Abstract

This paper reviews the econometrics of static games, with a focus on discrete-choice cases. These models have been used to study a rich variety of empirical problems, ranging from labor force participation to entry decisions. We outline the components of a general game and we describe the problem of doing robust inference in the presence of multiple solutions, and the different econometric approaches that have been applied to tackle this problem. We then describe the specific challenges that arise in different variations of these models depending on whether players are assumed to have complete or incomplete information, as well as whether or not non-equilibrium play is allowed. We describe the results in 2×2 games (the most widely studied game in econometrics) and we present extensions and recent results in games with richer action spaces. Areas for future research are also discussed.

JEL Codes: C57, C13, C14, C18, C35.

1 Introduction

This paper provides an overview of the general problem and the state of the literature on econometric inference in static games, with an emphasis on discrete-choice games. These models have been applied to study problems including labor force participation, entry, technology adoption, product differentiation, advertising and analyst stock recommendations, among others. We begin by outlining the general inferential problem and the challenges faced by the econometrician; in particular, the issue of robust inference when the game has multiple solutions. We also compare the issues and challenges of complete-information and incomplete-information games, as well as equilibrium and nonequilibrium models. The paper then illustrates inference in all these cases in a 2×2 game, the canonical application in econometric models of games. We then proceed to illustrate recent advances in discrete games with richer action spaces. The results we discuss include parametric and nonparametric models, complete and incomplete-information games, as well as...
equilibrium and non-equilibrium behavior. Following our review and analysis, we conclude by suggesting areas of future research.

2 A general normal-form game

This paper will focus on econometric inference of a normal form (or strategic form) static game where players move simultaneously (i.e., before observing the realized choices of others). The generic game consists of the following.

Players and actions: The game consists of a set of \{1, 2, \ldots, P\} players. Each player \(p\) has to select an action \(Y_p\) from within an action space \(S_p\). The subscript \(-p\) will denote “\(p\)’s opponents”. We denote a generic action for \(p\) as \(y_p \in S_p\) and \(y_{-p} \equiv (y_q)_{q \neq p} \in S_{-p}\) denotes a generic action profile for \(p\)’s opponents. \(Y_p\) will denote the actual action (the choice) made by \(p\) and \(Y_{-p} \equiv (Y_q)_{q \neq p}\) will denote the choice profile of \(p\)’s opponents. We will refer to \(Y \equiv (Y_p)_{p=1}^P\) as the outcome of the game.

Payoff functions: The (von Neumann–Morgenstern) payoff function of player \(p\) is denoted as
\[
u_p(Y_p, Y_{-p}), \tag{1}\]
Payoffs are treated as random functions, with \(\nu_p \in \mathcal{U}_p\) (a space of functions). We assume that player \(p\) observes the realization of her payoff function \(\nu_p\) prior to the game being played. Let \(\mathcal{U}\) denote the joint space of payoff functions for all the players in the game.

We will consider static games where players choose their strategies simultaneously and we will focus on cases where actions are scalar and the action space is discrete. The game has the following additional components.

Beliefs: Players are expected-payoff maximizers. They construct their expected payoffs by forming beliefs about the distribution of actions chosen by their opponents. Beliefs for player \(p\) are denoted as \(\pi_{-p}\), a distribution function over \(S_{-p}\). For a given set of beliefs \(\pi_{-p}\), the expected payoff function for \(p\) is given by
\[
u_p(Y_p, \pi_{-p}) = \int_{y_{-p} \in S_{-p}} \nu_p(Y_p, y_{-p}) d\pi_{-p}(y_{-p}). \tag{2}\]
We denote the space of possible beliefs for player \(p\) as \(\mathcal{B}_{-p}\), with \(\mathcal{B}\) denoting the combined space of beliefs for all players in the game. Degenerate beliefs that assign probability-mass one to a particular action profile are always allowed.
Strategies: A strategy $\sigma_p$ for player $p$ is a distribution over $S_p$. Let $\Sigma_p$ denote the strategy space for player $p$ and let $\Sigma$ denote the combined strategy space for all players in the game. For a given set of beliefs $\pi_{-p}$, the expected payoff for $p$ from playing strategy $\sigma_p$ is

$$\pi_p(\sigma_p, \pi_{-p}) = \int_{y_p \in S_p} \pi_p(y_p, \pi_{-p}) d\sigma_p(y_p) \quad (3)$$

Pure strategies assign probability-mass one to a particular action. We refer to all others as mixed-strategies.

Best-responses: The strategy $\sigma_p$ is a best-response for a particular set of beliefs $\pi_{-p}$ if $\pi_p(\sigma_p, \pi_{-p}) \geq \pi_p(\sigma_p', \pi_{-p})$ for all $\sigma_p' \in S_p$. Given a payoff function $u_p$, we denote this as $\sigma_p = BR_p(\pi_{-p} | u_p)$.

Solution: Let $\Omega_p$ denote the joint space of beliefs and strategies $(\pi_{-p}, \sigma_p)$ for player $p$. Given a payoff function $u_p$, let

$$\Omega_p^*(u_p) = \{(\pi_{-p}, \sigma_p) \in \Omega_p: \sigma_p = BR_p(\pi_{-p} | u_p)\}.$$ 

For a given strategy profile $y = (y_p)_{p=1}^P$, define

$$\Omega^*(y | u) = \{(\pi_{-p}, \sigma_p)_{p=1}^P \in \Omega^*(u): \sigma_p(y_p) > 0 \ \forall \ p\}.$$ 

$\Omega^*(y | u)$ is the set of all existing solutions where $y$ can be chosen with nonzero probability. Given $u$, we have $Y = y$ only if $\Omega^*(y | u) \neq \emptyset$. If the game has multiple existing solutions, the model can be completed by adding a solution selection mechanism.

Solution selection: Given $u$, there exists a mechanism $M$ that selects a solution from within $\Omega^*(u)$. Denote the solution selected by $M$ as

$$\{(\pi_{-p}^*, \sigma_p^* | u)\}_{p=1}^P$$

Then,

$$Pr(Y_p \in A | u) = \int_{y_p \in A} d\sigma_p^*(y_p | u) \quad \forall \ A \subseteq S_p \quad (5)$$

All models we will review here can be expressed as variations of the general normal-form game described above, with different assumptions about:
• **Solution concept:** The majority of econometric studies assume equilibrium behavior, which presupposes **correct beliefs**, and the most widely assumed equilibrium concept is Nash equilibrium (NE). Nonequilibrium solution concepts have also been considered. For example, motivated by empirical evidence of deviations from equilibrium behavior (see, e.g., Stahl and Wilson (1994)), experimental economists have used models of cognitive hierarchy (among others) to model behavior. These models are characterized by very precise assumptions about how exactly agents deviate from equilibrium. Examples include Costa-Gomes and Crawford (2001), Camerer (2004) and Costa-Gomes and Crawford (2006).

Econometric methods that do not impose equilibrium behavior include Aradillas-Lépez and Tamer (2008), Klein and Tamer (2012), Kline (2015), Kline (2018), and Aradillas-Lépez (2019). Unlike the experimental methods which assume a precise form in which players deviate from equilibrium, these papers rely on general restrictions on behavior and beliefs that include NE as a special case and, in some instances (e.g., Aradillas-Lépez and Tamer (2008) and Kline (2015)), also nest experimental models such as cognitive hierarchy.

• **Information:** Complete information models assume that the realization of payoff functions \( \{u_p\} \) is common knowledge, while incomplete information settings assume that the exact realization of payoff functions is only privately observed by each player. Some examples of complete-information games include Bjorn and Vuong (1984), Bresnahan and Reiss (1990, 1991a, 1991b), Berry (1992), Tamer (2003), Bajari, Hong, and Ryan (2010), Ciliberto and Tamer (2009), Aradillas-Lépez (2011), Kline (2015), Kline (2016), Aradillas-Lépez (2019) and Aradillas-Lépez and Rosen (2019). Incomplete-information games have been studied, for example, in Bajari, Hong, Krainer, and Nekipelov (2010), Aradillas-Lépez (2010), de Paula and Tang (2012), Aradillas-Lépez (2012), Xu and Wan (2014), Xu (2014), Lewbel and Tang (2015), Aradillas-Lépez and Gandhi (2016), Liu, Xu, and Vuong (2017) and Xiao (2018). The presence of multiple equilibria tends to be more pervasive in games with complete information (see Morris and Shin (2003)). Incomplete-information games require assumptions about the information possessed by players to construct their beliefs.

• **Beliefs:** Equilibrium models presuppose correct beliefs. In equilibrium games of complete information, beliefs are completely characterized by the realization of payoff functions. Incomplete-information games require specific assumptions about the information observed by players. A crucial assumption is whether players’ private-information is assumed to be independent across players conditional on observables to the econometrician. Examples that assume conditional independence include Bajari, Hong, Krainer, and Nekipelov (2010), de Paula and Tang (2012), Aradillas-Lépez (2012), Lewbel and Tang (2015), Aradillas-Lépez (2014) and Grieco (2014) analyzes a parametric game where unobservables have a particular structure that can combine (nest) the complete and incomplete information frameworks depending on the value of a subset of parameters.
and Gandhi (2016) and Xiao (2018). Papers that allow correlation in players’ private information under additional restrictions include Aradillas-López (2012), Grieco (2014), Xu and Wan (2014), Xu (2014) and Liu, Xu, and Vuong (2017). Methods that allow for incorrect beliefs typically restrict the space of beliefs according to weaker restrictions, such as iterated dominance or rationalizability.

- **Action spaces**: We will focus mainly on discrete action spaces. Within these models, binary-choice games have received the majority of attention (see, e.g., Berry and Tamer (2007) and the literature cited there). Discrete games with a richer-than-binary action space typically have a more complicated set of solutions and therefore are more challenging to analyze. Complete-information games with more than two actions have been analyzed in Bajari, Hong, and Ryan (2010) (multinomial), Davis (2006), Aradillas-López (2011) and Aradillas-Lopez and Rosen (2019) (ordinal). Nonparametric results in ordinal incomplete-information games have been obtained in Aradillas-López and Gandhi (2016).

- **Strategy spaces**: Here, the main distinction is whether mixed-strategies are allowed or attention is restricted to pure-strategies. Identification results in binary-choice games with complete information have been obtained by ruling out mixed-strategies. Under symmetry conditions, this leads to a unique prediction for \( \sum_{p=1}^{P} Y_p \) in all co-existing equilibria. This result has been exploited in Bresnahan and Reiss (1990, 1991a, 1991b), Berry (1992) and Tamer (2003), and it is discussed in detail in Berry and Tamer (2007, Section 2.4). Incomplete-information econometric methods typically assume beliefs that produce to a unique optimal action w.p.1 and therefore lead to pure-strategies.

