Discerning Solution Concepts for Discrete Games*

Nail Kashaev† Bruno Salcedo‡

First version: January 3, 2014
This version: March 16, 2020

Abstract The empirical analysis of discrete complete-information games has relied on behavioral restrictions in the form of solution concepts, such as Nash equilibrium. Choosing the right solution concept is crucial not just for the identification of payoff parameters, but also for the validity and informativeness of counterfactual exercises and policy implications. We say that a solution concept is discernible if it is possible to determine whether it generated the observed data on the players' behavior and covariates. We propose a set of conditions that make it possible to discern solution concepts. In particular, our conditions are sufficient to tell whether the players' choices emerged from Nash equilibria. We can also discriminate between rationalizable behavior, maxmin behavior, and collusive behavior. Finally, we identify the correlation structure of unobserved shocks in our model using a novel approach.

Keywords Discrete Games · Testability · Identification · Incomplete models · Market entry

JEL classification C52 · C72

*A previous version of this paper was circulated under the title "Identification of Solution Concepts for Discrete Games". We are thankful for the supervision of Joris Pinkse, Sung Jae Jun and Andrés Aradillas-López, as well as the useful comments from Victor H. Aguiar, Roy Allen, Bulat Gafarov, Paul Grieco, Marc Henry, Brendan Kline, Robert Marshall, Francesca Molinari, Salvador Navarro, and Mark Roberts. We also thank the attendants of the 2014 Spring Midwest Theory and Trade Conference at IUPUI, the 2014 Summer Meeting of The Econometric Society at the University of Minnesota, the 2015 International Game Theory Conference at Stony Brook University, and the 11th World Congress of the Econometric Society. We gratefully acknowledge the Human Capital Foundation (http://hcfoundation.ru/en/), and particularly Andrey P. Vavilov, for research support through the Center for the Study of Auctions, Procurement, and Competition Policy (http://capcp.psu.edu/) at the Pennsylvania State University. All remaining errors are our own.

†Department of Economics, Western University, nail.kashaev.ru, nkashaev@uwo.ca.

‡Department of Economics, Western University, brunosalcedo.com, bsalcedo@uwo.ca.

1. Introduction

In Game Theory, solution concepts impose restrictions on the behavior of players given their payoffs. The most popular solution concept is Nash equilibrium (NE) (Nash, 1951). Solution concepts are often used to establish theoretical results, to identify payoff parameters, and to derive policy and welfare implications from counterfactual analyses.[1] However, there may exist different solution concepts that are observationally equivalent but yield different payoff parameters, theoretical implications, and counterfactual predictions (see Section 2 for an example). Thus, it is important to understand when one can tell a solution concept apart from other distinct solution concepts. In such cases, we say that the solution concept is *discernible*.

We consider multiplayer binary-action games of complete information, similar to the classic entry game from Bresnahan and Reiss (1990). We maintain the assumption that the players' choices can display any form of rationalizable behavior in the sense of Bernheim (1984) and Pearce (1984). We provide a set of conditions that are sufficient to establish discernibility of any solution concept stronger than rationalizability. For instance, it is possible to determine whether the players' decisions arise from NE. Moreover, if they do arise from NE, then they cannot be consistent with any other form of rationalizable behavior that is not NE. We also identify all the payoff parameters, including those governing the correlation structure of the unobserved heterogeneity. To the best of our knowledge, this is the first formal result that identifies the correlation parameters using a solution concept weaker than pure-strategy NE.

Usually, the testable implications of a solution concept (e.g., NE) can be used to determine whether it is consistent with or *could have* generated the observed data.[2] Our results allow the researcher to answer the question of whether the

[1]The classic revealed-preference approach to the identification of payoff parameters in discrete games of complete information assumes that the choice of each player is a best response to the observed choices of other players (Bjorn and Vuong, 1984, Jovanovic, 1989, Bresnahan and Reiss, 1990). This is tantamount to assuming that the players' choices constitute pure strategy NE. The approach can be generalized to allow mixed strategy NE (Tamer, 2003, Bajari et al., 2010), rationalizable strategies (Aradillas-López and Tamer, 2008, Kline, 2015), or general convex solution concepts (Beresteanu et al., 2011, Galichon and Henry, 2011). See De Paula (2013) for a review of the literature.

[2]For example, one can construct a model specification test based on the results from Beresteanu et al. (2011) or Galichon and Henry (2011).

solution concept actually generated the data. This question is important because of several reasons. A solution concept can be consistent with the data and, at the same time, yield misleading counterfactual predictions. For example, this could happen if an alternative solution concept generated the data, and the two solution concepts are observationally equivalent. We provide an example in Section 2. Establishing discernibility of NE precludes this possibility and helps to establish the validity of counterfactual analysis and policy implications.

Discernibility is also useful in making sharper counterfactual predictions. One can always assume a less restrictive solution concept (in our case rationalizabilty), build the confidence set for the payoff parameters, and then construct robust confidence bands for the counterfactual of interest. However, these bands can be uninformative because of the weakness of the restrictions imposed on behavior. If one shows that a stronger solution concept (e.g., NE) generated the data, and this solution concept is discernible, then one can build more informative bounds for the counterfactual predictions. In other words, our methodology allows the researcher to determine the strongest restrictions on behavior that are still consistent with the observed data.

Discernibility also has practical implications that may reduce the computational burden. Suppose that the researcher is considering several distinct solution concepts, and establishes discernibility of all of them. In this case, discernibility implies that at most one of the solution concepts under consideration can explain the data. Hence, if a given solution concept explains the data, then the researcher can automatically rule out the other alternatives.

A potential difficulty is that not all solution concepts are distinct relative to one another. For example, every pure-strategy Nash equilibrium (PNE) is also a NE. Hence, the players choices could be both NE and PNE at the same time, in which case it would be impossible to tell PNE and NE apart. We can tell two discernible solution concepts apart as long as they are distinct, in that they give different predictions with positive probability. For example, we can tell apart PNE from Nash behavior that assigns positive probability to strictly mixed Nash equilibria. See Section 4.4.

Our strategy to establish discernibility relies on two assumptions. First, we assume that the researcher observes covariates with full support satisfying an exclusion restriction. Second, we assume that the excluded covariates generate enough variation in the conditional distribution of payoffs conditional on covariates. In particular, we require the family of these conditional distributions to be boundedly complete.[3] Using these assumptions, we identify the distribution of payoffs and the distribution of outcomes conditional on both the observed and unobserved characteristics of the environment. Knowing these distributions allows us to establish discernibility of solution concepts.

We are not the first to exploit the power of completeness assumptions coupled with exclusion restrictions to discriminate between behavior patterns. Berry and Haile (2014) apply a strategy similar to ours to a model of oligopolistic competition that allows, among other things, to discriminate between different models of competition. A significant difference between their setting and ours is that they consider continuous games, while we consider discrete games. They crucially rely on having an uncountable set of outcomes to relax the completeness assumption to some extent.

Identification of the payoff parameters is not necessary for discernibility. In Section 5, we relax rationalizability and allow for some forms of collusive behavior and ambiguity aversion in the sense of Gilboa and Schmeidler (1989). This comes at the expense that some of the payoff parameters are no longer point identified. However, we still can establish discernibility of a large class of solution concepts.

2. Motivating Example

We begin with a simple example to motivate the meaning and the importance of discernibility of solution concepts. First, we show that two distinct solution concepts—pure strategy Nash equilibrium (PNE) and a behavioral solution concept called strategic ambiguity aversion (SAA)—can be observationally equivalent. That is, they can generate the same distribution over observables. Hence, it is impossible to discern PNE and SAA in our example. Next, we show that PNE and SAA not being discernible can lead to incorrect quantitative and qualitative policy recommendations.

[3] Completeness of a family distribution is a well-known concept both in Statistical and Econometrics literature. See Andrews (2011). Newey and Powell (2003) and Darolles et al. (2011) use a completeness assumption to establish non-parametric identification for conditional moment restrictions. Blundell et al. (2007) use bounded completeness to achieve identification of Engel curves. Hoderlein et al. (2012) impose bounded completeness in the context of structural models with random coefficients.

Two firms $i \in \{1, 2\}$ simultaneously choose whether to enter a market ($y_i = 1$) or not ($y_i = 0$). Firm i's profit is given by[4]

$$y_i \cdot \Big[\eta_0(1 - y_{-i}) - \mathbf{e}_i\Big],$$

where (i) y_{-i} is the choice of i's competitor; (ii) $\eta_0 \geq 0$ is a fixed parameter that measures the effect of competition and is unknown by the researcher; and (iii) $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$ is a vector of payoff shocks unobserved by the researcher. We assume that $\mathbf{e}$ is supported on $\mathbb{R}^2$ and admits a probability density function that is symmetric around the 45-degree line (e.g., $\mathbf{e}_1$ and $\mathbf{e}_2$ are independent standard normal random variables). The firms observe both η_0 and $\mathbf{e}$. That is, the game is of complete information. The researcher observes (can consistently estimate) the distribution of outcomes $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)$. Many of the simplifications we make in this example are for exposition purpose only and are relaxed in subsequent sections.

2.1. Failure of Discernibility

Following Bresnahan and Reiss (1990), the classic approach to analyze entry games is to assume that the firms' choices always constitute PNE. Suppose that a researcher wants to test this assumption under the milder assumption that firm behavior is rationalizable. In our example, rationalizability is equivalent to assuming that the firms can choose any action that survives two rounds of elimination of strictly dominated strategies. When $e_i < 0$, entering the market is strictly dominant for firm i. When $e_i > \eta_0$, staying out of the market is strictly dominant for firm i. This results in four regions of the payoff space in which the game has a unique rationalizable outcome. In the remaining region—the *multiplicity region*—rationalizability imposes no restrictions on behavior.

Consider the following solution concept. Firm behavior is rationalizable, but firms never enter when there are multiple rationalizable outcomes. This could happen, for instance, if the firms were ambiguity averse and used maxmin strategies when facing strategic uncertainty. This solution concept is called SAA and is analyzed in Mass (2019). The predictions of SAA with $\eta_0 = \eta$ for different realizations of $\mathbf{e}$ are illustrated in the left panel of Figure 1.

[4]Throughout the paper, we use boldface font to denote random variables, and regular font for realizations of random variables and deterministic objects.

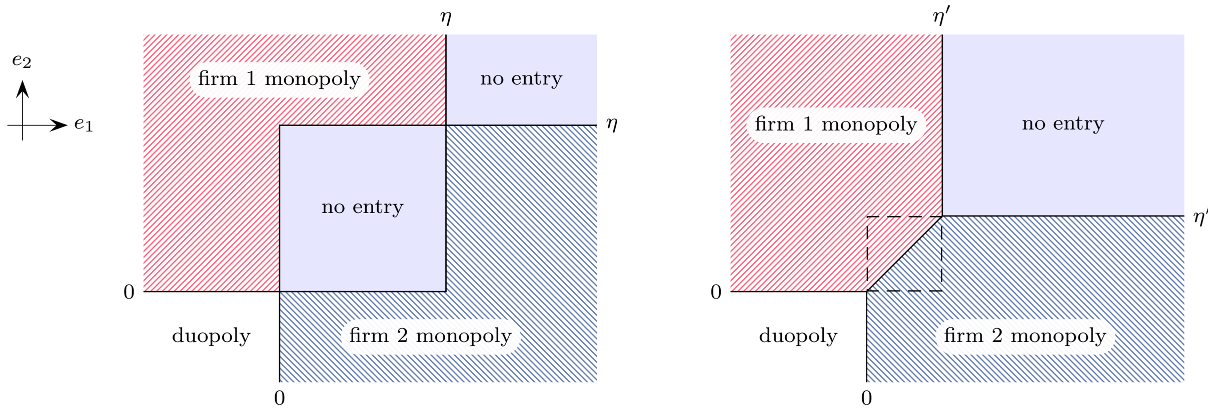

Figure 1 – Two observationally equivalent models: strategic ambiguity aversion (left) and pure-strategy Nash equilibrium (right).

