + \alpha_{10} \pi_L^2 (\theta_0|k = 2; \mathcal{I}_1)), H_1(X'_1 \beta_{10} + \Delta_{10} + \alpha_{10} \pi_U^2 (\theta_0|k = 2; \mathcal{I}_1)) \right] \quad \text{(for } \theta_0 \text{)}. \tag{7.69}
\]

It follows from our results that the level-2 rationalizable bounds for \( \theta \) are disjoint from those of \( \theta_0 \) with positive probability. Since the bounds for \( k > 2 \) are contained in those of \( k = 2 \) w.p.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, any \( \theta \) with \( \theta_1 \neq \theta_{10} \) will produce level-2 bounds that are violated with positive probability. Thus, no such \( \theta \) can be observationally equivalent to one that has \( \theta_1 = \theta_{10} \) and consequently, \( \theta_{10} \) is identified. Naturally, if the same conditions of Theorem 2 hold when we exchange the subscripts “1” and “2”, then \( \theta_{20} \) will be identified. \( \square \)
Notes

1This is important since there is theoretical work that challenges the multiplicity issues that arise under rationalizability. For example, Weinstein and Yildiz (2007) show that for any rationalizable set of strategies in a given game, there is a local disturbance of that game where these are the unique rationalizable strategies. This ambiguity about what is the exact game that is being played is exactly the reason why it is important to study the identified features of a model in the presence of multiplicity.

2The model of demand and supply uses equilibrium to equate the quantity demanded with quantity supplied thus obtaining the classic simultaneous equation model. Other literatures in econometrics, like job search models and hedonic equilibrium models explicitly use equilibrium as a “moment condition.”

3There might be ways to obtain sharper inference in these classes of games. This was pointed out to us by Francesca Molinari.

4The probabilities can be given a “structural” interpretation where 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 (4.2) for some well-defined beliefs. Expected utility maximization explains (through Equation 4.2) how players’ choices are produced in an incomplete information environment given beliefs.

5Throughout, we will focus on the case where \( \mu_{-p} \) is continuously distributed conditional on \( I_p \), and will ignore the distinction between strictly and weakly dominated strategies.

6The results that follow only require, for each player \( p \), that the support of \( \varepsilon_p \) be larger than that of \( X_p^\beta_p \) for all possible realizations of \( I_{-p} \).

7Recall that we are studying the case \( \alpha_p \leq 0 \) for \( p = 1, 2 \).

8Here, we refer to the identified set as the set of values of \( \alpha_p \) that are observationally equivalent, conditional on observables, to the true value \( \alpha_{p0} \).

9Strictly speaking, this would be a joint test of the rationality hypothesis and all other maintained assumptions.

10More precisely, the distance (in the unit square) between the smallest and the largest equilibrium beliefs.

11Interim rationalizability will only naturally produce upper bounds for rationalizable bids. Additional, ad-hoc assumptions could be made to characterize a lower bound.

12The results we analyze here do not depend on the specific choice of the sigma algebra, as long as it satisfies the singleton-measurability mentioned here. See footnote 10 in BS.

13Strictly speaking, what matters is that ties have probability zero for the most pessimistic conjecture.

14These and more properties are enumerated in BS, who focus on a more general case which allows for interdependent values.

15It appears that even if B1 is assumed to hold for all bids, 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.

Note trivially that since $\alpha_p \leq 0$ everywhere in $\Theta$, we have

\[
\min \left\{ E[H_2(X'_{20}\beta_2 + \Delta_2 + \alpha_2)|I_1], E[H_2(X'_{20}\beta_2 + \Delta_2 + \alpha_2)|I_1] \right\} \\
\leq \max \left\{ E[H_2(X'_{20}\beta_2 + \Delta_2)|I_1], E[H_2(X'_{20}\beta_2 + \Delta_2)|I_1] \right\} \quad \text{w.p.1.}
\]