- **Multiple solutions and selection mechanism**: Multiple equilibria are more prevalent in complete-information games (see Morris and Shin (2003)). As a result, an important body of work has been devoted to doing robust inference in complete-information games without any explicit assumptions about the underlying equilibrium selection mechanism. The list of such studies is extensive, but most are purely econometric papers that include a game as an example. Perhaps the most significant effort to date explicitly devoted to a nontrivial game is Ciliberto and Tamer (2009). On the other extreme we would have papers with an explicit model of equilibrium selection, and one of the most notable complete-information examples is Bajari, Hong, and Ryan (2010).

In econometric studies of incomplete-information games with equilibrium behavior, a commonly made assumption is that the underlying selection mechanism is degenerate, so that it chooses a unique equilibrium w.p.1. or, in other words “the data comes from a single equilibrium”. This assumption can often help identify the remaining parameters of the model.

---

\footnote{Ciliberto and Tamer (2009) allow for asymmetries and ruling out mixed-strategies does not lead to a unique prediction for \( \sum_{p=1}^{P} Y_p \) in all co-existing equilibria.}
(a result that does not hold in complete-information games). This type of assumption has led to identification results, for example, in Seim (2006), Pesendorfer and Schmidt-Dengler (2008), Bajari, Hong, Krainer, and Nekipelov (2010), Aradillas-López (2012). Studies that do not assume a degenerate selection mechanism include de Paula and Tang (2012), Aradillas-López and Gandhi (2016) and Xiao (2018). Examples that include an explicit model of equilibrium selection include Ackerberg and Gowrisankaran (2006) and Sweeting (2009).

In many instances, the presence of multiple equilibria has identification power, for example, to infer the direction of strategic-interaction (de Paula and Tang (2012), Aradillas-López and Gandhi (2016)) or the parameters of the model (Ackerberg and Gowrisankaran (2006), Sweeting (2009)). Testing for the presence of multiple equilibria has been studied, for example, in Otsu, Pesendorfer, and Takahashi (2016), Hahn, Moon, and Snider (2017) and Marcoux (2018). A survey of econometric methods with multiple equilibria in games can be found in de Paula (2013). Nonequilibrium inference typically makes no assumptions about the selection mechanism but focuses on doing robust inference involving the rest of the parameters (payoffs, beliefs).

- **Payoff functions**: Methods can be classified according to the parametric assumptions made about payoff functions. The list of studies that parametrize payoffs is vast and includes the majority of existing work. Nonparametric payoff functions are considered, for example, in Aradillas-López (2011), de Paula and Tang (2012), Kline and Tamer (2012), Lewbel and Tang (2015), Aradillas-López and Gandhi (2016), Liu, Xu, and Vuong (2017), Aradillas-López (2019), among others. In many nonparametric studies, the goal is to test or infer qualitative features of the game, such as the sign of strategic interaction, the information possessed by players or the presence of multiple equilibria.

3 The inferential problem

3.1 Data observed

The setting we consider is one where the econometrician observes \( i = 1, \ldots, n \) realizations of the game described above. The researcher observes \( Y_i = (Y_{p,i})_{i=1}^n \) and \( X_i \), where \( X_i \) is a collection of covariates that contain information for \( (u_{p,i}, \pi_{-p,i}, \sigma_{p,i})_{p=1}^P \). The models we will review assume that \( (Y_i, X_i)_{i=1}^n \) is a random sample, with \( (Y, X) \sim P_{X,Y} \) and \( P_{X,Y} \) nonparametrically identified.

3.2 Inference when the game produces unique predictions

If the game has a unique solution almost surely, then the selection mechanism becomes redundant and inference can be based on [5] with a trivial selection mechanism that chooses the only
existing solution. Alternatively, the game can have multiple solutions but there may exist an aggregate function $T(Y)$ such that all co-existing solutions yield a unique prediction for $T(Y)$. For example, under some symmetry conditions this happens in binary-choice games of complete information and NE behavior for $T(Y) = \sum_{p=1}^{P} Y_p$ if mixed strategies are ruled out (see Berry and Tamer (2007, Section 2.4)). In these cases, inference can proceed by applying (5) to $T(Y)$ and the specific selection mechanism becomes irrelevant.

### 3.3 Inference with an explicit selection mechanism

If the model is equipped with an explicit selection mechanism (or if there is a unique solution almost surely), inference would be based on (5).

### 3.4 Inference without a selection mechanism and multiple solutions

Let $\Psi_p \equiv (u_p, \pi_{-p}, \sigma_p)$ and $\Psi = \{\Psi_p\}_{p=1}^{P} \in \Upsilon$. Without an explicit model for the selection mechanism, the goal is to do robust inference on $\Psi$ based solely on the implications of the solution concept employed. Take any collection of action profiles $A \subseteq S$ and let

$$R(A) = \{\Psi: \Omega^*(y|u) \neq \emptyset \text{ for some } y \in A\},$$

where $\Omega^*(y|u)$ is as defined in (4). Note that $Y \in A$ only if $\Psi \in R(A)$.

**A notational convention:** For a random variable $\xi$ we will use the statements “a.e $\xi$”, “a.s in $\xi$” or “w.p.1. in $\xi$” interchangeably to refer to an event that is satisfied “for almost every realization of $\xi$”, “almost surely in $\xi$” or “with probability one in $\xi$”.

#### 3.4.1 Parameters and the identified set

The identified set is the collection of all $\Psi \in \Upsilon$ that could have generated the data observed for some selection mechanism. Suppose $(Y, \Psi)|X \sim P_X \in \mathcal{P}$ (a space of distributions). Any particular $P_X$ reflects implicit properties about the underlying selection mechanism, but the latter is not explicitly modeled. The space $\mathcal{P}$ implicitly restricts the class of selection mechanisms allowed. Since $P_{Y,X}$ is nonparametrically identified, the joint distribution of $(Y, X, \Psi)$ is completed by $P_X$ and we can describe the unknown parameters of the model as $(\Psi, P_X) \in \Theta$, where $\Theta = \{ (\Psi, P_X): \Psi \in \Upsilon, P_X \in \mathcal{P}\}$ denotes the parameter space. Using the above definition for $R$,

$$\Pr(Y \in A|X) \leq P_X(\Psi \in R(A)|X) \quad \forall A \subseteq S, \text{ a.e } X.$$  

---

Footer note:

3One of the first papers to formalize this notion was Jovanovic (1989) in equilibrium models.
The identified set of parameters can be defined from the above condition. However, it can be characterized in an alternative way. Note that, for any set \( C \subseteq \Upsilon \), \( R(Y) \subseteq C \) implies \( \Psi \in C \) (since \( \Psi \in R(Y) \) a.s). The identified set can be defined as

\[
\Theta_I = \{ (\Psi, \mathcal{P}_X) \in \Theta : \mathcal{P}_X (R(Y) \subseteq C | X) \leq \mathcal{P}_X (\Psi \in C | X) \ \forall C \subseteq \Upsilon, \ \text{a.e. X.} \}
\]

This definition is sharp, meaning that a parameter is in \( \Theta_I \) if and only if there exists an underlying selection mechanism (not necessarily unique) that makes it consistent with the data observed. The characterization in (8) may be infeasible in practice because it involves all possible subsets in \( \Upsilon \). If we pre-specify a class of subsets \( \bar{C} \subseteq \Theta \) and define

\[
\Theta_{I_{\bar{C}}} = \{ (\Psi, \mathcal{P}_X) \in \Theta : \mathcal{P}_X (R(Y) \subseteq \bar{C} | X) \leq \mathcal{P}_X (\Psi \in \bar{C} | X) \ \forall \bar{C} \subseteq \bar{C}, \ \text{a.e. X} \}. \]

Note that \( \Theta_I \subseteq \Theta_{I_{\bar{C}}} \). A definition such as \( \Theta_{I_{\bar{C}}} \) can therefore be conservative but computationally feasible to implement. An important body of econometric work has been devoted to the issue of feasible sharp inference, aimed at characterizing the smallest class of sets \( \bar{C} \) such that \( \Theta_I = \Theta_{I_{\bar{C}}} \). Such a class is often referred to as the class of core determining sets (Galichon and Henry (2011)).

Sharp inference methods have used results from random set theory (see Beresteanu and Molinari (2008), Beresteanu, Molchanov, and Molinari (2011), Chesher and Rosen (2017)) and optimal transportation theory (Galichon and Henry (2011)).

### 3.4.2 Moment inequalities characterizations of the identified set

Inference can be performed bypassing entirely the selection mechanism. Instead of \( \mathcal{P}_X \) (the joint distribution of \((Y, \Psi)\mid X\)), we can focus merely on \( \mathcal{Q}_X \), the distribution of \( \Psi \mid X \), and define

\[
\tilde{\Theta}_I = \{ (\Psi, \mathcal{Q}_X) \in \Theta : \text{Pr}(Y \in A \mid X) \leq \mathcal{Q}_X (\Psi \in R(A) \mid X) \ \forall A \in S, \ \text{a.e. X} \}. \]

Clearly, \( \Theta_I \subseteq \tilde{\Theta}_I \), but \( \tilde{\Theta}_I \) may include \( \Psi \)'s that are excluded from \( \Theta_I \). Inferential methods based on moment inequalities can focus only on a class of subsets in \( S \). For example, in discrete games, we could focus only on the class of singletons in \( S \),

\[
\tilde{\Theta}_{I_S} = \{ (\Psi, \mathcal{Q}_X) \in \Theta : \text{Pr}(Y = y \mid X) \leq \mathcal{Q}_X (\Psi \in R(y) \mid X) \ \forall y \in S, \ \text{a.e. X} \}. \]

Econometric inference with moment inequalities has been an area of active research, and game-theoretic models have been among the most important applications. Examples include Cher- nozhukov, Hong, and Tamer (2007), Romano and Shaikh (2010), Andrews and Soares (2010), Bugni (2010), Andrews and Jia-Barwick (2010), Chernozhukov, Lee, and Rosen (2013), Andrews and Shi (2013), Romano, Shaikh, and Wolf (2014), Chetverikov (2012), Andrews and Shi (2013).
3.4.3 Using a superset for the set of solutions $\mathcal{R}$

In some cases, a full characterization of the set $\mathcal{R}$ (defined in (6)) may require finding all existing solutions to the game or it may necessitate stronger assumptions about payoff functions (e.g., parametrization) than the researcher is willing to make. In such instances, it may be possible to characterize a superset $\overline{\mathcal{R}}$ that is feasible to compute under the specific assumptions of the model, with the property that $\mathcal{R}(A) \subseteq \overline{\mathcal{R}}(A)$ for all $A$. From here, any of the previous characterizations of the identified set can be constructed replacing $\mathcal{R}$ with $\overline{\mathcal{R}}$. This approach is taken, for example, in Aradillas-López (2011), Kline and Tamer (2012), de Paula and Tang (2012) and Aradillas-López and Gandhi (2016).