We will show that PNE and SAA can produce the exact same distributions over observables. According to PNE, only one firm enters in the multiplicity region. Since either of the two firms could be the one that enters in equilibrium, we need to specify an equilibrium selection rule. We assume that firms always play the equilibria according to which the most profitable firm is the one that enters. The predictions of PNE with such selection rule and $\eta_0 = \eta'$ are illustrated in the right panel of Figure 1.

The duopoly region is the same under both solution concepts. For a fixed value of the competition effect, the no-entry region is smaller under PNE. However, since η_0 is unknown to the researcher, it is possible to set $\eta' < \eta$ so that both models assign the same probability to no entry. Since both models imply the same probability of both duopoly and no entry, they also imply the same probability of having a monopoly. Note that both the distribution of shocks and each of the two models are symmetric around the 45-degree line. Hence, each of the two monopolies are equally likely under both solution concepts. The formal proof is in Appendix C.2 in the online supplement.

Despite being very different, SAA and PNE can imply identical distributions over outcomes. Hence, in this example, it is impossible to determine whether the data *is* generated by PNE. At best, the researcher can tell whether the data *can be* explained by PNE. That is, PNE is not discernible and, for the exact same reason, neither is SAA. The fact that we use different parameter values for each of the two models is unimportant. For instance, Proposition 4.5 in Section 4.3 establishes a general nondiscernibility result that can be applied even if the payoff parameters are assumed to be the same under different solution concepts.

2.2. Importance of Discernibility

Next, we show that PNE and SAA can generate opposite counterfactual predictions in our example. Therefore, the failure of discernibility can lead to incorrect policy recommendations. Suppose that a policymaker wants to increase the number of markets that are served by at least one firm. As a policy instrument, she can choose to offer a subsidy $\tau > 0$ to one firm, say Firm 1, for entering markets in which Firm 2 does not enter. Entry subsidies are commonly used to incentivize the provision of strategic infrastructure such as broadband internet access (Goolsbee, 2002). The specific subsidy scheme we analyze allows for a stark and simple exposition. Appendix C.3 in the online supplement presents a similar result with a more realistic subsidy scheme.

Suppose that the data is generated by SAA, but the policymaker evaluates the policy assuming that firms always play PNE. We have already established that the policymaker cannot refute her assumption because PNE and SAA are observationally equivalent. However, under PNE the size of the competition effect must be smaller than the one under SAA (i.e., $\eta' < \eta$). Thus, assuming the incorrect solution concept would lead to inconsistent estimates of η_0. This, in turn, might lead to flawed welfare evaluations. Since both models are examples of rationalizable behavior, one may expect that an inconsistent estimator of the competition effect will not affect the qualitative implications of different policy interventions. Here, this is not the case.

PNE in the multiplicity region always predict monopolies. Hence, under the PNE hypothesis, all markets are served except for those in which not entering is dominant for both firms. The policy being evaluated decreases the probability of the latter region (see Figure 2). Therefore, under the policymaker's assumptions, the policy unambiguously reduces the number of markets without service independently of the parameter values.

However, under SAA, the effect of the policy is always smaller than under PNE, and it can even have the opposite direction for some parameter values. This can happen because the policy also increases the probability of the multiplicity region and, under strategic ambiguity aversion, firms never enter in this region. A firm might be willing to forego the subsidy for fear of another firm entering the market, which would result in negative profits. The net effect of the policy on the probability of monopolies is given by the difference between the probabilities of regions E^+ and E^- in Figure 2. The adverse effect can actually dominate and

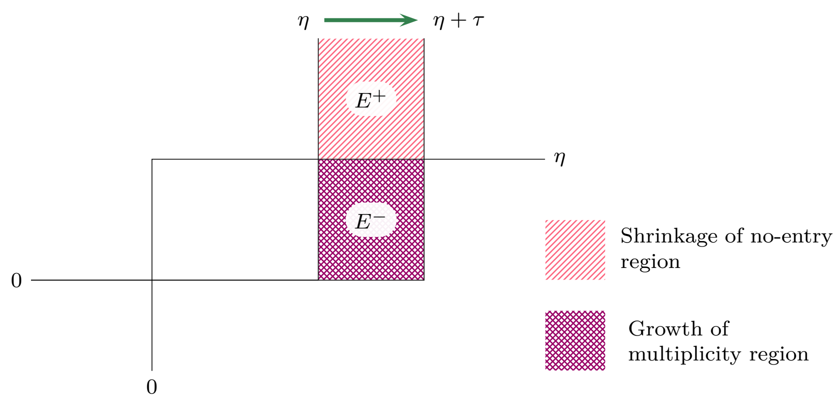

Figure 2 – Effect of the proposed policy.

the policy can increase the probability that a market is not served. For example, one can verify that this is the case whenever $\mathbf{e}_1$ and $\mathbf{e}_2$ are independent standard normal random variables and $\Phi(\eta) > 3/4$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

3. Framework

The motivating example from Section 2 shows that it is possible for distinct solution concepts to be observationally equivalent while implying different policy recommendations. In what follows, we introduce a framework that rules out that possibility. Our assumptions on the data generating process and the covariates observed by the researcher guarantee that a large class of solution concepts are discernible, in a formal sense to be defined.

3.1. Payoffs

There are $d_I < +\infty$ players (firms) indexed by $i \in I = \{1, \ldots, d_I\}$. Each player chooses an action $y_i \in Y_i = \{0, 1\}$. Appendix C.1 in the online supplement generalizes our analysis to games with many actions. The set of outcomes is $Y = \times_{i \in I} Y_i$. Let $y_{-i} = (y_j)_{j \neq i}$ denote the vector of actions from i's opponents.

Player i's payoffs from outcome y are given by [5]

$$y_i \cdot \Big(\alpha^0_{i,y_{-i}}(\mathbf{w}) + \beta^0_i(\mathbf{w})\mathbf{z}_i - \mathbf{e}_i\Big),$$

where (i) $\mathbf{w}$ is a vector of observed player and market characteristics with support $W \subseteq \mathbb{R}^{d_W}$; (ii) $\mathbf{z} = (\mathbf{z}_i)_{i \in I}$ is a vector of player-specific covariates that does not share any common components with $\mathbf{w}$; (iii) $\mathbf{e} = (\mathbf{e}_i)_{i \in I}$ is a vector of payoff shocks unobserved by the researcher; and (iv) $\beta^0_i, \alpha^0_{i,y_{-i}} : W \to \mathbb{R}$ are unknown functions. The covariate $\mathbf{z}_i$ is assumed to be a scalar for simplicity. If there are additional player specific covariates, we include them in $\mathbf{w}$. We assume that all the parameters and payoff shocks are common knowledge among the players. That is, the game is of complete information.

Example 1 If $\alpha^0_{i,y_{-i}}(w) = \delta^0_{ii}(w) + \sum_{j \neq i} \delta^0_{ij}(w) y_j$, then the specification corresponds to an entry game as in Bresnahan and Reiss (1990) or Berry (1992). The value of $\delta^0_{ii}(w) + \beta^0_i(w) z_i - e_i$ measures the monopoly payoff of firm i. For $i \neq j$ the value of $\delta^0_{ij}(w)$ measures the strategic effect of the market presence of firm j on i's profits. The strategic effect of the presence of a firm on payoffs of competitors (the signs of $\delta^0_{ij}(w)$, $i \neq j$) is unrestricted.

Example 2 If $\alpha^0_{i,y_{-i}}(w) = \delta^0_i(w) \cdot \mathbb{1}(\sum_{j \in I} y_j \geq \tau^0(w))$, with $\tau^0(w) \in [0, d_I]$, the model corresponds to a regime-change game. The term $\beta^0_i(w) z_i - e_i$ captures the individual cost of participating in a revolt. The threshold $\tau^0(w)$ determines the number of participants required for the revolt to be successful. And $\delta^0_i(w)$ captures the benefit to i from participating in a successful revolt. This payoff structure can also be used to study coordinated-action problems (Rubinstein, 1989), bank runs and currency attacks (Morris and Shin, 2003), or tacit collusion in oligopolistic markets (Green et al., 2014).

We impose standard assumptions on payoffs (see, for instance, Jia, 2008, Ciliberto and Tamer, 2009, Bajari et al., 2010, and Ciliberto et al., 2018).

Assumption 1

(i) The support of $\mathbf{z}$ conditional on $\mathbf{w} = w$ is $Z = \mathbb{R}^{d_I}$ for all $w \in W$.

(ii) $\beta^0_i(w) \neq 0$ for all i and $w \in W$.

[5] We use boldface font to denote random variables and vectors.

Assumption 1 requires the player-specific covariates $\mathbf{z}$ to have full support and be relevant conditional on all other covariates $\mathbf{w}$. In entry games, examples of continuous firm-specific covariates could be the logarithm of the distance of the market to the existing network of each firm, or to the firms' headquarters. These distances have been used by Ciliberto and Tamer (2009) and Ciliberto et al. (2018) to analyze the airline industry. Examples of shared covariates in $\mathbf{w}$ could be demographic variables such as market size (e.g., market population) or market income (e.g., per capita income). While there is a restriction on $\beta_i^0(\cdot)$, we do not impose any restrictions on $\alpha_{i,-y}^0(\cdot)$.

Assumption 2 $\mathbf{e}|(\mathbf{z}=z,\mathbf{w}=w) \sim N(0,\Sigma^0(w))$ for all z and w, where $\Sigma^0 : W \to \mathbb{R}^{d_I \times d_I}$ is such that $\Sigma^0(w)$ is a positive definite symmetric matrix and $\Sigma^0_{ii}(w) = 1$ for all $w \in W$ and $i \in I$.

The normality assumption is common in applied work and helps to simplify the exposition. In Appendix A.1, we replace it with two weaker assumptions. The first one imposes restrictions on the tails of the distribution of $\mathbf{e}$. The second one requires the distribution of $\mathbf{e}$ to constitute a boundedly complete family of distributions. The normality assumption also implies that the probability that a player obtains the same payoffs from different outcomes is zero. Thus, we do need to worry about situations when players may be indifferent between actions. The requirement $\Sigma^0_{ii}(w) = 1$ is a scale normalization. Note that we allow the payoff shocks to be correlated across players.