4 $2 \times 2$ games

Binary choice games with $Y_p \in \{0, 1\}$ have been studied extensively, particularly to model simultaneous participation or “entry” decisions. Some examples include Bjorn and Vuong (1984), Brennan and Reiss (1990), Berry (1992), Tamer (2003), Sweeting (2009), Ciliberto and Tamer (2009), Aradillas-López (2010), Klein and Tamer (2012), de Paula and Tang (2012). Surveys of entry games can be found in Berry and Reiss (2007) and Berry and Tamer (2007).

The canonical example of binary choice games in econometrics is the $2 \times 2$ case. This was the focus of Bjorn and Vuong (1984), the paper that pioneered the econometric analysis of games and a contribution that was well ahead of its time. Consider the following matrix form game.

\[
\begin{array}{c|c|c}
Y_2 = 1 & Y_2 = 0 \\
Y_1 = 1 & X'_1 \beta_1 + \Delta_1 - \varepsilon_1, X'_2 \beta_2 + \Delta_2 - \varepsilon_2 & X'_1 \beta_1 - \varepsilon_1, 0 \\
Y_1 = 0 & 0, X'_2 \beta_2 - \varepsilon_2 & 0, 0 \\
\end{array}
\]

$\gamma \equiv (\beta_1, \beta_2, \Delta_1, \Delta_2)$ are unknown parameters. $X \equiv (X_1, X_2)$ and $\varepsilon \equiv (\varepsilon_1, \varepsilon_2)$ are non-strategic payoff shifters. $X$ is observed by the econometrician but $\varepsilon$ is not. The space of payoff functions in this parametric model is indexed by $\gamma \in \Gamma$ (the parameter space).

Assume for simplicity that $\Delta_p \leq 0$, corresponding strategic substitutes. We will also maintain that $\varepsilon \mid X$ is jointly continuously distributed with unbounded support $\mathbb{R}^2$. We will describe inference of this game under the cases of complete and incomplete information, and in each instance we will consider two alternative solution concepts: Nash equilibrium (NE) and iterated-dominance.

4.1 Inference in the complete information case

Suppose the true value of $\gamma$ and the realizations of $X$ and $\varepsilon$ are observed by both players.
4.1.1 Nash equilibrium (NE) behavior

A strategy profile \((\sigma_1, \sigma_2)\) is a NE in this game if \(u_p(\sigma_p, \sigma_{-p}) \geq u_p(y_p, \sigma_{-p})\) for \(y_p \in \{0,1\}\).

**Pure-strategy NE**

For a given \((X, \gamma)\) and any profile \(y = (y_1, y_2)\) let \(\mathcal{R}_{PSNE}(y|X, \gamma)\) denote the region of values of \((\epsilon_1, \epsilon_2)\) such that \(y\) is a pure-strategy NE (PSNE). The regions are as follows.

\[
\mathcal{R}_{PSNE}(1,1|X, \gamma) = \{(\epsilon_1, \epsilon_2): X'_1 \beta_1 + \Delta_1 - \epsilon_1 \geq 0, X'_2 \beta_2 + \Delta_2 - \epsilon_2 \geq 0\},
\]

\[
\mathcal{R}_{PSNE}(1,0|X, \gamma) = \{(\epsilon_1, \epsilon_2): X'_1 \beta_1 - \epsilon_1 \geq 0, X'_2 \beta_2 + \Delta_2 - \epsilon_2 \leq 0\},
\]

\[
\mathcal{R}_{PSNE}(0,1|X, \gamma) = \{(\epsilon_1, \epsilon_2): X'_1 \beta_1 + \Delta_1 - \epsilon_1 \leq 0, X'_2 \beta_2 - \epsilon_2 \geq 0\},
\]

\[
\mathcal{R}_{PSNE}(0,0|X, \gamma) = \{(\epsilon_1, \epsilon_2): X'_1 \beta_1 - \epsilon_1 \leq 0, X'_2 \beta_2 - \epsilon_2 \leq 0\},
\]

**Mixed-strategy NE**

A strategy profile \((\sigma_1, \sigma_2)\) is a mixed-strategy NE if and only if \(u_1(1, \sigma_2) = u_1(0, \sigma_2)\) and \(u_2(1, \sigma_1) = u_2(0, \sigma_1)\). That is,

\[
X'_1 \beta_1 - \epsilon_1 + \sigma_2 \cdot \Delta_1 = 0, \quad \text{and} \quad X'_2 \beta_2 - \epsilon_2 + \sigma_1 \cdot \Delta_2 = 0.
\]

This yields

\[
\sigma_1^M (1) = \frac{X'_2 \beta_2 - \epsilon_2}{-\Delta_2} \quad \text{and} \quad \sigma_2^M (1) = \frac{X'_1 \beta_1 - \epsilon_1}{-\Delta_1}.
\]

(9)

We have \(0 < \sigma_1(1) < 1\) and \(0 < \sigma_2(1) < 1\) if and only if \(X'_1 \beta_1 + \Delta_1 < \epsilon_1 < X'_1 \beta_1\) and \(X'_2 \beta_2 + \Delta_2 < \epsilon_2 < X'_2 \beta_2\). Thus,

\[
\mathcal{R}_{MSNE}(\sigma_1^M, \sigma_2^M |X, \gamma) = \{(\epsilon_1, \epsilon_2): X'_1 \beta_1 + \Delta_1 < \epsilon_1 < X'_1 \beta_1, X'_2 \beta_2 + \Delta_2 < \epsilon_2 < X'_2 \beta_2\},
\]

is the region of mixed-strategy NE. Figure 1 shows the NE regions.

Define \(\mathcal{R}_{NE}(y|X, \gamma)\) as the region of values of \((\epsilon_1, \epsilon_2)\) such that there exists a NE where \(y\) can be chosen with nonzero probability. We have

\[
\mathcal{R}_{NE}(y|X, \gamma) = \mathcal{R}_{PSNE}(y|X, \gamma) \cup \mathcal{R}_{MSNE}(\sigma_1^M, \sigma_2^M |X, \gamma).
\]

And therefore

\[
Pr(Y = y|X) \leq Pr(\epsilon \in \mathcal{R}_{NE}(y|X, \gamma)|X) \text{ for all } y, \text{ a.e } X.
\]

If we assume \(\epsilon|X \sim G(\cdot|X, \rho)\) (a parametric family of distributions) and we let \(\theta \equiv (\gamma, \rho) \in \Theta\), with

\[
\int_{\epsilon \in \mathcal{R}_{NE}(y|X, \gamma)} dG(\epsilon|X, \rho) = H_{NE}(y|X, \theta),
\]
we can perform inference on $\theta$ based on the moment inequalities implied by the model. Define

$$\overline{\Theta}_I = \{\theta \in \Theta : Pr(Y = y|X) \leq H_{NE}(y|X, \theta) \forall y, \text{ a.e } X\}.$$ 

A sharp characterization of the identified set can proceed in the manner described in Section 3.4.1. Suppose $(Y, \varepsilon)|X \sim \mathcal{P}_X \in \mathcal{P}$ (a class of distributions). The parameters of the game are $(\gamma, \mathcal{P}_X) \in \Theta$ and the identified set can be described as

$$\Theta_I = \{(y, \mathcal{P}_X) \in \Theta : \mathcal{P}_X(\mathcal{R}_{NE}(Y|X, \gamma) \subseteq \mathcal{C}|X) \leq \mathcal{P}_X(\varepsilon \in \mathcal{C}|X) \forall \mathcal{C} \subset \mathbb{R}^2, \text{ a.e } X\}.$$ 

### 4.1.2 Ruling out mixed-strategies

As Figure 1 shows, if we rule out mixed-strategies, every co-existing NE predicts the same value for $Y_1 + Y_2$. We have,

$$Pr(Y_1 + Y_2 = 0|X) = Pr(\varepsilon \in \mathcal{R}_{PSNE}((0, 0)|X, \gamma)|X),$$
$$Pr(Y_1 + Y_2 = 1|X) = Pr(\varepsilon \in \mathcal{R}_{PSNE}((1, 0)|X, \gamma) \cup \mathcal{R}_{PSNE}((0, 1)|X, \gamma)|X),$$
$$Pr(Y_1 + Y_2 = 2|X) = Pr(\varepsilon \in \mathcal{R}_{PSNE}((1, 1)|X, \gamma)|X).$$

Inference for the parameters of the model can be performed from here. This result extends beyond two players under some symmetry conditions of strategic effects and has been widely used for
identification purposes (see Berry and Tamer (2007) Section 2.4)).

4.1.3 Iterated dominance behavior

NE behavior presupposes that players’ beliefs are correct. A weaker requirement that is still consistent with rationality and expected-payoff maximization is to allow players to have incorrect beliefs, but assume that these beliefs assign zero probability to dominated strategies. Iteratively removing dominated strategies in this fashion leads to the solution concept of Iterated Dominance. The following analysis was done in Aradillas-López and Tamer (2008). The steps of iterated dominance are as follows.

**Step 1:** Eliminate all dominated strategies in $S$. Define $S^1$ to be the strategy profiles that remain.

**Step 2:** Eliminate all strategies that are not best-responses to beliefs concentrated on $S^1$. Define $S^2$ to be the strategy profiles that remain.

: 

**Step k:** Eliminate all strategies that are not best-responses to beliefs concentrated on $S^{k-1}$. Define $S^k$ to be the strategy profiles that remain.

: 

This process stops when there are no more strategies left to eliminate. The set of strategies $S^*$ that survive is the rationalizable strategies. By construction, all NE strategies must be rationalizable and, therefore, contained in $S^k$ for any $k$. In the $2 \times 2$ game convergence is achieved after at most two steps.

**Step 1:** Dominated strategies are described as follows:

\[ Y_p = 1 \text{ is dominated if and only if } X_p' \beta_p - \epsilon_p < 0 \text{ (i.e., } \epsilon_p > X_p' \beta_p) \]

\[ Y_p = 0 \text{ is dominated if and only if } X_p' \beta_p + \Delta_p - \epsilon_p > 0 \text{ (i.e., } \epsilon_p < X_p' \beta_p + \Delta_p) \]

---

4The concept of rationalizability and its relationship with iterated dominance was developed and analyzed in Pearce (1984) and Bernheim (1984).
Step 2: If there are no dominated strategies, the process ends in Step 1. Otherwise,

If $Y_p = 0$ is dominated:

1. $Y_p = 1$ is dominated if and only if $X'_p\beta_p + \Delta_p - \varepsilon_p < 0$ (i.e., $\varepsilon_p > X'_p\beta_p + \Delta_p$).
2. $Y_p = 0$ is dominated if and only if $X'_p\beta_p + \Delta_p - \varepsilon_p > 0$ (i.e., $\varepsilon_p < X'_p\beta_p + \Delta_p$).

If $Y_p = 1$ is dominated:

1. $Y_p = 1$ is dominated if and only if $X'_p\beta_p - \varepsilon_p < 0$ (i.e., $\varepsilon_p > X'_p\beta_p$).
2. $Y_p = 0$ is dominated if and only if $X'_p\beta_p - \varepsilon_p > 0$ (i.e., $\varepsilon_p < X'_p\beta_p$).

Figure 2 shows the action profiles that survive $k$ rounds of iterated dominance, with $k = 1, 2$.

Let $\mathcal{S}^k(X, \varepsilon, \gamma)$ denote the set of strategies that survive $k$ iterated dominance steps and define

$$\mathcal{R}_{ID}^k(y|X, \gamma) = \{(\varepsilon_1, \varepsilon_2) : y \in \mathcal{S}^k(X, \varepsilon, \gamma)\}$$

Then, under the assumption that strategies survive $k$ steps of iterated dominance,

$$\Pr(Y = y|X) \leq \Pr(\varepsilon \in \mathcal{R}_{ID}^k(y|X, \gamma)|X)$$

for all $y$, a.e $X$.

Assuming iterated dominance instead of NE behavior, the sets $\Theta_I$ and $\Theta_{II}$ described in the NE case would be re-defined replacing $\mathcal{R}_{NE}$ with $\mathcal{R}_{ID}^k$.

4.2 Inference in the incomplete information case

Suppose now that the realization of a subset of payoff shifters is only private information. For simplicity, suppose $X \equiv (X_1, X_2)$ is observed by both players but $\varepsilon_p$ is only privately observed by player $p$. Econometric analysis of incomplete information games typically starts by assuming that beliefs are conditioned on a certain information observed by players. Let $W_p$ denote the information used by player $p$ to condition her beliefs. Assume throughout that $X \in W_p$ (since both players observe $X$). Let $\pi_{-p}(W_p)$ denote $p$’s subjective beliefs for the probability that $Y_{-p} = 1$.

Player $p$’s expected payoff of choosing $Y_p = 1$ is given by

$$\pi_p(1, \pi_{-p}, X_p, \varepsilon_p) = X'_p\beta_p + \Delta_p \cdot \pi_{-p}(W_p) - \varepsilon_p,$$

and $\pi_p(0, \pi_{-p}, X_p, \varepsilon_p) = 0$. We will assume that players follow a threshold-crossing decision rule

$$Y_p = 1 \left\{ X'_p\beta_p + \Delta_p \cdot \pi_{-p}(W_p) - \varepsilon_p \geq 0 \right\}.$$

This can be done without loss of generality under the assumption that $\varepsilon_p|X_p, W_p$ is continuously distributed with unbounded support, so that $\Pr(X'_p\beta_p + \Delta_p \cdot \pi_{-p}(W_p) - \varepsilon_p = 0) = 0$. 

13
Econometric inference in incomplete information games typically proceeds by making more precise assumptions about $W_p$. Given that $X$ is observed by both players, the prototypical econometric study would rely on the following two assumptions in this case:

1. **Conditional independence of players’ private information**: Conditional on $X$, we have $\varepsilon_1 \perp \varepsilon_2$.
2. **Information used by players**: $W_p = X$ for both $p = 1, 2$. 

Figure 2: Action profiles that survive $k$ rounds of iterated dominance

$k = 1$:

$k = 2$ (rationalizable actions):
Thus, expected payoffs \((10)\) are of the form

\[
\pi_p(1, \pi_{-p}, X_p, \varepsilon_p) = X_p' \beta_p + \Delta_p \cdot \pi_{-p}(X) - \varepsilon_p,
\]

Conditional independence of private information is a common assumption in the econometric analysis of incomplete information games, but recent efforts have been made to relax this restriction (see the papers cited in Section 2). Most existing methods can be extended to cases where \(W_p = (X, \xi_p)\) (where \(\xi_p\) is informative for \(\varepsilon_{-p}\)) as long as \(\xi_p\) is observable to the econometrician. For simplicity we consider the case \(W_p = X\).

### 4.2.1 Bayesian Nash equilibrium (BNE) behavior

Suppose \(\varepsilon|X \sim G_X\) and \(\varepsilon_p|X \sim G_{X,p}(\cdot)\). For a given \(\pi \equiv (\pi_1, \pi_2) \in [0,1]^2\), let

\[
\Lambda(\pi|X, \gamma, G_X) = \begin{pmatrix} \frac{\partial \pi_1}{\partial \gamma} \\ \frac{\partial \pi_2}{\partial \gamma} \end{pmatrix} - \begin{pmatrix} G_{1,X}(X_1' \beta_1 + \Delta_1 \cdot \pi_2) \\ G_{2,X}(X_2' \beta_2 + \Delta_2 \cdot \pi_1) \end{pmatrix}
\]

Given \((X, \gamma, G_X)\), a BNE is a pair of beliefs \(\pi^*\) for players 1 and 2 that solves the fixed-point condition

\[
\Lambda(\pi|X, \gamma, G_X) = 0.
\]

Let \(g_{p,X}\) denote the density function corresponding to the distribution \(G_{p,X}\) and let

\[
\nabla_\pi \Lambda(\pi|X, \gamma, G_X) = \begin{pmatrix} 1 \\ -\Delta_1 \cdot g_{1,X}(X_1' \beta_1 + \Delta_1 \cdot \pi_2) \\ -\Delta_2 \cdot g_{2,X}(X_2' \beta_2 + \Delta_2 \cdot \pi_1) \\ 1 \end{pmatrix}
\]

denote the Jacobian of the BNE system with respect to \(\pi\). We say that a solution \(\pi^*\) to the BNE system is **regular** if \(\det(\nabla_\pi \Lambda(\pi^*|X, \gamma, G_X)) \neq 0\). That is,

\[
1 - \Delta_1 \cdot \Delta_2 \cdot g_{1,X}(X_1' \beta_1 + \Delta_1 \cdot \pi_2) g_{2,X}(X_2' \beta_2 + \Delta_2 \cdot \pi_1) \neq 0
\]

By the Implicit Function Theorem, regular BNE are locally unique, well-defined functionals of \((X, \gamma, G_X)\).

### Existence and multiplicity of BNE

Existence of a solution to \((13)\) follows from Brouwer’s Fixed Point Theorem. Sufficient conditions for uniqueness of a solution can be obtained by verifying the Gale-Nikaido conditions (Gale and Nikaido (1965)). If all the principal minors of \(\nabla_\pi \Lambda(\pi|X, \gamma, G_X)\) are positive, then the BNE solution
to (13) will be unique. This will be satisfied if
\[
1 - \Delta_1 \cdot \Delta_2 \cdot g_{1,x}(X'_1 \beta_1 + \Delta_1 \cdot \pi_2) \cdot g_{2,x}(X'_2 \beta_2 + \Delta_2 \cdot \pi_1) > 0 \tag{15}
\]
While (14) ensures local uniqueness, (15) ensures global uniqueness.\footnote{Note that (15) is immediately satisfied if \(\Delta_1 \cdot \Delta_2 \leq 0\), but it may not hold if \(\Delta_1 \cdot \Delta_2 > 0\). Also note that if \(X_1\) or \(X_2\) contain an element with rich enough support, then by restricting attention to regions where such a covariate is sufficiently negative or sufficiently positive, we can make \(g_{1,x}(X'_1 \beta_1 + \Delta_1 \cdot \pi_2) \cdot g_{2,x}(X'_2 \beta_2 + \Delta_2 \cdot \pi_1)\) sufficiently small that (15) is satisfied and uniqueness is achieved. This idea was explored in Aradillas-López (2010). Figure 3 illustrates cases with unique BNE and multiple BNE.}

Figure 3: An illustration of Bayesian-Nash equilibria (BNE)

Identified set
For a given \((X, \gamma, G_X)\) let \(\Pi_{BNE}(X, \gamma, G_X)\) denote the set of all solutions to the BNE system (13). Let
\[
\mathcal{R}_{BNE,p}(1|X, \gamma, G_X) = \left\{ \varepsilon_p : X'_p \beta_p + \Delta_p \pi^*_p - \varepsilon_p \geq 0 \text{ for some } \pi^*_p \in \Pi_{BNE}(X, \gamma, G_X) \right\},
\]
\[
\mathcal{R}_{BNE,p}(0|X, \gamma, G_X) = \left\{ \varepsilon_p : X'_p \beta_p + \Delta_p \pi^*_p - \varepsilon_p < 0 \text{ for some } \pi^*_p \in \Pi_{BNE}(X, \gamma, G_X) \right\},
\]
\[
\mathcal{R}_{BNE}(y_1, y_2|X, \gamma, G_X) = \mathcal{R}_{BNE,1}(y_1|X, \gamma, G_X) \times \mathcal{R}_{BNE,2}(y_2|X, \gamma, G_X)
\]

Note that
\[
Pr(Y = y|X) \leq Pr(\varepsilon \in \mathcal{R}_{BNE}(y|X, \gamma, G_X)) \text{ for all } y, \text{ a.e } X.
\]
The identified set can be characterized from here in the manner described previously. Moment inequalities can be obtained as follows. For a given \((X, \gamma, G_X)\) let

\[
\pi^*_p(X, \gamma, G_X) = \min\{\pi^*_p \in \Pi_{BNE}(X, \gamma, G_X)\}, \quad \pi^-_p(X, \gamma, G_X) = \max\{\pi^-_p \in \Pi_{BNE}(X, \gamma, G_X)\}.
\]

Note that

\[
\begin{align*}
1 \{ X'_p \beta_p + \Delta_p \cdot \pi^*_p(X, \gamma, G_X) - \epsilon_p \geq 0 \} \leq Y_p & \leq 1 \{ X'_p \beta_p + \Delta_p \cdot \pi^-_p(X, \gamma, G_X) - \epsilon_p \geq 0 \} \quad \text{a.e. } X. \\
E[Y_p|X] & \leq G_X \left( X'_p \beta_p + \Delta_p \cdot \pi^*_p(X, \gamma, G_X) \right) \quad \text{a.e. } X. \\
E[Y_p|X] & \geq G_X \left( X'_p \beta_p + \Delta_p \cdot \pi^-_p(X, \gamma, G_X) \right) \quad \text{a.e. } X.
\end{align*}
\]

Inference can proceed from here in the ways described previously.