To keep the notation tractable, we group covariates and payoff parameters as follows. Let $\mathbf{x} = (\mathbf{z},\mathbf{w})$ be the vector of all observed covariates. Let $\alpha = (\alpha_{i,y_{-i}}(\cdot))_{i\in Y, y\in Y}$, $\beta = (\beta_i(\cdot))_{i\in I}$, and $\theta = (\alpha,\beta,\Sigma(\cdot)) \in \Theta$. Hence, we can define the *payoff indices* $\pi(x,e,\theta) = (\pi_{i,y}(x,e,\theta))_{i\in I, y\in Y}$ by

$$\pi_{i,y}(x,e,\theta) = y_i \cdot \Big(\alpha_{i,y_{-i}}(w) + \beta_i(w) z_i - e_i\Big).$$

The true value of the payoff parameters is denoted by $\theta_0 = (\alpha_0,\beta_0,\Sigma^0(\cdot))$.

3.2. Distribution of Play

An important object for our analysis is the *distribution of play* h_0, defined as the conditional distribution of $\mathbf{y}$ given $\mathbf{x}$ and $\mathbf{e}$. That is,

$$h_0(y, x, e) = \Pr(\mathbf{y} = y | \mathbf{x} = x, \mathbf{e} = e).$$

The distribution of play describes the joint behavior of the players as a function of market and player characteristics. It is a nonparametric latent parameter. Let $h_0(x, e) = (h_0(y, x, e))_{y \in Y}$, and let H be the set of all possible distributions of play. Given $h, h' \in H$, we say that $h = h'$ if and only if $h(\mathbf{x}, \mathbf{e}) = h'(\mathbf{x}, \mathbf{e})$ a.s.. Note that, by construction, $h_0(x, e)$ belongs to the d_Y-dimensional simplex.

We impose the following restriction on the distribution of play. It limits the way the player-specific covariates and shocks can affect the behavior of the players.

Assumption 3 (Exclusion Restriction) There exists a measurable function $\tilde{h}_0$ such that $h_0(\mathbf{x}, \mathbf{e}) = \tilde{h}_0(\mathbf{w}, \mathbf{v})$ a.s., where $\mathbf{v} = (\mathbf{v}_i)_{i \in I}$ and $\mathbf{v}_i = \beta_i^0(\mathbf{w})\mathbf{z}_i - \mathbf{e}_i$.

Assumption 3 is a joint assumption on h_0 and θ_0. It requires that the player-specific covariates and shocks can affect choices *only* via the index $\mathbf{v}$, whose distribution depends on the value of θ_0. It says that $\mathbf{y}$ is independent of $\mathbf{z}$ and $\mathbf{e}$ conditional on $\mathbf{v}$ and $\mathbf{w}$. Note that, given θ_0, the realizations of $\mathbf{v}$ and $\mathbf{w}$ are sufficient to pin down the payoff indices. Hence, Assumption 3 can be interpreted as requiring that, after conditioning on the realization of $\mathbf{w}$, the players are payoff driven. If there are two markets with the exact same payoff indices and the same realization of $\mathbf{w}$, then the distribution over outcomes should be the same. Assumption 3 is implied by the assumptions made in Bajari et al. (2010).[6]

Under Assumption 3 the distribution of play can be robust to policy interventions. For example, if one wants to evaluate policies that only affect firms indirectly through the prices of inputs. The firms might care about the changes in prices, but not about the source of these changes. In situations where this assumption is reasonable, knowing θ_0 and h_0 is sufficient to analyze policies that

[6] More specifically, Bajari et al. (2010) assume that players make choices randomizing among the different NE of the game. Their Assumption 6 requires the selection probabilities to be measurable with respect to the latent utility indices, in our notation. An analogous assumption could be imposed on the selection mechanisms of any model satisfying Assumptions 2.2–2.4 in Beresteanu et al. (2011). Doing so would imply our Assumption 3.

only operate through the payoff indices.

3.3. Solution Concepts

Although Assumption 3 imposes some structural restrictions to the behavior of players, one may still want to impose additional economic restrictions. These restrictions come in the form of solution concepts such as rationalizability or NE. Solution concepts often depend on the characteristics of the environment. Hence, we allow for the restrictions arising from solution concepts to depend on the payoff parameters.

Definition 1 A *solution concept* is a function $S : \Theta \to 2^H$.

For example, suppose that players choose actions simultaneously and only use rationalizable strategies, i.e., strategies that survive the iterated elimination of strictly dominated strategies. Let $S_R(\theta)$ be the set of h such that, given the payoff indices $\pi(x, e, \theta)$, $h(x, e)$ assigns positive probability only to rationalizable outcomes for all x and e.[7]

The Nash hypothesis is that the behavior of the players always constitutes NE of the simultaneous-move game in pure or mixed strategies. There can be multiple NE and, in such cases, there is no consensus on which equilibria are more likely to arise. In order to assume as little as possible about the equilibrium selection, one must allow for arbitrary mixtures of equilibria. The actual distribution of outcomes could be any point in the convex hull of the set of the distribution over outcomes implied by NE. Let $S_N(\theta)$ be the set of h such that, for all x and e, $h(x, e)$ belongs to the convex hull of the set of distributions over outcomes implied by NE of the game given the payoff indices $\pi(x, e, \theta)$. $S_N(\theta)$ exactly captures the predictions of the Nash hypothesis.

Note that the NE solution concept is *nested* into rationalizability. That is, $S_N(\theta) \subseteq S_R(\theta)$ for all θ. Also, S_N is a *convex* solution concept in that $S_N(\theta)$ is a convex set for all θ.

The distribution of play completely characterizes behavior. However, in general, there are at least two reasons to work with restrictions that are coming

[7]Since $S(\theta)$ is a collection of distributions of play that depend on covariates, defining solution concepts in terms of θ implicitly allows behavior to depend on the covariates. See also Section 3.4.

from Economic Theory, that is, solution concepts. First, solution concepts might make for more credible counterfactual analyses because they are also supported by nonempirical arguments (Dawid, 2016). For instance, one may argue that the policy intervention considered in Section 2.2 would not affect whether firms play PNE. However, since only one firm is subsidized, it is possible that the subsidized firm will be more likely to enter in markets with multiple PNE. In this case, the distribution of play would not be policy invariant, but the predictions based on the solution concept would remain accurate.[8]

A second reason to focus on solution concept is portability. A solution concept can make predictions related to changes in some fundamental characteristics of the environment. For example, NE is well defined for two-player and three-player games. In contrast, the distribution of play cannot be easily extrapolated to make predictions if the number of players changes.

Moreover, a solution concept might be relevant beyond the specific application being considered. Many solution concepts from Economic Theory are general theories of behavior. Finding evidence in support for a solution concept in one setting, provides support for its use in other settings. For instance, Walker and Wooders (2001), Chiappori et al. (2002), and Gauriot et al. (2016) have tested the implications of the Nash hypothesis in the context of penalty kicks and tennis serves.[9] And their analysis is often used to justify the use of NE in general settings unrelated to sports.

For most of the text, we assume that the players' behavior is rationalizable. This assumption is common in the literature (see, for instance, Aradillas-López and Tamer, 2008 and Kline, 2015). Section 5 relaxes this assumption.

Assumption 4 (Rationalizability) $h_0 \in S_R(\theta_0)$.

[8]Suppose we define a dummy covariate to indicate whether the policiy is active. The distribution of play would be policy invariant if such covariate was excluded in that it affected behavior only through payoffs. Even if the distribution of play is not policy invariant, it can still be consistent with Assumption 3. This is because Assumption 3 only requires $\mathbf{z}$ to be excluded. Policy dummies can be included as part of the non-excluded covariates.

[9]These papers consider only zero-sum games. It is not entirely clear whether it is possible to generalize their methodology to general-sum games.

3.4. Relation Between Solution Concepts, Distributions of Play, and Selection Mechanisms

This section clarifies the relation between solution concepts, the distribution of play, and selection mechanisms. Our distribution of play is a complete econometric model in the sense of Tamer (2003) and Manski (1988) in that it "asserts that a random variable $\mathbf{y}$ is a function of a random pair ($\mathbf{x}$,[$\mathbf{e}$]) where $\mathbf{x}$ is observable and [$\mathbf{e}$] is not" (Tamer, 2003, pp. 150). In other words, h_0 is an "empirical" solution concept that completely describes behavior of players without any economic restrictions. In contrast, many solution concepts arising from Economic Theory are generally incomplete in that, even knowing the value of the parameters and all the characteristics of the environment, there can be multiple solutions. Both NE and rationalizability fall under this category. Our approach is to take the distribution of play as a primitive, and model incomplete solution concepts as sets of complete models that depend on the parameters of the environment.

An alternative approach to ours is to define a solution concept as a random set $\mathrm{Sol}(\mathbf{x}, \mathbf{e}, \theta)$ consisting of possible distributions p over outcomes, which depend on the characteristic of the environment and the payoffs (e.g., Beresteanu et al., 2011 and Bajari et al., 2010). Then, one would complete the model with a selection mechanism that assigns probabilities $\mathrm{sel}(\,\cdot\,|x, e, \theta)$ to the different possible distributions emerging from the solution concept. The distribution of play could be defined as a weighted average of the different distributions with weights determined by the selection mechanism. For instance, if $\mathrm{Sol}(x, e, \theta)$ is finite,

$$h(y, x, e) = \sum_{p \in \mathrm{Sol}(x,e,\theta)} p(y) \cdot \mathrm{sel}(p|x, e, \theta).$$

Under some technical measurability assumptions, both approaches are mathematically equivalent in terms of the relation between solution concepts, distributions of play, and the data (see Section 2 in Beresteanu et al., 2011). The difference between the two approaches is that our approach emphasizes the distribution of play, which is well defined independently of any solution concept. In contrast, selection mechanisms can only be defined relative to a specific solution concept. The following example demonstrates the relation between them.

Example 1 (continued) Fix some x and suppose that $I = \{1, 2\}$, the strategic effects have negative signs ($\delta^0_{12}(w), \delta^0_{21}(w) < 0$), and the firms always play NE, selecting each NE with equal probabilities. Depending on the realizations of $\mathbf{e}$ there are at most three NE. When there is a unique NE, then the distribution of play assigns full probability to the equilibrium outcome. When there are three NE (firm 1 monopoly, firm 2 monopoly, and a mixed one), then $h_0(x, e)$ is the equally-weighted mixture of the distributions over outcomes implied by these three NE. For instance,

$$\begin{aligned} h_0((1,0), x, e) &= p_1((1,0), x, e) \cdot 1/3 + p_2((1,0), x, e) \cdot 1/3 + p_m((1,0), x, e) \cdot 1/3 \\ &= 1 \cdot 1/3 + 0 \cdot 1/3 + p_m((1,0), x, e) \cdot 1/3 \\ &= \frac{1}{3} + \frac{1}{3} \cdot \frac{e_2 - \delta_{22}^0(w) - \beta_2^0(w) z_2}{\delta_{21}^0(w)} \cdot \left(1 - \frac{e_1 - \delta_{11}^0(w) - \beta_1^0(w) z_1}{\delta_{12}^0(w)}\right), \end{aligned}$$

where $p_i(y, x, e)$ and $p_m(y, x, e)$ are the probability that outcome y is played under firm i monopoly NE and the mixed strategy NE, respectively. Note that $h_0((1,0), x, e)$ is a function of w, $\beta_1^0(w) z_1 - e_1$, and $\beta_2^0(w) z_2 - e_2$. Hence, the distribution of play in this example satisfies our exclusion restriction.