**Inference assuming a degenerate equilibrium selection mechanism**

In incomplete-information games, point-identification can be obtained by assuming that the underlying equilibrium selection mechanism \(M\) is degenerate (i.e., does not randomize across existing BNE) without having to assume which equilibrium is chosen. This is an important distinction with complete-information games, where simply assuming a degenerate selection mechanism is not enough: one would need to assume which equilibrium is chosen. To illustrate, suppose \(\Pi_{BNE}(X, \gamma, G_X)\) is a finite set (e.g., suppose the regularity condition \(\text{(14)}\) is satisfied). Then,

\[
E[Y_p|X] = E_X \left[ \sum_{\pi_j \in \Pi_{BNE}(X, \gamma, G_X)} 1 \{ X'_p \beta_p + \Delta_p \cdot \pi^-_{-p}(X, \gamma, G_X) - \epsilon_p \geq 0 \} \cdot Pr(M \text{ selects } \pi_j|X, \epsilon_p) \bigg| X \right] \quad (16)
\]

Suppose \(M\) selects one BNE w.p.1. We say in this case that the data is *generated from a single equilibrium*. Let \(\pi^*\) denote this BNE. The previous equation becomes,

\[
E[Y_p|X] = E \left[ 1 \{ X'_p \beta_p + \Delta_p \cdot \pi^*_p(X, \gamma, G_X) - \epsilon_p \geq 0 \} \bigg| X \right] = G_X \left( X'_p \beta_p + \Delta_p \pi^*_p(X, \gamma, G_X) \right) \quad \text{for } p = 1, 2.
\]

But, since \(\pi^*\) is a BNE solution, it follows that

\[
E[Y_p|X] = G_X \left( X'_p \beta_p + \Delta_p E[Y_{-p}|X] \right) \quad \text{for } p = 1, 2. \quad (17)
\]
Since \(E[Y_p|X]\) is nonparametrically identified, inference for the payoff parameters \(\gamma\) can proceed from here using a two-step procedure. In the first step we can nonparametrically estimate \(E[Y_p|X]\) and in the second step we would plug them into the above equation to estimate the parameters of the model.\(^7\) A variety of econometric methods can be applied, including those that allow for a nonparametrically specified \(G_X\). See, for example, Powell, Stock, and Stoker (1989), Klein and Spady (1993), Ahn and Manski (1993), Ichimura (1993), Ahn (1995). Examples of econometric studies that assume a degenerate equilibrium selection mechanism include Seim (2006), Pesendorfer and Schmidt-Dengler (2008), Bajari, Hong, Krainer, and Nekipelov (2010), Aradillas-López (2012). Note that counterfactual analysis would still require assumptions about which equilibrium is selected.

### Dropping the conditional-independence assumption

Assuming independent across players’ private information is a strong limitation shared by most papers in this literature. In general, this assumption has been dropped by either: (i) assuming a fully parametric joint distribution for privately observed shocks (e.g., Xu (2014)) or (ii) assuming that the source of correlation is a shock with just a finite number of support points (e.g, Marcoux (2018)). A completely different approach was proposed in Aradillas-López (2010), where beliefs were assumed to be of the type \(E[Y_p|X,Y_p]\), instead of \(E[Y_p|X,\varepsilon_p]\). This allows, for example, to treat the joint distribution of \((\varepsilon_1, \varepsilon_2)\) nonparametrically without restricting their joint support. Whether beliefs are of the form \(E[Y_p|X,Y_p]\) or \(E[Y_p|X,\varepsilon_p]\) is testable. More work should be devoted to testing conditional-independence of private information.

#### 4.2.2 Iterated dominance behavior

Once again, we can replace BNE with the weaker requirement of iterated dominance. This is done in Aradillas-López and Tamer (2008). We maintain that players follow threshold-crossing decision rules

\[
Y_p = \mathbb{1}\left\{X_p'\beta_p + \Delta_p \cdot \pi_{-p}(X) - \varepsilon_p \geq 0\right\}
\]

but we no longer impose the requirement that beliefs solve the BNE conditions. Iterated dominance can restrict the range of possible beliefs. This can be done naturally because beliefs are bounded in the \([0,1]\) interval.\(^9\) In what follows, let \(G_p(\cdot|X)\) denote the distribution of \(\varepsilon_p|X\).

---

\(^7\)Assuming that the underlying selection mechanism is degenerate and all the data comes from the same equilibrium is an almost universal assumption in econometric models of dynamic games. See, for example, Aguirregabiria and Mira (2007), Pakes, Ostrovsky, and Berry (2007) and Pesendorfer and Schmidt-Dengler (2008).

\(^8\)The approach in Aradillas-López (2010) relies on using, as an inferential range, a subset of values of \(X\) where a Gale-Nikaido condition of the type given in (15) is satisfied and therefore the BNE is unique.

\(^9\)The approach we describe here can be generally applied to any game where actions have bounded support. The binary-choice game is the simplest such case, since \(Y_p \in \{0,1\}\).
**Step 1:** Since $\pi_p(X) \in [0,1]$ and actions are strategic-substitutes,

\[ 1\{X_p\beta_p + \Delta_p - \varepsilon_p > 0\} \leq Y_p \leq 1\{X_p\beta_p - \varepsilon_p \geq 0\}. \]

Therefore,

\[ \frac{G_p(X_p\beta_p + \Delta_p |X)}{\pi^1_p(X)} \leq Pr(Y_p = 1 |X) \leq \frac{G_p(X_p\beta_p |X)}{\pi^1_p(X)}. \]

**Step 2:** Eliminate all strategies that are not best-responses to beliefs that survive Step 1. It follows that

\[ 1\{X_p\beta_p + \Delta_p \cdot \pi^1_p(X) - \varepsilon_p > 0\} \leq Y_p \leq 1\{X_p\beta_p + \Delta_p \cdot \pi^1_p(X) - \varepsilon_p \geq 0\}. \]

Therefore,

\[ \frac{G_p(X_p\beta_p + \Delta_p \cdot \pi^1_p(X) |X)}{\pi^1_p(X)} \leq Pr(Y_p = 1 |X) \leq \frac{G_p(X_p\beta_p + \Delta_p \cdot \pi^1_p(X) |X)}{\pi^1_p(X)}. \]

\[ : \]

\[ : \]

**Step k:** Eliminate all strategies that are not best-responses to beliefs that survive Step $k-1$. It follows that

\[ 1\{X_p\beta_p + \Delta_p \cdot \pi^{k-1}_p(X) - \varepsilon_p > 0\} \leq Y_p \leq 1\{X_p\beta_p + \Delta_p \cdot \pi^{k-1}_p(X) - \varepsilon_p \geq 0\}. \]

Therefore,

\[ \frac{G_p(X_p\beta_p + \Delta_p \cdot \pi^{k-1}_p(X) |X)}{\pi^{k-1}_p(X)} \leq Pr(Y_p = 1 |X) \leq \frac{G_p(X_p\beta_p + \Delta_p \cdot \pi^{k-1}_p(X) |X)}{\pi^{k-1}_p(X)}. \]

\[ : \]

\[ : \]

Unlike the complete-information case, this process can continue indefinitely. Rationalizable choice probabilities (and therefore BNE choice probabilities) are always included in the interval

\[ \left[ \pi^k_p(X), \pi^{\bar{k}}_p(X) \right] \]
for any $k$. From here, a characterization of the identified set under $k$ steps of iterated dominance can proceed analogously to the previous cases discussed. Figure 4 illustrates the iterative procedure for $k = 1, 2, 3$.

**Figure 4: Illustration of iterated dominance in the incomplete-information game**

5 Discrete games with richer action spaces: recent results

Many real-world applications cannot be handled by a binary choice game. Recently, identification and inference results have been obtained in discrete games where the action space is rich but has an ordinal property. Suppose players’ actions are real-valued and the action spaces in our normal-form game (1) are discrete and ordinal,

$$S_p = \{s_p^1, s_p^2, \ldots, s_p^M_p\},$$

(18)

with $s_p^j < s_p^{j+1}$ for all $j$. None of the results we will discuss below presupposes precise knowledge of the elements in $S_p$, relying only on the ordinal nature and the assumption that either (i) the upper
and lower bounds of \( S_p \) are known (or can be estimated) or (ii) inference is focused in the interior of \( S_p \). Binary-choice games are always included as a special case of every model we discuss next.

5.1 Inference in games with complete information

5.1.1 Nonparametric results

Aradillas-López (2011) (henceforth AL11) derives testable implications for Nash Equilibrium outcomes in ordered discrete games with action spaces such as (18). The results obtained do not parametrize payoff functions but rely instead only on shape restrictions. The following properties are assumed to hold w.p.1.

- **Concavity of payoffs:** For any \( y_p \in S_{-p} \),
  \[
  u_p(s_p, y_p) - u_p(s_{p}^{j-1}, y_p) > u_p(s_{p}^{j+1}, y_p) - u_p(s_p, y_p) \quad \forall \ s_p \in S_p.
  \]

- **Nonincreasing differences:** For any \( s_p \in S_p \) and \( y_p, y_p' \in S_{-p} \), if \( u_p(s_p^j, y_p) \geq u_p(s_p^j, y_p') \), then
  \[
  u_p(s_{p}^{j+1}, y_p) - u_p(s_p^j, y_p) \geq u_p(s_{p}^{j+1}, y_p') - u_p(s_p^j, y_p').
  \]

Figure 5 illustrates the nonincreasing differences condition. It is reminiscent of properties in supermodular games, and it results in a no-crossing property for payoff functions.