4. Discernibility of Rationalizable Solution Concepts

4.1. Definition of Discernibility

Recall that θ_0 and h_0 denote the true payoff parameters and the true distribution of play. Let $\Psi \subseteq \Theta \times H$ denote a set of possible values that (θ_0, h_0) can take. We are interested in whether solution concepts, in particular NE, are discernible according to the following definition:

Definition 2 Given a set $\Psi \subseteq \Theta \times H$, a solution concept S is said to be *discernible* relative to Ψ if there do *not* exist (θ, h), $(\theta', h') \in \Psi$ such that $h \in S(\theta)$, $h' \notin S(\theta')$, and

$$\mathbb{E}[\, h(\mathbf{x}, \mathbf{e}) | \mathbf{x}; \theta \,] = \mathbb{E}[\, h'(\mathbf{x}, \mathbf{e}) | \mathbf{x}; \theta' \,] \text{ a.s..}$$

Note that

$$\mathbb{E}[\, h_0(y, \mathbf{x}, \mathbf{e}) | \mathbf{x}; \theta_0 \,] = \Pr(\mathbf{y} = y | \mathbf{x}) \text{ a.s.,}$$

for all $y \in Y$. Hence, $\mathbb{E}[\, h_0(\mathbf{x}, \mathbf{e})|\mathbf{x}; \theta_0 \,]$ is identified (can be consistently estimated from observed data on outcomes and covariates). Thus, the definition of discernibility leads to two properties that fully characterize the relationship between the observed (estimable) distribution of $\mathbf{y}$ conditional on $\mathbf{x}$ and any given solution concept S:

(i) If the data is generated by a given solution concept, then it cannot be explained by something else: if $h_0 \in S(\theta_0)$, then

$$(\Pr(\mathbf{y} = y|\mathbf{x}))_{y \in Y} \neq \mathbb{E}[\, h(\mathbf{x}, \mathbf{e})|\mathbf{x}; \theta \,]$$

with positive probability for all $(\theta, h) \in \Psi$ such that $h \notin S(\theta)$.

(ii) If the data is not generated by a given solution concept, then it cannot be explained by it: if $h_0 \notin S(\theta_0)$, then

$$(\Pr(\mathbf{y} = y|\mathbf{x}))_{y \in Y} \neq \mathbb{E}[\, h(\mathbf{x}, \mathbf{e})|\mathbf{x}; \theta \,]$$

with positive probability for all $(\theta, h) \in \Psi$ such that $h \in S(\theta)$.

4.2. Identification of θ_0 and h_0

In order to establish discernibility of solution concepts, we first establish identification of the payoff parameters and the distribution of play. Later on, we also consider environments where identification of payoff parameters fails to hold.

Proposition 4.1 *Under assumptions 1–4, θ_0 is identified.*

The proof of the proposition is in Appendix A. The identification of the payoff parameters β_0 and α_0 follows standard arguments that exploit the large support of our player-specific covariates. However, to the best of our knowledge, Proposition 4.1 is the first result in the literature that identifies the unknown correlation structure of payoffs in games of complete information under solution concepts weaker than PNE.[10] The proof involves looking at the limits of the partial derivative of $\Pr(\mathbf{y} = y|\mathbf{x} = x)$ with respect to one of the excluded covariates for a specific

[10]See Kline (2015) for an identification result assuming either PNE or independent shocks.

outcome vector y along specific rays in the support of excluded covariates.[11]

To identify h_0 note that, under Assumption 3, the excluded covariates $\mathbf{z}$ generate variation in the observed conditional distribution over outcomes without changing h_0. This is because h_0 is affected by $\mathbf{z}$ *only* via the conditional distribution of the index $\mathbf{v}$ conditional on $\mathbf{z}$. This exogenous variation yields the following identification result.

Proposition 4.2 *Under assumptions 1–3, if θ_0 is identified, then h_0 is identified.*

Proof. Suppose towards a contradiction that there exist $h' \neq h_0$ such that

$$\Pr(\mathbf{y} = y|\mathbf{x} = x) = \mathbb{E}[\, h_0(y, \mathbf{x}, \mathbf{e})|\mathbf{x} = x; \theta_0 \,] = \mathbb{E}[\, h'(y, \mathbf{x}, \mathbf{e})|\mathbf{x} = x; \theta_0 \,]\,.$$

for all $y \in Y$ and $x \in X$. By Assumption 3 there exist $\tilde{h}_0$ and $\tilde{h}'$ such that

$$\mathbb{E}\left[\, \tilde{h}_0(y, \mathbf{w}, \mathbf{v}) - \tilde{h}'(y, \mathbf{w}, \mathbf{v})|\mathbf{x} = x; \theta_0 \,\right] = 0,$$

for all $y \in Y$ and $x \in X$, where $\mathbf{v}$ denotes the index from Assumption 3. The collection of the conditional distributions of $\mathbf{v}|\mathbf{x} = x$ forms a complete exponential family (see Theorem 2.12 in Brown, 1986). Hence,

$$\tilde{h}_0(y, \mathbf{w}, \mathbf{v}) - \tilde{h}'(y, \mathbf{w}, \mathbf{v}) = 0 \quad \text{a.s.}$$

for all $y \in Y$. This establishes identification of h_0. ■

The proof of Proposition 4.2 uses bounded completeness of the family of normal distributions. This property is not unique to normal distributions and is satisfied by many other parametric families. See Appendix A for more details.

4.3. Discernibility of Solution Concepts

Note that the definition of discernibility has the form "there do *not* exist (θ, h), $(\theta', h') \in \Psi$ such that…." Hence, the smaller Ψ is (the more restrictions are imposed), the easier it is to establish discernibility of a solution concept. In

[11] Although the rays themselves have probability zero, the partial derivatives use information in a neighborhood of those rays. Because the covariates are continuous, these open neighborhoods have positive probability and thus observable implications.

particular, if Ψ is a singleton, then every solution concept is trivially discernible. On the other hand, if Ψ is very big (very few restrictions are imposed), then one should not expect many solution concepts to be discernible. In this section, we provide two results. The first one establishes discernibility of solution concepts under a small number of restrictions (Theorem 4.3). The second one shows absence of discernibility despite Ψ being very small (Proposition 4.5).

Theorem 4.3 *Let S be any solution concept nested into S_R and let Ψ_R be the set of parameters that satisfy assumptions 1–4. S is discernible relative to Ψ_R.*

Proof. Assume towards a contradiction that there exists (θ, h) and (θ', h') that satisfy assumptions 1–4, and such that $h \in S(\theta)$, $h' \notin S(\theta')$, and

$$\mathbb{E}[\, h(\mathbf{x}, \mathbf{e})|\mathbf{x}; \theta \,] = \mathbb{E}[\, h'(\mathbf{x}, \mathbf{e})|\mathbf{x}; \theta' \,] \text{ a.s..}$$

Propositions 4.1 and 4.2 then imply that $h = h'$ and $\theta = \theta'$. The latter is not possible since by assumption $h' \notin S(\theta') = S(\theta)$. ■

Let us revisit the assumptions that define the set of admissible data generating processes captured by Ψ_R. Assumption 2 imposes standard restrictions on the payoff parameters. Assumptions 1 and 3 endow the econometrician access to excluded continuous covariates. These covariates play a crucial role. For example, Appendix C.4 in the online supplement shows that the solution concepts from Section 2 become discernible in the presence of such covariates. Assumption 4 allows any form of payoff-driven rationalizable behavior. Under these assumptions, Theorem 4.3 implies that we can discern any solution concept which implies rationalizable behavior, including S_N.

Corollary 4.4 *S_N is discernible relative to Ψ_R.*

Strengthening Assumptions 1–4 would not change the conclusion of Corollary 4.4, because it would only shrink the set Ψ_R. In contrast, the following proposition establishes the importance of our exclusion restrictions (Assumption 3).

Proposition 4.5 *Suppose Assumptions 1, 2, and 4 hold. Either $S_N(\theta_0) = S_R(\theta_0)$ or S_N is* not *discernible relative to $\Psi^* = \{\theta_0\} \times S_R(\theta_0)$.*

It is important to note that the payoff parameters are known under Ψ^*. That is, without Assumption 3, the NE solution concept is not discernible even in settings with a fully known or identifiable payoff structure (e.g., laboratory experiments).

4.4. Discriminating Between Solution Concepts

Discernibility is not sufficient to tell apart solution concepts with a non-empty intersection. Consider two solution concepts S and S'. If $h_0 \in S(\theta_0) \cap S'(\theta_0)$, then we cannot tell whether S or S' generated the data, because *both* S and S' generated the data. For example, whenever the players play a NE, they are also playing rationalizable strategies. Hence, S_R and S_N cannot be told apart in general even if S_N is discernible.

In order to guarantee that we can tell S and S' apart, we require them to be *distinct* in that $S(\theta_0) \cap S'(\theta_0) = \emptyset$. In other words, we require that S and S' make different predictions with positive probability. That is,

$$\Pr\left(h(\mathbf{x}, \mathbf{e}) \neq h'(\mathbf{x}, \mathbf{e})\right) > 0$$

for all $h \in S(\theta_0)$ and $h' \in S'(\theta_0)$. This is a mild restriction. For example, let $S_{R\setminus N}$ be the solution concept that results from assuming that behavior is rationalizable, but the players make choices that are not NE with positive probability. We have $S_{R\setminus N}(\theta_0) = S_R(\theta_0) \setminus S_N(\theta_0)$, which implies that $S_{R\setminus N}$ and S_N are distinct.

Without discernibility, it is possible to have distinct solution concepts that cannot be told apart. For example, the solution concepts from Section 2 are distinct and are observationally equivalent. The following corollary shows that, if the payoff parameters θ_0 are known or can be point identified, then discernibility is sufficient to tell distinct solution concepts apart. Section 5 shows that point identification of the payoff parameters is not necessary.

Corollary 4.6 *Consider any $\Psi \subseteq \{\theta_0\} \times H$, and any two solution concepts S and S' such that either S or S' is discernible relative to Ψ. If $S(\theta_0) \cap S'(\theta_0) = \emptyset$, then the data can be consistent with either S or S' but not both.*

Proof. Suppose without loss of generality that S is discernible, and take any $h \in S(\theta_0)$ and $h' \in S'(\theta_0)$. Since, $h' \notin S(\theta_0)$, it must be the case that (h, θ_0) and

(h', θ_0) imply different distributions over observables, i.e.,

$$\mathbb{E}[\,h(\mathbf{x},\mathbf{e})|\mathbf{x};\theta_0\,] \neq \mathbb{E}[\,h'(\mathbf{x},\mathbf{e})|\mathbf{x};\theta_0\,]\,\text{a.s.}.$$

Hence, at most one of S and S' can be consistent with the data. ■

5. Beyond Rationalizability and Identification of θ_0

The sole purpose of Assumption 4 is to establish identification of the payoff parameters. However, point identification of θ_0 is not necessary for discernibility of solution concepts. In this section, we consider two departures from rationalizability that lead to partial identification of the payoff parameters.