Figure 5: Nonincreasing differences
AL11 assumes complete-information NE behavior without ruling out mixed-strategies. Concavity of payoffs and the independent-mixing nature of NE imply that, in any NE, players can randomize across at most two possible actions, and these actions must be adjacent. Fix an action profile \( y \equiv (y_p)_{p=1}^P \), where \( y_p = s_p^j \in \mathcal{S}_p \) for each \( p \) (we write \( j \) instead of \( j_p \) for notational simplicity). Let

\[
S(y) = \{ (a_p)_{p=1}^P : a_p \in \{ s_p^{j-1}, s_p^j, s_p^{j+1} \} \}
\]  

(19)

The support of any NE where \( y \) is played with positive probability must be a subset of \( S(y) \). Next, for each \( p \) let\(^\text{10}\)

\[
\pi_p(\cdot, S(y)) = \max_{y_p \in S(y)} u_p(\cdot, y_p), \quad \text{and} \quad \pi_p(\cdot, S(y)) = \min_{y_p \in S(y)} u_p(\cdot, y_p).
\]

By concavity and nonincreasing differences, the action profile \( y \) can be played with positive probability in a NE only if

\[
\pi_p(s_p^j, S(y)) < \pi_p(s_p^{j+1}, S(y)) \quad \text{and} \quad u_p(s_p^j, S(y)) > u_p(s_p^{j+1}, S(y)) \quad \forall \ p.
\]

Our shape conditions restrict which NE can co-exist. Suppose there exists a NE where \( y = (y_p)_{p=1}^P \) is played with positive probability. Now, take another action profile \( y' = (y'_p)_{p=1}^P \). Then, there co-exists a NE where \( y' \) is played with positive probability only if, for each \( p \), one of the following holds

\[
\begin{align*}
\text{(i)} & \quad \pi_p(\cdot, S(y')) > \pi_p(\cdot, S(y)) \quad \text{and} \quad u_p(\cdot, S(y')) < u_p(\cdot, S(y)), \\
\text{(ii)} & \quad \pi_p(\cdot, S(y')) \leq \pi_p(\cdot, S(y)) \quad \text{and} \quad y'_p \leq y_p, \\
\text{(iii)} & \quad u_p(\cdot, S(y')) \geq \pi_p(\cdot, S(y)) \quad \text{and} \quad y'_p \geq y_p.
\end{align*}
\]

(20)

Let

\[
\Pi_p^a(y, y') = 1 - I\{ \pi_p(\cdot, S(y')) \leq \pi_p(\cdot, S(y)) \quad \text{and} \quad y'_p > y_p \},
\]

\[
\Pi_p^b(y, y') = 1 - I\{ u_p(\cdot, S(y')) \leq \pi_p(\cdot, S(y)) \quad \text{and} \quad y'_p < y_p \},
\]

\[
\Pi_p^c(y, y') = \min\{ \Pi_p(y, y'), \Pi_p^b(y, y') \},
\]

\[
\Pi(y, y') = \prod_{p=1}^P \Pi_p^c(y, y').
\]

Note that \( \Pi(y, y') \) is the indicator for the event that at least one of the conditions in (20) is satisfied for each \( p \).

**Definition 1 (NE outcome)** We say that \( y \equiv (y_p)_{p=1}^P \) is a NE outcome if there exists a NE where \( y \) is played with positive probability. Let \( \mathcal{E} \) denote the collection of all NE outcomes in the game.

\(^{10}\text{Concavity and nonincreasing differences imply that payoff functions do not cross. See Figure 5.}\)
Pick any action profile $y = (y_p)_{p=1}^P$. If we maintain that $Y \equiv (Y_p)_{p=1}^P \in \mathcal{E}$, then

$$1[Y = y] \leq 1\{y \in \mathcal{E}\} \leq \Pi^*(Y, y) \text{ w.p.} 1.$$ \hspace{1cm} (21)

If payoffs are completely unrestricted beyond the shape restrictions, $\Pi^*$ is unobserved and (21) cannot be used directly. AL11 proposes a way to exploit (21) by assuming the existence of a strategic-interaction index, a known function $f_p$ (possibly multi-valued) that captures the direction of strategic interaction.

- **Existence of a strategic-interaction index:** For each $p$ there exists a function $f_p : \mathcal{S}_{-p} \rightarrow \mathbb{R}^{d_p}$ ($d_p \geq 1$) known to the researcher such that, for any $y_{-p}, y'_{-p} \in \mathcal{S}_{-p}$
  
  (i) If $f_p(y'_{-p}) = f_p(y_{-p})$, then $u_p(\cdot, y_{-p}) = u_p(\cdot, y'_{-p})$.
  
  (ii) If $f_p(y'_{-p}) \geq f_p(y_{-p})$ (element-wise), then $u_p(\cdot, y_{-p}) \leq u_p(\cdot, y'_{-p})$.
  
  (iii) If $f_p(y'_{-p}) \not\geq f_p(y_{-p})$ and $f_p(y_{-p}) \not\geq f_p(y'_{-p})$, then nothing is implied about the ordinal relationship between $u_p(\cdot, y_{-p})$ and $u_p(\cdot, y'_{-p})$.

Thus, ranking the strategic index $f_p$ enables us to rank payoff functions.

Let

$$H_p(y, y') = \mathbb{1}\{f_p(u_{-p}) \geq f_p(v_{-p}) \forall u_{-p} \in S(y), v_{-p} \in S(y')\}.$$ 

$H_p(y, y') = 1$ implies $u_p(\cdot, u_{-p}) \leq u_p(\cdot, v_{-p}) \forall u_{-p} \in S(y), v_{-p} \in S(y')$. Therefore,

$$H_p(y, y') \leq \mathbb{1}\{\Pi_p(\cdot, S(y)) \leq u_p(\cdot, S(y'))\}.$$ 

And,

$$\Pi^p_p(y, y') \leq 1 - H_p(y', y) \cdot \mathbb{1}\{y'_p > y_p\} \equiv \Pi^b_p(y, y'),$$

$$\Pi^p_p(y, y') \leq 1 - H_p(y', y) \cdot \mathbb{1}\{y'_p < y_p\} \equiv \Pi^b_p(y, y'),$$

$$\Rightarrow \Pi^p_p(y, y') \leq \min\{\Pi^a_p(y, y'), \Pi^b_p(y, y')\} \equiv \Pi_p(y, y'),$$

$$\Rightarrow \Pi^*(y, y') \leq \prod_{p=1}^P \Pi^p_p(y, y') \equiv \Pi^*(y, y').$$

And from (21) we obtain

$$1[Y = y] \leq 1\{y \in \mathcal{E}\} \leq \Pi^*(Y, y) \text{ w.p.} 1.$$
Let $C$ denote a collection of outcomes. Then,
\[
1 \{Y \in C\} \leq 1 \{C \cap \mathcal{E} \neq \emptyset\} \leq \max_{y \in C} \{I^*(Y, y)\},
\]
\[
1 \{C \subseteq \mathcal{E}\} \leq \min_{y \in C} \{I^*(Y, y)\}.
\]

The first and second lines relate to the event that some $C$ contains a NE outcome and that every outcome in $C$ is a NE outcome, respectively. Let $X$ denote the set of observable covariates by the econometrician. Bounds for the probability of equilibrium outcomes are obtained from the previous inequalities,
\[
\Pr(Y = y | x) \leq \Pr(y \in \mathcal{E} | X) \leq E[I^*(Y, y) | X],
\]
\[
\Pr(Y \in C | X) \leq \Pr(C \cap \mathcal{E} \neq \emptyset | X) \leq E\left[\max_{y \in C} \{I^*(Y, y)\} \bigg| X\right],
\]
\[
\Pr(C \subseteq \mathcal{E} | X) \leq E\left[\min_{y \in C} \{I^*(Y, y)\} \bigg| X\right],
\]

Since the bounds in (22) are nonparametrically identified, confidence intervals (CI) for $\Pr(y \in \mathcal{E} | X)$, $\Pr(C \cap \mathcal{E} \neq \emptyset | X)$ and $\Pr(C \subseteq \mathcal{E} | X)$ can be constructed using, e.g, the methods described in Imbens and Manski (2004) and Stoye (2009). AL11 also describe bounds for other probabilities of interest, such as the propensity to select a particular $y$ conditional on being a NE profile.

### 5.1.2 Parametric results

A parametric ordered-response game of complete information was first analyzed in Davis (2006) under parametrization restrictions that yield a unique pure-strategy NE. A much more general model is studied in Aradillas-Lopez and Rosen (2019) (henceforth AR19). Action spaces are as described in (18). AR19 impose restrictions on payoff functions that turn it effectively into a simultaneous ordered-response model. First, they assume that we can express payoffs as,
\[
u_p(y_p, y_{-p}, X_p, \epsilon_p),
\]

where $X_p$ is observed by the econometrician while $\epsilon_p$ is an unobserved scalar. Let $X \equiv (X_p)_{p=1}^P$ and $\epsilon \equiv (\epsilon_p)_{p=1}^P$. Strict concavity of $u_p$ with respect to $y_p$ is maintained. AR19 then impose an assumption about how payoff functions “shift” with $\epsilon_p$.

- **Increasing differences in $(y_p, \epsilon_p)$:** For any $y_{-p} \in \mathcal{S}_{-p}$ and a.e $X_p$, the following holds: If $\epsilon'_p > \epsilon_p$ and $y'_p > y_p$, then
  \[
u_p(y'_p, y_{-p}, X_p, \epsilon'_p) - \nu_p(y_p, y_{-p}, X_p, \epsilon_p) < \nu_p(y'_p, y_{-p}, X_p, \epsilon'_p) - \nu_p(y_p, y_{-p}, X_p, \epsilon_p).
\]
Figure 6 illustrates the increasing differences property.

Figure 6: Increasing differences in \((y_p, \epsilon_p)\)

Fix \(X_p, \epsilon_p\) and take any \(y_p\). Let \(y_p^*(y_p, X_p, \epsilon_p)\) denote \(p\)'s best-response. The key implication of the increasing-differences shape restriction is the following. There exists a sequence of non-overlapping thresholds

\[ \epsilon_p^*(s_j^1, y_p, X_p) < \epsilon_p^*(s_j^2, y_p, X_p) < \cdots < \epsilon_p^*(s_j^{M_p}, y_p, X_p), \]

such that

\[ y_p^*(y_p, X_p, \epsilon_p) = s_j^i \iff \epsilon_p^*(s_j^i, y_p, X_p) < \epsilon_p \leq \epsilon_p^*(s_j^{i+1}, y_p, X_p). \]  

Fix an action profile \(y \equiv (y_p)_{p=1}^P\), where \(y_p = s_j^i \in S_p\) for each \(p\) (again, we write \(j\) instead of \(j_p\) for notational simplicity). Let \(S(y)\) be as described in (19). Strict concavity and independent-mixing imply that if \(y\) is played with positive probability in a NE, then the support of this NE must be a subset of \(S(y)\). By definition, the profile \(y\) can be played in a NE only if \(s_j^i = y_p^*(a_{-p}, X_p, \epsilon_p)\) for some \(a_{-p} \in S(y)\) and this is true for all \(p\). Let \(X \equiv (X_p)_{p=1}^P\) and \(\epsilon = (\epsilon_1, \ldots, \epsilon_P)'\)

\[ R_p(y, X_p) = \bigcup_{a_{-p} \in S(y)} \left[ \epsilon_p^*(s_j^i, a_{-p}, X_p), \epsilon_p^*(s_j^{i+1}, a_{-p}, X_p) \right], \]

\[ R(y, X) = R_1(y, X_1) \times \cdots \times R_P(y, X_P) \]

(24)
Let the NE outcomes $E$ be as described in Definition 1. Pick any action profile $y$. If we maintain the assumption that the observed choice profile $Y$ is a NE outcome, then
\[
1\{Y = y\} \leq 1\{y \in E\} \leq 1\{\epsilon \in \mathcal{R}(y, X)\}.
\]
In particular,
\[
1\{\mathcal{R}(Y, X) \in \mathcal{C}\} \leq 1\{\epsilon \in \mathcal{C}\} \quad \forall \mathcal{C} \subset \mathbb{R}^P.
\]
And since $X$ is observable to the econometrician, inference can ultimately be based on
\[
\Pr(\mathcal{R}(Y, X) \subseteq \mathcal{C} | X) \leq \Pr(\epsilon \in \mathcal{C} | X) \quad \forall \mathcal{C} \subset \mathbb{R}^P, \text{ a.e } X.
\]
(25)

If we limit attention to pure-strategy NE (PSNE) and assume that $Y$ is always the realization of a PSNE, the sets in (24) would simply become
\[
\mathcal{R}_p(y, X_p) = \left[ \epsilon^*_p(s^j_{p, y_p}, X_p, \epsilon_p | \gamma_p), \epsilon^*_p(s^{j+1}_{p, y_p}, X_p, \epsilon_p | \gamma_p) \right],
\]
A parametric model

Suppose we parametrize payoff functions as $u_p(y_p, y_{-p}, X_p, \epsilon_p | \gamma_p)$. Denote $\gamma_p \equiv (\gamma_p^p)_{p=1}^P$. We have now parametric expressions for the thresholds described in (23). Denote them by $\epsilon^*_p(s^j_{p, y_p}, X_p, \epsilon_p | \gamma_p)$ and let
\[
\mathcal{R}_p(y, X_p | \gamma_p) = \bigcup_{a \in S(y)} \left[ \epsilon^*_p(s^j_{p, y_p, a_p, X_p | \gamma_p}), \epsilon^*_p(s^{j+1}_{p, y_p, a_p, X_p | \gamma_p}) \right],
\]
\[
\mathcal{R}(y, X | \gamma) = \mathcal{R}_1(y, X_1 | \gamma_1) \times \cdots \times \mathcal{R}_P(y, X_P | \gamma_P)
\]
Next, suppose the joint distribution of $\epsilon | X$ is also parametrized. For simplicity, suppose this parametrization assumed $\epsilon \perp X$ and $\epsilon \sim G(\rho)$. The parameters of the model are $\theta \equiv (\gamma, \rho) \in \Theta$ ($\Theta$ is parameter space). For a given set $\mathcal{C}$, let
\[
H(C, X, \theta) = \Pr(\mathcal{R}(Y, X | \gamma) \subseteq \mathcal{C} | X) - \Pr(\epsilon \in \mathcal{C} | X).
\]
Let $\mathcal{C}$ denote a pre-specified class of sets in $\mathbb{R}^P$. Based on (25), the identified set for $\theta$ can be described as
\[
\Theta_I = \left\{ \theta \in \Theta: H(C, X, \theta) \leq 0 \quad \text{a.e } X, \forall C \in \mathcal{C} \right\}
\]
(26)

Example: A two-player game

As an illustration, AR19 focus on a two-player game where payoff functions are modeled as
\[
u_p(y_p, y_{-p}, X_p, \epsilon_p | \gamma_p) = y_p \times \left( \delta + X_p' \beta - \Delta_p \cdot y_{-p} - \eta \cdot y_p + \epsilon_p \right),
\]
with $\gamma_p \equiv (\delta, \beta', \Delta_p, \eta)'$. Concavity will be satisfied by specifying $\eta > 0$ in the parameter space. Increasing-differences is satisfied by this parametrization for any $\gamma_p$. AR19 focus, for illustration purposes, on a strategic-substitutes case. Therefore, $\Delta_p \geq 0$ in the parameter space for $p = 1, 2$. The action space in this example is assumed to be $S_p = \{0, 1, 2, \ldots\}$. Given this specification, the thresholds in (23) have the functional form

$$
\varepsilon^*_p(y_p, y_{-p}, X_p | \gamma_p) = \begin{cases} 
\infty & \text{if } y_p = 0, \\
\eta \cdot (2y_p - 1) + \Delta_p y_{-p} - \delta - X_p' \beta & \text{if } y_p \geq 1.
\end{cases}
$$

The joint distribution of $\varepsilon = (\varepsilon_1, \varepsilon_2)'$ is parameterized as an FGM copula $G(\varepsilon | \rho)$ with logistic marginal distributions.

**Identification results assuming PSNE**

AR19 assume that players always choose a pure-strategy Nash equilibrium (PSNE). As discussed above, this simplifies the set $R(y|X, \gamma)$ to

$$
R(y|X, \gamma) = 2 \times \prod_{p=1}^{k} \left[ \varepsilon^*_p(y_p, y_{-p}, X_p | \gamma_p), \varepsilon^*_p(y_p + 1, y_{-p}, X_p | \gamma_p) \right]
$$

Ruling out mixed-strategies in the binary-choice version of this game produces a unique prediction for $Y_1 + Y_2$ for all co-existing equilibria (see Section 4.1.2 above and Berry and Tamer (2007, Section 2.4)) and all payoff parameters can be identified from there. However, this is no longer the case as soon as we introduce more than two possible actions. With our shape restrictions, ruling out mixed-strategies only pins down a parametric expression for $Pr(Y = (0, 0)|X)$ because, if $(0, 0)$ is a PSNE, it is the only PSNE. Let $\lambda \equiv \eta - \delta$. Then,

$$
Pr(Y = (0, 0)|X) = G(\lambda - X_1' \beta, \lambda - X_2' \beta | \rho).
$$

Let $\bar{\Theta} \equiv (\lambda, \beta', \rho)'$. Then, $\bar{\Theta}$ can be point-identified and estimated using, for example, MLE.

**Inference**

Let $\Theta^* \equiv (\eta, \Delta_1, \Delta_2)$ and note that $\Theta = (\bar{\Theta}, \Theta^*)$. AR19 propose an inferential procedure where $\bar{\Theta}$ is first estimated using MLE and then inference for the remaining subset of parameters $\Theta^*$ is conducted using the information from the conditional moment inequalities described in (26) for a pre-specified class of sets $C$. An outline of their procedure is as follows. Let $f_c$ denote the

---

11 Existence of a PSNE w.p.1 in this case follows from Tarski’s Fixed Point Theorem, see e.g. Theorem 2.2 of Vives (1999) or Section 2.5 of Topkis (1998).

12 Using the results in Chesher and Rosen (2017), AR19 characterize the class of core-determining sets $C^*$ that leads to a sharp characterization of the identified set in (26). In their empirical application, for computational simplicity they
density of $X$. For a given $X$ and $\theta$ let

$$T(C, X, \theta) = \left( Pr(\mathcal{R}(Y, X)|\gamma) \subseteq C|X) - Pr(\varepsilon \in C|\rho) \right) \cdot f_x(X)$$

$$\mathcal{M}(C, \theta) = E_X \left[ \max \{ T(C, X, \theta), 0 \} \right],$$

$$\mathcal{M}(\theta) = \sum_{C \in \mathcal{C}} \mathcal{M}(C, \theta)$$

Note that $\mathcal{M}(\theta) \geq 0$ for all $\theta$ and $\mathcal{M}(\theta) = 0$ if and only if $T(C, X, \theta) \leq 0$ a.e $X$, for all $C \in \mathcal{C}$. Density-weighting has computational and theoretical advantages. The identified set in (26) can be rewritten as

$$\Theta_I = \{ \theta \in \Theta: \mathcal{M}(\theta) = 0 \}.$$ 

Suppose for illustration purposes that all the elements in $X \in \mathbb{R}^d$ are continuously distributed. AR19 propose an estimator for $\mathcal{M}(\theta)$ of the following form

$$\hat{T}(C, x, \theta) = \frac{1}{n \cdot h^d} \sum_{i=1}^{n} \left( 1 \{ \mathcal{R}(Y_i, x)|\gamma) \subseteq C \} - Pr(\varepsilon \in C|\theta) \right) K \left( \frac{X_i - x}{h_n} \right),$$

$$\hat{\mathcal{M}}(C, \theta) = \frac{1}{n} \sum_{i=1}^{n} H(C|X_i, \theta) \cdot 1 \{ H(C|X_i, \theta) \geq -b_n \},$$

$$\hat{\mathcal{M}}(\theta) = \sum_{C \in \mathcal{C}} \hat{\mathcal{M}}(C, \theta),$$

where $b_n \to 0$ is a nonnegative sequence converging to zero at an appropriate rate. Let

$$\Theta_I^* = \{ \theta \in \Theta_I: H(C|X, \theta) < 0 \ \text{a.e} \ X, \forall C \in \mathcal{C} \}$$

This is the set of all parameter values that satisfy (26) as strict inequalities almost surely. AR19 describe conditions (smoothness and regularity conditions of conditional moments, manageability and empirical-process conditions for the functional forms assumed for payoffs and for the parametric distribution of $\varepsilon$, as well as kernel conditions and bandwidth convergence restrictions) such that

$$\sqrt{n} \cdot \hat{\mathcal{M}}(\theta): \begin{cases} \overset{p}{\longrightarrow} 0 & \forall \theta \in \Theta_I^*, \\ \overset{d}{\longrightarrow} \mathcal{N} \left( 0, E \left[ \varphi_n(Y_i, X_i, \theta)^2 \right] \right) & \forall \theta \in \Theta_I \setminus \Theta_I^*, \\ \overset{p}{\longrightarrow} +\infty & \forall \theta \in \Theta_I. \end{cases}$$

use a subset of this class.