Suppose that firms are ambiguity averse in the sense of Gilboa and Schmeidler (1989). That is, suppose that each firm ranks its actions in terms of its minimum possible payoff, and then chooses the action that maximizes this minimum. This action is called the *maxmin* action, and is generically unique (indifference between actions is ruled out by continuity of the distribution of the utility shocks). Let $S_M(\theta_0)$ be the set of distributions of play that assign full probability to maxmin actions given the payoff parameter θ_0. The middle panel of Figure 3 illustrates the predictions of this solution concept for the two-firm entry-game from Example 1.

Another possibility is that the firms are colluding. Suppose that firms can compensate each other via transfers that are not observed by the researcher. In this case, the firms could agree to choose *collusive* outcomes that maximize the sum of their individual profits, even if doing so does not maximize the individual profits of some of them. Let $S_C(\theta_0)$ be the set of distributions of play that assign full probability to collusive outcomes. The predictions of the collusive solution concept are illustrated in the left panel of Figure 3 given the parametrization from Example 1.

Formally, we assume that the behavior of players is consistent with either rationalizability, maxmin, or collusive behavior.

Assumption 5 $h_0 \in \bar{S}(\theta_0) = S_R(\theta_0) \cup S_M(\theta_0) \cup S_C(\theta_0)$.

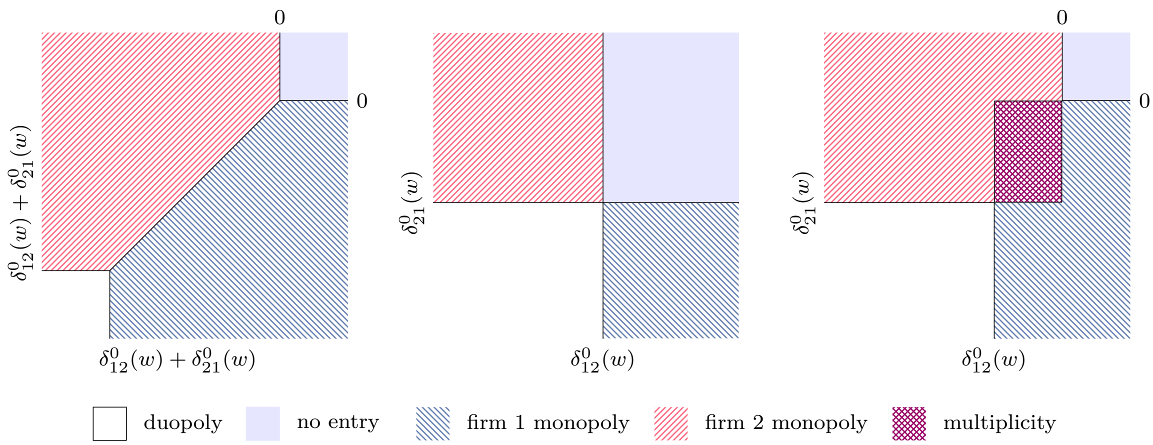

Figure 3 – Three solution concepts for two-firm entry games with $\delta_{12}(w), \delta_{21}(w) < 0$ and $\delta^0_{ii}(w) + \beta^0_i(w)z_i = 0$ for $i = 1, 2$: collusion (left), maxmin (middle), and rationalizability (right).

Proposition 5.1 *Suppose that assumptions 1–3, and 5 hold. Then*

(i) h_0, β_0, *and* Σ^0, *are identified.*

(ii) If $h_0 \in S_M(\theta_0) \cup S_C(\theta_0)$, *then* α_0 *is not point identified.*

(iii) Any solution concept nested into $\bar{S}$ *is discernible relative to the set of parameters that satisfy assumptions 1–3, and 5.*

(iv) One can discriminate between S_R, S_M, *and* S_C.

The proof of Proposition 5.1 is in Appendix A. Under collusive behavior, we can also identify some linear combinations of α^0 parameters. But, without imposing assumptions that would reduce the dimensionality, we cannot identify all of them. See Proposition A.1 in Appendix A to get a sense of which linear combinations can be identified.

We conclude this section by noting that some of our results establish discernibility of solution concepts and the ability to discriminate between them without point identification of either θ_0 or h_0. In particular Proposition 5.1 does not require point identification θ_0, and Proposition A.1 in the appendix establishes discernibility of S_R, S_M, and S_C without point identification of h_0.

6. Conclusion

We have defined discernibility of a given solution concept to mean that the solution concept can explain observed data if and only if the data was generated by it. And we have shown that any solution concept stronger than rationalizability is discernible under commonly imposed assumptions. We have also established identification of payoff parameters, including the correlation between unobserved payoff shocks, allowing for any form of rationalizable behavior. Our results are robust to some departures from rationalizability including ambiguity aversion and collusive behavior. Our exclusion restriction is necessary for discernibility of the NE solution concept in some settings, even when the payoff parameters are known.

It is possible to determine whether the data can be generated by any given convex solution concept (e.g., NE). For instance, one can construct the sets of conditional moment inequalities characterizing the solution concept (see Beresteanu et al., 2011 or Galichon and Henry, 2011). Then, the identified set (i.e., the set of parameters that satisfy these moment inequalities) is empty if and only if the data can be generated by the solution concept. Our results imply that one can substantially strengthen this conclusion. The identified set of payoff parameters is empty if and only if the data is in fact generated by the solution concept.

References

Andrews, D. W. K. (2011). Examples of L2-complete and boundedly-complete distributions. Discussion Paper 1801, Cowles Foundation.

Aradillas-López, A. and Tamer, E. (2008). The identification power of equilibrium in simple games. *Journal of Business & Economic Statistics*, 26(3):261–283.

Bajari, P., Hong, H., and Ryan, S. P. (2010). Identification and estimation of a discrete game of complete information. *Econometrica*, 78(5):1529–1568.

Beresteanu, A., Molchanov, I., and Molinari, F. (2011). Sharp identification regions in models with convex moment predictions. *Econometrica*, 79(6):1785–1821.

Bernheim, B. D. (1984). Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797.

Bjorn, P. A. and Vuong, Q. H. (1984). Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation. Social Science Working Paper 537, California Institute of Technology.

Blundell, R., Chen, X., and Kristensen, D. (2007). Semi-nonparametric IV estimation of shape-invariant Engel curves. *Econometrica*, 75(6):1613–1669.

Bresnahan, T. F. and Reiss, P. C. (1990). Entry in monopoly market. *The Review of Economic Studies*, 57(4):531–553.

Brown, L. D. (1986). *Fundamentals of statistical exponential families with applications in statistical decision theory*, volume 9 of *Lecture notes – Monograph series*. Institute of Mathematical Statistics.

Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: the case of penalty kicks in soccer. *American Economic Review*, 92(4):1138–1151.

Ciliberto, F., Murry, C., and Tamer, E. T. (2018). Market structure and competition in airline markets. SSRN 2777820.

Ciliberto, F. and Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.

Darolles, S., Fan, Y., Florens, J.-P., and Renault, E. (2011). Nonparametric instrumental regression. *Econometrica*, 79(5):1541–1565.

Dawid, R. (2016). Modelling non-empirical confirmation. In Ippoliti, E., Sterpetti, F., and Nickles, T., editors, *Models and Inferences in Science*, volume 25 of *Studies in Applied Philosophy, Epistemology and Rational Ethics*, pages 191–205. Springer.

De Paula, A. (2013). Econometric analysis of games with multiple equilibria. *Annual Review of Economics*, 5(1):107–131.

Galichon, A. and Henry, M. (2011). Set identification in models with multiple equilibria. *The Review of Economic Studies*, 78(4):1264–1298.

Gauriot, R., Page, L., and Wooders, J. (2016). Nash at Wimbledon: evidence from half a million serves. SSRN 2850919.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Goolsbee, A. (2002). Subsidies, the value of broadband, and the importance of fixed costs. In Crandall, R. and Alleman, J. H., editors, *Broadband: Should we Regulate High-Speed Internet Access?*, pages 278–294. Brookings Institution Press.

Green, E. J., Marshall, R. C., and Marx, L. M. (2014). Tacit collusion in oligopoly. In Blair, R. D. and Sokol, D. D., editors, *Oxford Handbook of International Antitrust*

Economics, volume 2, pages 464–497. Oxford University Press.

Hoderlein, S., Nesheim, L., and Simoni, A. (2012). Semiparametric estimation of random coefficients in structural economic models. Working Paper CWP09/12, CeMMAP.

Jia, P. (2008). What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica*, 76(6):1263–1316.

Jovanovic, B. (1989). Observable implications of models with multiple equilibria. *Econometrica*, 56(6):1431–1437.

Kline, B. (2015). Identification of complete information games. *Journal of Econometrics*, 189(1):117–131.

Manski, C. F. (1988). *Analog Estimation Methods in Econometrics*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman and Hall.

Mass, H. (2019). Strategies under strategic uncertainty. Mimeo.

Morris, S. and Shin, H. S. (2003). Global games: theory and applications. In Dewatripont, M., Hansen, L. P., and Turnovsky, S. J., editors, *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 1*, volume 34 of *Econometric Society Monographs*, pages 56–114. Cambridge University Press.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, 52(2):286–295.

Newey, W. K. and Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050.

Rubinstein, A. (1989). The electronic mail game: strategic behavior under almost common knowledge. *American Economic Review*, 79(3):385–391.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165.

Walker, M. and Wooders, J. (2001). Minimax play at Wimbledon. *American Economic Review*, 91(5):1521–1538.

A. Omitted Proofs

In this section, we first prove a version of Propositions 4.1 for games with two players and without nonexcluded covariates $\mathbf{w}$, where we relax the normality and rationalizability assumptions. This result shows that one does not need to point identify θ_0 and h_0 in order to establish the discernibility of some solution concepts. Then, we show how this results can be applied to games with covariates, with many players, and with many actions (Propositions 4.1, 5.1, and C.1).

A.1. Binary Two-Player Game

Suppose that $I = \{1, 2\}$, $Y_i = \{0, 1\}$ for $i \in I$, and i's payoffs are given by

$$y_i \cdot (\alpha^0_{i,y_{-i}} + \beta^0_i \mathbf{z}_i - \mathbf{e}_i).$$

Assumption 6

(i) $\mathbf{z} = (\mathbf{z}_i)_{i=1,2}$ and $\mathbf{e} = (\mathbf{e}_i)_{i=1,2}$ are independent.

(ii) $\mathbf{e}$ admits a probability density function (p.d.f.) $f_{\mathbf{e}}$ that is continuously differentiable and strictly positive on $\mathbb{R}^2$.

(iii) $\mathbb{E}[\,\mathbf{e}_i\,] = 0$ and $\mathbb{E}[\,\mathbf{e}_i^2\,] = 1$, $i = 1, 2$.

Part (ii) of Assumption 6 is a regularity condition needed for invertibility of the marginal cumulative distribution functions (c.d.f.) $F_{\mathbf{e}_i}$, $i = 1, 2$. Part (iii) is a location and scale normalization. Assumption 6 is implied by Assumption 2.

Assumption 7 $\alpha^0_{1,1} - \alpha^0_{1,0} \neq \alpha^0_{2,1} - \alpha^0_{2,0}$.

Assumption 7 requires the strategic effects of players' actions to be asymmetric. Geometrically, it means that the multiplicity region has different height and width.

The next proposition shows that rationalizability, collusion, and maxmin are discernible under minimal restrictions on the distribution of shocks. It also illustrates that point identification of neither θ_0 nor h_0 is necessary for solution concepts to be discernible and for the ability to discriminate between solution concepts.

Proposition A.1 *Suppose that assumptions 1 and 5–7 hold. Then*

(i) β^0_i *and* $F_{\mathbf{e}_i}$ *are point identified for all* $i = 1, 2$.

(ii) If $h_0 \in S_R(\theta_0)$, *then* α^0_{i,y_i} *is identified for all* i *and* y_i.

(iii) *If* $h_0 \in S_M(\theta_0)$*, then only* $\min_{y_i}\{\alpha^0_{i,y_i}\}$ *is identified for all* i*.*

(iv) *If* $h_0 \in S_C(\theta_0)$*, then only* $\alpha^0_{i,0}$*,* $i = 1, 2$*, and* $\alpha^0_{1,1} + \alpha^0_{2,1}$ *are identified.*

(v) S_R*,* S_C*, and* S_M *are discernible relative to the set of parameters that satisfy assumptions 1 and 5–7.*

(vi) *One can discriminate between* S_R*,* S_M*, and* S_C*.*

Proof. (*Step 1—Identification of* β^0_i *and* $F_{\mathbf{e}_i}$) Define $\delta^0_i = \alpha^0_{i,1} - \alpha^0_{i,0}$, $i = 1, 2$. Note that under rationalizability and collusive behavior

$$\Pr(\mathbf{y}_1 = 0, \mathbf{y}_2 = 0 | \mathbf{z} = z) = \int_{\beta^0_1 z_1}^{\infty} \int_{\beta^0_2 z_2}^{\infty} f_{\mathbf{e}}(e)\, de + q(z),$$

where under collusive behavior

$$q(z) = \begin{cases} 0, & \delta^0_1 + \delta^0_2 \leq 0, \\ -\int_{\beta^0_1 z_1 + \delta^0_1 + \delta^0_2}^{\beta^0_1 z_1} \int_{\beta^0_2 z_2 + \delta^0_1 + \delta^0_2}^{\beta^0_2 z_2} h(\beta^0_1 z_1 - e_1, \beta^0_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de, & \delta^0_1 + \delta^0_2 > 0, \end{cases}$$

and under rationalizability

$$q(z) = \begin{cases} \int_{\beta^0_1 z_1 + \delta^0_1}^{\beta^0_1 z_1} \int_{\beta^0_2 z_2 + \delta^0_2}^{\beta^0_2 z_2} h_0(\beta^0_1 z_1 - e_1, \beta^0_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de, & \delta^0_1, \delta^0_2 \leq 0, \\ \int_{\beta^0_1 z_1}^{\beta^0_1 z_1 + \delta^0_1} \int_{\beta_2 z_2 + \delta^0_2}^{\beta^0_2 z_2} h_0(\beta^0_1 z_1 - e_1, \beta^0_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de, & \delta^0_1 > 0, \delta^0_2 \leq 0, \\ \int_{\beta^0_1 z_1 + \delta^0_1}^{\beta^0_1 z_1} \int_{\beta^0_2 z_2}^{\beta^0_2 z_2 + \delta^0_2} h_0(\beta^0_1 z_1 - e_1, \beta^0_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de, & \delta^0_1 \leq 0, \delta^0_2 > 0, \\ -\int_{\beta^0_1 z_1}^{\beta^0_1 z_1 + \delta^0_1} \int_{\beta^0_2 z_2}^{\beta^0_2 z_2 + \delta^0_2} h_0(\beta^0_1 z_1 - e_1, \beta^0_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de, & \delta^0_1, \delta^0_2 > 0. \end{cases}$$

Under maxmin

$$\Pr(\mathbf{y}_1 = 0, \mathbf{y}_2 = 0 | \mathbf{z} = z) = \int_{\beta^0_1 z_1 + \min\{\delta^0_1, 0\}}^{\infty} \int_{\beta^0_2 z_2 + \min\{\delta^0_2, 0\}}^{\infty} f_{\mathbf{e}}(e)\, de.$$

Hence, Lemma B.1 can be applied to $p(z) = \Pr(\mathbf{y}_1 = 0, \mathbf{y}_2 = 0 | \mathbf{z} = z)$ under all three solution concepts. We can thus identify β^0_i and the marginal c.d.f.s $F_{\mathbf{e}_i}$, $i = 1, 2$, independently of the solution concept. Since z_i, $i = 1, 2$, can be rescaled, assume without loss of generality that $\beta^0_i = 1$, $i = 1, 2$.

(*Step 2—Identification of* $\alpha^0_{i,y_{-i}}$) Next, let

$$\mu_i(z) = \Pr(\mathbf{y}_i = 1 | \mathbf{z} = z)$$

denote the (known) probability of firm i entering, conditional on the value of $\mathbf{z}$. We will take limits as $\mathbf{z}_{-i}$ goes to $\pm\infty$. Under any of the three solution concepts,

$$\lim_{z_{-i}\to+\infty} \mu_{-i}(z) = 1 \quad \text{and} \quad \lim_{z_{-i}\to-\infty} \mu_{-i}(z) = 0.$$

Hence, under rationalizability,

$$\lim_{z_{-i}\to-\infty} \mu_i(z) = F_{\mathbf{e}_i}(\alpha^0_{i,0} + z_i) \implies \alpha^0_{i,0} = F^{-1}_{\mathbf{e}_i}\Big(\lim_{z_{-i}\to-\infty} \mu_i(z)\Big) - z_i,$$
$$\lim_{z_{-i}\to+\infty} \mu_i(z) = F_{\mathbf{e}_i}(\alpha^0_{i,1} + z_i) \implies \alpha^0_{i,1} = F^{-1}_{\mathbf{e}_i}\Big(\lim_{z_{-i}\to+\infty} \mu_i(z)\Big) - z_i,$$

where we used the facts that $\lim_{z_{-i}\to+\infty} \mu_i(z)$ and $\lim_{z_{-i}\to-\infty} \mu_i(z)$ are well defined, and $F_{\mathbf{e}_i}$ is known (from Step 1) and invertible (from part (ii) of Assumption 6). Similarly, under collusive behavior,

$$\lim_{z_{-i}\to-\infty} \mu_i(z) = F_{\mathbf{e}_i}(\alpha^0_{i,0} + z_i) \implies \alpha^0_{i,0} = F^{-1}_{\mathbf{e}_i}\Big(\lim_{z_{-i}\to-\infty} \mu_i(z)\Big) - z_i,$$

and

$$\lim_{z_{-i}\to+\infty} \mu_i(z) = F_{\mathbf{e}_i}(\alpha^0_{1,1} + \alpha^0_{2,1} - \alpha^0_{-i,0} + z_i)$$
$$\implies \alpha^0_{1,1} + \alpha^0_{2,1} = F^{-1}_{\mathbf{e}_i}\Big(\lim_{z_{-i}\to+\infty} \mu_i(z)\Big) - z_i + \alpha^0_{-i,0}.$$

Finally, under maxmin, each player i chooses between a payoff of 0 and

$$\min\Big\{\alpha^0_{i,0} + \mathbf{z}_i - \mathbf{e}_i,\ \alpha^0_{i,1} + \mathbf{z}_i - \mathbf{e}_i\Big\} = \mathbf{z}_i - \mathbf{e}_i + \min\{\alpha^0_{i,0},\ \alpha^0_{i,1}\}.$$

Hence, under maxmin,

$$\mu_i(z) = F_{\mathbf{e}_i}\Big(\min\{\alpha^0_{i,0}, \alpha^0_{i,1}\} + z_i\Big) \implies \min\{\alpha^0_{i,0}, \alpha^0_{i,1}\} = F^{-1}_{\mathbf{e}_i}\Big(\mu_i(z)\Big) - z_i.$$

(*Step 3—Discriminating Between Solution Concepts*) Recall that we have defined $\delta^0_i = \alpha^0_{i,1} - \alpha^0_{i,0}$. We can discriminate between rationalizability, collusive behavior, and maxmin by examining the statistics

$$t_i = F^{-1}_{\mathbf{e}_i}\left(\lim_{z_{-i}\to+\infty} \mu_i(z)\right) - F^{-1}_{\mathbf{e}_i}\left(\lim_{z_{-i}\to-\infty} \mu_i(z)\right),$$

$i = 1, 2$. Under maxmin, we have $t_1 = t_2 = 0$. Under collusion, we have $t_1 = t_2 =$

$\delta_1^0 + \delta_2^0$. Under rationalizability, we have $t_1 = \delta_1^0$ and $t_2 = \delta_2^0$.

Assumption 7 implies that $\delta_1^0 \neq \delta_2^0$. Hence, the observed data can be used to discriminate between S_R and S_C, and between S_R and S_M. If $\delta_1^0 + \delta_2^0 \neq 0$, then S_M and S_C can also be discriminated. Otherwise, $S_C(\theta) = S_M(\theta)$ for all θ and hence S_M and S_C are discernible. ■

Proposition A.1 still does not identify the correlation structure nor the distribution of play. For that purpose, we impose the following assumptions on the distribution of the unobserved payoff shocks. Let $f_{\mathbf{e}_i}$, and $F_{\mathbf{e}_i|\mathbf{e}_j}$ denote marginal p.d.f. of $\mathbf{e}_i$ and the conditional c.d.f. of $\mathbf{e}_i$ conditional on $\mathbf{e}_j$, respectively. We use ∂_{e_i} to denote the partial derivatives with respect to e_i.

Assumption 8

(i) For all $\tau \in (-1,1)$ and all $\bar{e}, \underline{e} \in \mathbb{R}$ such that $\bar{e} \geq \underline{e}$, there exists a constant $e^* \in \mathbb{R}$ such that

$$\Big[e_1 < e^* \text{ and } \tau e_1 + \underline{e} \leq e_2 \leq \tau e_1 + \bar{e}\Big] \implies \partial_{e_1} f_{\mathbf{e}}(e) \geq 0.$$

(ii) For almost all (with respect to the Lebesgue measure) $\tau \in (-1,1)$, and all real numbers $\gamma_1, \bar{\gamma}_2, \underline{\gamma}_1 \in \mathbb{R}$ such that $\bar{\gamma}_2 \geq \underline{\gamma}_2$,

$$\lim_{z_1 \to -\infty} \frac{f_{\mathbf{e}_1}(z_1+\gamma_1)}{f_{\mathbf{e}_1}(z_1)} \Big[F_{\mathbf{e}_2|\mathbf{e}_1}(\tau z_1 + \bar{\gamma}_2 | z_1 + \gamma_1) - F_{\mathbf{e}_2|\mathbf{e}_1}(\tau z_1 + \underline{\gamma}_2 | z_1 + \gamma_1)\Big] = 0.$$

Condition (i) in Assumption 8 requires the tail of the joint density to be convex in e_i along directions $e_{-i} = \tau e_i$. Condition (ii) controls the rates of convergence to zero of the marginal p.d.f. and the conditional c.d.f.. Both conditions are satisfied by distributions that have exponential tails.