Notice that density-weighting in the definition of $T(C, X, \theta)$ helps produce a simple estimator.
Let $\hat{\theta}_{MLE}$ be the MLE estimator of $\theta$ and, for a given $\theta = (\bar{\theta}, \theta^*)$ let

$$\hat{V}(\theta) = \sqrt{n} \left( \hat{\theta}_{MLE} - \bar{\theta} \right).$$

The proposal in AR19 is to construct a Wald-type statistic based on $\hat{V}(\theta)$. However, this will require some form of regularization since the asymptotic variance of $\sqrt{n} \cdot \hat{m}(\theta)$ will be zero if $\theta \in \Theta^*_I$. Let

$$\hat{t}(\theta) = \hat{V}(\theta) \hat{\Sigma}^{-1}(\theta_{MLE}, \theta^*) \hat{V}(\theta).$$

where $\hat{\Sigma}(\theta_{MLE}, \theta^*)$ is a regularized estimator of the asymptotic variance of $\hat{V}(\theta)$. Let $r \equiv \dim(\theta)$. AR19 show that $\hat{t}(\theta)$:

$$\sim \chi^2_r \text{ if } \theta \in \Theta^*_I,$$

$$\sim \chi^2_{r+1} \text{ if } \theta \in \Theta_I \setminus \Theta^*_I,$$

$$\to +\infty \text{ if } \theta \not\in \Theta_I.$$

These asymptotic properties immediately suggest how to construct a CS for $\theta$. For a pre-specified target coverage probability $1 - \alpha$ let $c(\chi^2_r, 1 - \alpha)$ denote the $(1 - \alpha)$-th quantile of the $\chi^2_r$ distribution. By the properties of the chi-squared distribution, we have $c(\chi^2_{r+1}, 1 - \alpha) > c(\chi^2_r, 1 - \alpha)$. Let

$$\hat{S}^\theta_{1-\alpha} = \{ \theta \in \Theta : \hat{t}(\theta) \leq c(\chi^2_{r+1}, 1 - \alpha) \}$$

AR19 show that $\hat{S}^\theta_{1-\alpha}$ satisfies $\lim \inf_{n \to \infty} \Pr(\theta \in \hat{S}^\theta_{1-\alpha}) \geq 1 - \alpha$. This inferential approach has the computational advantage that critical values do not have to be obtained by resampling methods due to the asymptotically pivotal properties of the statistic used in its construction.

### 5.2 Inference in games with incomplete information

Aradillas-López and Gandhi (2016) (henceforth AG16) study ordered discrete games with incomplete information under the assumption of BNE behavior. Their first assumption is that payoff functions can be expressed as $u_p(y_p, y_{-p}, Z_p)$ with $Z_p = (X, \varepsilon_p)$, where $\varepsilon_p$ is privately observed by player $p$ and $X$ is publicly observed by all players as well as the econometrician ($\varepsilon_p$ is not restricted to be a scalar). Their next assumption is that payoffs can be expressed as

$$u_p(Y_p, Y_{-p}, Z_p) = u_p^a(y_p, Z_p) - u_p^b(y_p, Z_p) \cdot \eta_p(y_{-p}, X).$$

where $u_p^a, u_p^b$ and $\eta_p$ are scalar functions. The key feature of (28) is that $\eta_p$ depends on $Z_p$ solely through $X$. The inferential object of interest in AG16 is the strategic index $\eta_p$. Players’ private information is assumed to be mutually independent conditional on $X$ and beliefs are assumed to
be conditioned on $X$. For any set of beliefs $\pi_{−p} : S_{−p} \rightarrow [0, 1]$, the expected utility of player $p$ of choosing $Y_p = y_p$ is given by\(^{14}\)

$$
\overline{u}_p(y_p, \pi_{−p}, Z_p) = \sum_{y_{−p} \in S_{−p}} \pi_{−p}(y_{−p}|X) \cdot u_p(y_p, y_{−p}, Z_p) = u_p^b(y_p, Z_p) - u_p^b(y_p, Z_p) \cdot \eta_p(\pi_{−p}, X),
$$

where $\eta_p(\pi_{−p}, X) = \sum_{y_{−p} \in S_{−p}} \pi_{−p}(y_{−p}|X) \cdot \eta_p(y_{−p}, X)$. The following shape restriction normalizes the “strategic meaning” of the index $\eta_p$.

- **Marginal benefit of** $Y_p$ **is nonincreasing in** $\eta_p$: For a.e $Z_p$ and any $y, y' \in S_p$,

$$
y > y' \implies u_p^b(y, Z_p) \geq u_p^b(y', Z_p).
$$

This implies that the index $\eta_p$ and player $p$’s optimal strategy are strategic substitutes, but does not presuppose that the game itself is of strategic substitutes, since the index $\eta_p$ can shift in many possible ways with opponents’ actions. The constructive implication of the previous assumption is as follows. Take any pair of beliefs $\pi_{−p}$ and $\pi'_{−p}$. Then,

$$
\left[\overline{u}_p(y, \pi_{−p}, Z_p) - \overline{u}_p(y', \pi_{−p}, Z_p)\right] - \left[\overline{u}_p(y, \pi'_{−p}, Z_p) - \overline{u}_p(y', \pi'_{−p}, Z_p)\right] \\
= \left(\eta_p(\pi_{−p}, X) - \eta_p(\pi'_{−p}, X)\right) \cdot \left(u_p^b(y, Z_p) - u_p^b(y', Z_p)\right)
$$

It follows that, if $\eta_p(\pi'_{−p}, X) \geq \eta_p(\pi_{−p}, X)$, then

$$
\overline{u}_p(y, \pi'_{−p}, Z_p) - \overline{u}_p(y', \pi'_{−p}, Z_p) \leq \overline{u}_p(y, \pi_{−p}, Z_p) - \overline{u}_p(y', \pi_{−p}, Z_p) \quad \forall y > y'.
$$

Now, consider two sets of beliefs, $\pi_{−p}$ and $\pi'_{−p}$, each having a unique optimal choice for the corresponding expected-utility functions. Denote them as $y^*_p(\pi_{−p}, Z_p)$ and $y'^*_p(\pi'_{−p}, Z_p)$ respectively. From our previous assumption, for a.e $Z_p$,

$$
\eta_p(\pi'_{−p}, X) > \eta_p(\pi_{−p}, X) \implies 1\{y^*_p(\pi_{−p}, Z_p) \leq y_p\} \leq 1\{y'^*_p(\pi'_{−p}, Z_p) \leq y_p\} \quad \forall y_p \in S_p. \quad (29)
$$

To see why, note first that, by definition, $\overline{u}_p(y, \pi_{−p}, Z_p) - \overline{u}_p(y^*_p(\pi_{−p}, Z_p), \pi_{−p}, Z_p) < 0$ for all $y > y^*_p(\pi_{−p}, Z_p)$. But from our assumptions above, $\eta_p(\pi'_{−p}, X) > \eta_p(\pi_{−p}, X)$ implies

$$
\overline{u}_p(y, \pi'_{−p}, Z_p) - \overline{u}_p(y'^*_p(\pi'_{−p}, Z_p), \pi'_{−p}, Z_p) \leq \overline{u}_p(y, \pi_{−p}, Z_p) - \overline{u}_p(y^*_p(\pi_{−p}, Z_p), \pi_{−p}, Z_p) < 0
$$

\(^{14}\)The independent-mixing nature of Nash equilibrium is once again key here. Without this property, beliefs could be conditioned on the potential choice $y_p$ and expected payoffs would not have the structure exploited to obtain the results that follow.
for all $y > y_p^*(\pi_{-p}, Z_p)$. Thus, we must have $y_p^*(\pi_{-p}', Z_p) \leq y_p^*(\pi_{-p}, Z_p)$ and the inequality in (29) follows.

The inequality in (29) will produce a testable implication for the model if we assume that players’ beliefs are such that they always produce a unique optimal choice. AG16 show that under these assumptions,

$$Cov\left(\mathbb{1}\{Y_p \leq y_p\}, \eta_p(Y_{-p}, X)\right|\left.X\right) \geq 0 \quad \forall \ y_p \in \mathcal{S}_p, \ a.e \ X. \quad (30)$$

The binary-choice version of this result was obtained in de Paula and Tang (2012) under assumptions of symmetry (a condition not required in AG16). Note that if the selection mechanism $\mathcal{M}$ is degenerate w.p.1 (e.g., if there is always a unique BNE), then $Cov\left(\mathbb{1}\{Y_p \leq y_p\}, \eta_p(Y_{-p}, X)\right|\left.X\right) = 0$ w.p.1. AG16 detail inferential and testing procedures based on (30) and test-statistics of the type used in AR19 and described previously. They show how to test a particular functional form for the index $\eta_p$ or parametrize it and construct a CS for the parameters.

6 Concluding remarks and directions for future research

Recent advancements in the econometrics of partially identified models have made it possible to do robust inference in static games with multiple solutions. However, more work needs to be done in order to make these models a better approximation to real world problems. Some lines of research that need to be explored further include:

(i) **Nonequilibrium models**: The assumption that economic agents have correct beliefs and perfect models about others is an elegant theoretical framework, but perhaps not a realistic approximation to real-world behavior in many instances. More work needs to be done to develop econometric models that allow for incorrect beliefs and, importantly, methods that can explore more precisely how agents choose from within a space of allowable (e.g., rationalizable) beliefs. Characterizing testable implications (preferably nonparametric) of different solution concepts to help researchers discriminate among competing behavioral models can produce valuable contributions to the literature.

(ii) **Allowing a more flexible information structure in incomplete-information games**: More work needs to be done to relax the assumption of conditional independence of private information. The current state of the art either assumes that the source of correlation is very simple (e.g., a commonly observed shock with finite support) or imposes a fully parametrized model. Endogeneity (correlation between observed and unobserved payoff shifters) must be

---

15 The first fully nonparametric, non-experimental test for different levels of rationality in game-theoretic models was developed by Kosenkova (2018) for first-price auctions.
allowed if these models are to be of any practical use. As in the previous point, developing nonparametric specification tests to help choose among different information structures remains an important topic.

(iii) **Exploring richer action spaces:** As we illustrated in this paper, recent advances have been made to extend binary choice games into models with richer strategy spaces. However, more work needs to be done to bring these methods closer to real world problems. For example, multi-valued actions (e.g., adding an element of strategic interaction to the single-agent models studied in [Heckman (1978)](#) or action spaces with unobserved, random censoring (e.g., capacity constraints).

(iv) **Counterfactual analysis:** As we discussed, many of these models are only partially identified. This can discourage practitioners from using robust-inferential methods that only product set predictions. The question then is, how to achieve a balance between robustness and predictions from counterfactual analysis. One way is to develop methods to refine the set-predictions for these outcomes (borrowing perhaps from recent advances in methods to do inference in a subset of parameters in partially identified models [Kaido, Molinari, and Stoye (2019)](#)). Alternatively, we can re-frame the problem entirely and do inference on a population objective function based on the counterfactual policy we want to analyze. Results from the theoretical literature on policy robustness (see [Hansen and Sargent (2016)](#)) can be explored.

**References**