Example 3 Assumption 8 is satisfied for the bivariate normal distribution. Indeed, for the bivariate distribution with unit variances and with correlation ρ_0

$$\partial_{e_1} f_{\mathbf{e}}(e_1, e_2) = \frac{\rho_0 e_2 - e_1}{1-\rho_0^2} \cdot f_{\mathbf{e}}(e_1,e_2) \geq \frac{(\rho_0\tau - 1)e_1 + \min\{\rho_0\underline{e}, \rho_0\bar{e}\}}{1-\rho_0^2} \cdot f_{\mathbf{e}}(e_1,e_2).$$

Hence, one can take $e^* = \min\{\rho_0\underline{e}, \rho_0\bar{e}\}/(1-\tau\rho_0)$. In order to verify condition (ii), note that for the normal distribution we have

$$\frac{f_{\mathbf{e}_i}(z_i+\gamma_i)}{f_{\mathbf{e}_i}(z_i)} = \exp(-\gamma_i^2/2)\cdot\exp(\gamma_i z_i),$$

and

$$\begin{aligned}
&F_{\mathbf{e}_{-i}|\mathbf{e}_i}(\tau z_i+\bar{\gamma}_{-i}|z_i+\gamma_i) - F_{\mathbf{e}_{-i}|\mathbf{e}_i}(\tau z_i+\underline{\gamma}_{-i}|z_i+\gamma_i)\\
&\qquad = \Phi\left(\frac{(\tau-\rho_0)e_i+\bar{\gamma}_{-i}-\rho_0\gamma_i}{\sqrt{1-\rho_0^2}}\right) - \Phi\left(\frac{(\tau-\rho_0)e_i+\underline{\gamma}_{-i}-\rho_0\gamma_i}{\sqrt{1-\rho_0^2}}\right),
\end{aligned}$$

where $\Phi(\,\cdot\,)$ denotes the standard normal c.d.f.. Thus, for $\gamma_i>0$, condition (iii) follows trivially. Now, suppose $\gamma_i<0$. If $\tau\neq\rho_0$, then L'Hôpital's yields:

$$\begin{aligned}
&\lim_{e_1\to-\infty}\frac{f_{\mathbf{e}_i}(z_i+\gamma_i)}{f_{\mathbf{e}_i}(z_i)}\Big[F_{\mathbf{e}_{-i}|\mathbf{e}_i}(\tau z_i+\bar{\gamma}_{-i}|z_i+\gamma_i) - F_{\mathbf{e}_{-i}|\mathbf{e}_i}(\tau z_i+\underline{\gamma}_{-i}|z_i+\gamma_i)\Big]\\
&= \lim_{e_1\to-\infty}\frac{(\tau-\rho_0)\left[\phi\left(\frac{(\tau-\rho_0)e_i+\bar{\gamma}_{-i}-\rho_0\gamma_i}{\sqrt{1-\rho_0^2}}\right) - \phi\left(\frac{(\tau-\rho_0)e_i+\underline{\gamma}_{-i}-\rho_0\gamma_i}{\sqrt{1-\rho_0^2}}\right)\right]}{-\gamma_i\exp(-\gamma_1^2/2)\cdot\exp(-\gamma_i e_i)}\\
&= \lim_{e_1\to-\infty}\frac{\exp\left(-\frac{(\tau-\rho_0)^2}{2(1-\rho_0^2)}e_i^2\right)}{\exp(-\gamma_i e_i)},
\end{aligned}$$

where $\phi(\,\cdot\,)$ denotes the standard normal p.d.f.. The latter limit equals to zero because $\tau\neq\rho_0$.

We also need to model the correlation between $\mathbf{e}_1$ and $\mathbf{e}_2$.

Assumption 9 $\mathbf{e}_2=\sqrt{1-\rho_0^2}\,\boldsymbol{\xi}+\rho_0\mathbf{e}_1$ a.s., where $\boldsymbol{\xi}$ is independent from $\mathbf{e}_1$ and $\rho_0\in(-1,1)$ is an unknown parameter.

Assumptions 8 and 9 allow us to identify the correlation between unobservables. Note that the roles of $\mathbf{e}_1$ and $\mathbf{e}_2$ are fully interchangeable in Assumptions 8 and 9 (e.g., one could define $\mathbf{e}_1=\sqrt{1-\rho_0^2}\,\boldsymbol{\xi}+\rho_0\mathbf{e}_2$ a.s.). In order to identify h_0, we use the following assumption.

Assumption 10 The family of distributions $\{F_{\mathbf{e}}(e-t)\mid t\in\mathbb{R}^2\}$ is boundedly complete. That is, for any bounded function $g : \mathbb{R}^2 \to \mathbb{R}$

$$\left[\forall t \in \mathbb{R}^2,\ \int_{-\infty}^{\infty} g(e) f_{\mathbf{e}}(e-t)\, de = 0\right] \implies g(\mathbf{e}) = 0 \text{ a.s..}$$

Assumption 10 is a richness condition and is satisfied by many multivariate distributions. For example, it is satisfied by the multivariate normal (see, for instance, Newey and Powell, 2003) and the Gumbel distributions.

Example 4 Assume that $\mathbf{e}_1$ and $\boldsymbol{\xi}$ are i.i.d. according to the Gumbell distribution with parameter $(0,1)$, that is, the p.d.f. of $\mathbf{e}_1$ is $f(e_1) = \exp(-e_1 - \exp(-e_1))$. The family of distributions $\{f_{\mathbf{e}}(e-t) \mid t \in \mathbb{R}^2\}$ is boundedly complete since it belongs to the two-parameter exponential family with t as a parameter (Brown, 1986). Indeed, letting $\gamma_0 := \sqrt{1-\rho_0^2}$,

$$\begin{aligned} f_{\mathbf{e}}(e-t) &= f_{\mathbf{e}_2|\mathbf{e}_1}(e_2 - t_2 | e_1 - t_1) f_{\mathbf{e}_1}(e_1 - t_1) \\ &= \frac{1}{\gamma_0} \cdot f(e_1 - t_1) f\left(\frac{e_2 - t_2 - \rho_0(e_1 - t_1)}{\gamma_0}\right) \\ &= G(e) \exp\left(\sum_{i=1}^{2} \eta_i(t) T_i(e) + \xi(t)\right), \end{aligned}$$

where

$$G(e) = \frac{1}{\gamma_0} \cdot \exp\left(\frac{(\rho_0 - \gamma_0)\, e_1 - e_2}{\gamma_0}\right), \quad \xi(t) = \frac{1}{\gamma_0} \cdot \left((\gamma_0 - \rho_0)\, t_1 + t_2\right),$$

$$T_1(e) = \exp(-e_1), \qquad T_2(e) = \exp\left(-\frac{e_2 - \rho_0 e_1}{\gamma_0}\right),$$

$$\eta_1(t) = \exp(t_1) \quad \text{and} \quad \eta_2(t) = \exp\left(\frac{t_2 - \rho_0 t_1}{\gamma_0}\right).$$

The following proposition establishes point identification of h_0 and the correlation parameter ρ_0.

Proposition A.2 *Suppose that assumptions 1, 3, 5, 6, and 8–10 hold. Then*

(i) All the conclusions of Proposition A.1 hold;

(ii) ρ_0 and h_0 are identified;

(iii) Any solution concept nested into $\bar{S}$ is discernible relative to the set of parameter values that satisfy assumptions 1, 3, 5, 6, and 8–10.

Proof. Validity of the conclusions (i)–(iv) of Proposition A.1 is trivial, because they do no rely on Assumption 7, and all the other assumption from Proposition A.1 are satisfied. Identification of ρ_0 follows from combining Step 1 of the proof of Proposition A.1 with Lemma B.1. Identification of h_0 and discernibility of any S nested into $\bar{S}$ follows from the same arguments used in the proofs of Proposition 4.2 and Theorem 4.3 in the main text. ■

Note that, unlike Proposition A.1, Proposition A.2 does not use Assumption 7 to discriminate between rationalizability, collusion, and maxmin. The main difference is that under the assumptions of Proposition A.1, h_0 may not be point identified, and the only way to discriminate between solution concepts is to use asymmetry of the multiplicity region.

A.2. Proof of Proposition 4.1

Fix some w. For notation simplicity we will drop w from the notation. Since conditional on $\mathbf{w} = w$ the support of $\mathbf{z}$ is full and $\mathbf{z}_i$ enters only payoffs of player i, we can identify the sign of β_i for every i since

$$\lim_{z_i \to -\infty} \Pr\left(\mathbf{y}_i = 0 | \mathbf{z} = z\right) = 0 \iff \beta_{i,y_i} > 0$$

for all i. Without loss of generality we assume that $\beta_i > 0$ for all i (if $\beta_i < 0$ we can always use $-z_i$ as a covariate). Next we pick any two different players i and j, and a profile of actions for all other players $\{y_k\}_{k\in I\setminus\{i,j\}}$. By sending z_k, $k \in I \setminus \{i, j\}$ either to $+\infty$ or to $-\infty$ we can guarantee that

$$\lim_{z_k \to C_k, k \in I\setminus\{i,j\}} \Pr(\mathbf{y}_k = y_k,\ k \in I \setminus \{i, j\} | \mathbf{z} = z) = 1,$$

where $C_k = +\infty$ if $y_k = 1$ and $C_k = -\infty$ if $y_k = 0$. Thus we end up having a two player game. Applying Proposition A.2 we identify β_i, β_j, $\alpha_{i,y_{-i}}$, $\alpha_{j,y_{-j}}$, and Σ_{ij}. The conclusion of the proposition follows from the fact that the choice of w, i, j, and $\{y_k\}_{k\in I\setminus\{i,j\}}$ was arbitrary. ■

A.3. Proof of Proposition 4.5

Suppose that $S_R(\theta_0) \neq S_N(\theta_0)$. We will construct two rationalizable distributions of play $h, h' \in S_R(\theta_0)$ such $h \in S_N(\theta_0)$, $h' \notin S_N(\theta_0)$, and

$$\mathbb{E}[\, h(\mathbf{x}, \mathbf{e}) | \mathbf{x}; \theta_0 \,] = \mathbb{E}[\, h'(\mathbf{x}, \mathbf{e}) | \mathbf{x}; \theta_0 \,] \text{ a.s..}$$

Since $F_{\mathbf{e}}$ is absolutely continuous, players are almost surely not indifferent between outcomes. Hence, the fact that $S_R(\theta_0) \neq S_N(\theta_0)$ implies that the payoffs of at least two players must depend on their opponents' actions. That is, there must exist $i, j \in I$ and a profile $y_{-ij} \in \{0, 1\}^{d_I - 2}$ such that

$$\alpha^0_{i,(1,y_{-ij})}(\mathbf{w}) \neq \alpha^0_{i,(0,y_{-ij})}(\mathbf{w}) \quad \text{and} \quad \alpha^0_{j,(1,y_{-ij})}(\mathbf{w}) \neq \alpha^0_{j,(0,y_{-ij})}(\mathbf{w})$$

with positive probability.

Since β_0 is known, we can make each player $k \notin \{i, j\}$ play the action specified in y_{-ij} with probability 1 by taking limits as z_k goes to either $+\infty$ or $-\infty$. Taking such limits for all $k \notin \{i, j\}$ is as if players i and j were playing a two-player game. Hence, we can assume without loss of generality that $d_I = 2$, and $i = 1$ and $j = 2$ are the only two players. Moreover, for exposition purposes, we will drop w from the notation.

With these simplifications, the multiplicity region is characterized by

$$E(z) = \left\{ e \in \mathbb{R}^{d_I} \;\middle|\; \min\{\alpha^0_{i,1}, \alpha^0_{i,0}\} \leq e_i - \beta^0_i z_i \leq \max\{\alpha^0_{i,1}, \alpha^0_{i,0}\} \text{ for } i = 1, 2 \right\}.$$

For $e \in E(z)$, the best response of each player depends on the action of its opponent. For example, if $\alpha^0_{1,0} > \alpha^0_{1,1}$, then player 1 prefers $y_1 = 1$ to $y_1 = 0$ if $y_2 = 0$, and prefers $y_1 = 0$ if $y_2 = 1$. Moreover, for $e \in E(z)$, the game has either zero or two PNEs, and one mixed-strategy NE. In the mixed-strategy NE, each firm i, $i = 1, 2$, chooses $y_i = 1$ with probability

$$\frac{a^0_{-i,0} + \beta^0_{-i} z_{-i} - e_{-i}}{a^0_{-i,0} - a^0_{-i,1}} \in (0, 1).$$

For $e \notin E(z)$, there is a unique NE almost everywhere.

Let h be such that the players play the unique NE when $e \notin E(z)$, and play the mixed-strategy NE when $e \in E(z)$. Let h' be given by $h'(z, e) = h(z, e)$ when

$e \notin E(z)$, and

$$h'(y,z,e) = \int_{E(z)} h(y,z,\epsilon) f_{\mathbf{e}}(\epsilon)\, d\epsilon$$

when $e \in E(z)$. By construction, $h \in S_N(\theta_0)$ and

$$\mathbb{E}[\, h(\mathbf{x},\mathbf{e})|\mathbf{x};\theta_0\,] = \mathbb{E}[\, h'(\mathbf{x},\mathbf{e})|\mathbf{x};\theta_0\,] \text{ a.s..}$$

Hence, it only remains to show that h' does not belong to $S_N(\theta_0)$.

In the multiplicity region, the game has less than four PNEs. Moreover, these PNEs are the same for all pairs (z,e) such that $e \in E(z)$, because they only depend on the sign of $\alpha^0_{i,0} - \alpha^0_{i,1}$, $i = 1,2$. Consequently, there exists an outcome y^* that is not played in any PNE of the multiplicity region. Therefore, for $e \in E(z)$, $h(y^*,z,e)$ is the maximum probability of y^* consistent with $S_N(\theta_0)$. By construction, for every $z \in Z$, there exists a set with positive Lebesgue measure $\tilde{E}(z) \subseteq E(z)$ such that $h'(y^*,z,e) > h(y^*,z,e)$ for all $e \in \tilde{E}(z)$. Therefore, $h' \notin S_N(\theta_0)$, and we can conclude that S_N is not discernible. ■

Note that the distribution of play h' in the proof of Proposition 4.5 does *not* satisfy our exclusion restriction. Assumption 3 requires that, conditional on $\mathbf{w}$, the joint distribution of distribution over outcomes conditional on the payoff indices does not depend on $\mathbf{z}$. In contrast, h' allows this distribution to be different for every $z \in Z$.

A.4. Proof of Proposition 5.1

Fix any $w \in W$ and players $i,j \in I$, $i \neq j$. As in the proof of Proposition 4.1, we can turn the many-player game into a two-player game for i and j and any profile $\{y_k\}_{k \in I \setminus \{i,j\}}$. Applying Proposition A.2, we can identify $\beta^0_i(w)$, $\beta^0_j(w)$, and $\Sigma^0_{ij}(w)$. Since the w, i, and j are arbitrary, we can identify β_0 and Σ^0. Assumption 3 together with normality of $\mathbf{e}$ identifies h_0. Hence, any S nested into $\bar{S}$ is discernible. The lack of identification of α_0 under S_M and S_C follows from parts (iii) and (iv) of Proposition A.1. ■

B. Auxiliary Results

The following lemma identifies the marginal effect of z_i (captured by β_i^0), the nonparametric marginal distributions of error terms ($F_{\mathbf{e}_i}$), and the correlation between $\mathbf{e}_1$ and $\mathbf{e}_2$ in binary games.

Lemma B.1 *Let $f_{\mathbf{e}}$ be the p.d.f. of $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$ and let $p : \mathbb{R}^2 \to \mathbb{R}$ be given by*

$$p(z) = \int_{\beta_1 z_1 + \delta_1}^{\infty} \int_{\beta_2 z_2 + \delta_2}^{\infty} f_{\mathbf{e}}(e)\, de_2\, de_1 + q(z),$$

with

$$|q(z)| = \int_{\beta_1 z_1 + \underline{\gamma}_1}^{\beta_1 z_1 + \bar{\gamma}_1} \int_{\beta_2 z_2 + \underline{\gamma}_2}^{\beta_2 z_2 + \bar{\gamma}_2} g(\beta_1 z_1 - e_1, \beta_2 z_2 - e_2) f_{\mathbf{e}}(e)\, de_2\, de_1,$$

for some unknown parameters $\beta_i \neq 0$, $\delta_i, \bar{\gamma}_i, \underline{\gamma}_i \in \mathbb{R}$ with $\bar{\gamma}_i \geq \underline{\gamma}_i$, $i = 1, 2$, and $g : \mathbb{R}^2 \to [0, 1]$. If (i) $\mathbb{E}[\mathbf{e}_i] = 0$ and $\mathbb{E}[\mathbf{e}_i^2] = 1$, $i = 1, 2$; and (ii) $f_{\mathbf{e}}$ is continuously differentiable and strictly positive on $\mathbb{R}^2$; then, each of β_i, δ_i, and the marginal c.d.f.s $F_{\mathbf{e}_i}$, $i = 1, 2$, are identified from knowing the function p. If, moreover, (iii) $\mathbf{e}_2 = \sqrt{1 - \rho^2}\boldsymbol{\xi} + \rho \mathbf{e}_1$ a.s., where $\boldsymbol{\xi}$ is independent from $\mathbf{e}_1$ and $\rho \in (-1, 1)$; and (iv) $f_{\mathbf{e}}$ satisfies the conditions from Assumption 8; then the correlation parameter ρ is also identified from p.

Proof. First note that

$$\lim_{|z_i| \to +\infty} |q(z)| \leq \lim_{|z_i| \to +\infty} F_{\mathbf{e}_i}(\beta_i z_i + \bar{\gamma}_i) - F_{\mathbf{e}_i}(\beta_i z_i + \underline{\gamma}_i) = 0,$$

for all i. Hence, for all i,

$$\lim_{z_{-i} \to +\infty} p(z) = \mathbb{1}(\beta_{-i} < 0)[1 - F_{\mathbf{e}_i}(\beta_i z_i + \delta_i)]$$

and

$$\lim_{z_{-i} \to -\infty} p(z) = \mathbb{1}(\beta_{-i} > 0)[1 - F_{\mathbf{e}_i}(\beta_i z_i + \delta_i)].$$

Thus, since $f_{\mathbf{e}}$ is strictly positive on $\mathbb{R}^2$, we can identify the sign of β_i and $p_i(z_i) =: F_{\mathbf{e}_i}(\beta_i z_i + \delta_i)$, $i = 1, 2$. Since $\beta_i \neq 0$ and we know the mean and the variance of

$\mathbf{e}_i$, for $i = 1, 2$, we can also identify

$$\int_{-\infty}^{\infty} t\, dp_i(t) = \frac{1}{\beta_i} \cdot \mathbb{E}[\, \mathbf{e}_i - \delta_i \,] = -\frac{\delta_i}{\beta_i},$$

and

$$\int_{-\infty}^{\infty} t^2\, dp_i(t) = \frac{1}{\beta_i^2} \cdot \mathbb{E}\Big[\, (\mathbf{e}_i - \delta_i)^2 \,\Big] = \frac{1 + \delta_i^2}{\beta_i^2},$$

where we used the change of variables $e_i = \beta_i t + \delta_i$. As a result, since we already learned the sign of β_i, we can identify β_i and δ_i from

$$\beta_i^2 = \left[\int_{-\infty}^{\infty} t^2\, dp_i(t) - \left(\int_{-\infty}^{\infty} t\, dp_i(t) \right)^2 \right]^{-1} \quad \text{and} \quad \delta_i = -\beta_i \int_{-\infty}^{\infty} t\, dp_i(t).$$

If $q(z) = 0$ for all z, then we identify the joint distribution of $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$ since β_i and δ_i, $i = 1, 2$, are identified and

$$\Pr(\mathbf{e}_1 \geq t_1, \mathbf{e}_2 \geq t_2) = p\left(\frac{t_1 - \delta_1}{\beta_1}, \frac{t_2 - \delta_2}{\beta_2} \right)$$

for all $t_1, t_2 \in \mathbb{R}$. Note that, in this case, we do not need to invoke Lemma B.2. If $q(z) \neq 0$ for some z, we can still identify the marginal distribution of the error terms given that

$$F_{\mathbf{e}_i}(t_i) = p_i\left(\frac{t_i - \delta_i}{\beta_i} \right).$$

It only remains to identify ρ in the case $q(z) \neq 0$ for some z. We can always rescale and shift z_i, $i = 1, 2$. Hence, we can assume without loss of generality that $\beta_i = 1$ and $\delta_i = 0$, $i = 1, 2$. Then, it follows from Lemma B.2 and the independence between $\boldsymbol{\xi}$ and $\mathbf{e}_1$ that

$$1 + \lim_{z_1 \to -\infty} \frac{\partial_{z_1} p(z)}{f_{\mathbf{e}_1}(z_1)} \Bigg|_{z_2 = \tau z_1} = \lim_{z_1 \to -\infty} F_{\mathbf{e}_2 | \mathbf{e}_1}(\tau z_1 | z_1) = \lim_{z_1 \to -\infty} F_{\boldsymbol{\xi}}\left(\frac{(\tau - \rho) z_1}{\sqrt{1 - \rho^2}} \right).$$

for almost all $\tau \in [-1, 1]$. Hence, ρ is the only point of jump discontinuity of the function $\psi : [-1, 1] \to \mathbb{R}$ given by

$$\psi(\tau) = 1 + \lim_{z_1 \to -\infty} \frac{\partial_{z_1} p(z)}{f_{\mathbf{e}_1}(z_1)} \Bigg|_{z_2 = \tau z_1}.$$

Hence, ρ is also identified. ■

Lemma B.2 *Let $\hat{p} : Z \to \mathbb{R}$ be given by*

$$\hat{p}(z) = \int_{z_1}^{\infty} \int_{z_2}^{\infty} f_{\mathbf{e}}(e)\, de_2\, de_1 + \hat{q}(z),$$

with

$$|\hat{q}(z)| = \int_{z_1+\underline{\gamma}_1}^{z_1+\bar{\gamma}_1} \int_{z_2+\underline{\gamma}_2}^{z_2+\bar{\gamma}_2} g(z_1 - e_1, z_2 - e_2) f_{\mathbf{e}}(e)\, de_2\, de_1,$$

where $\gamma, \tilde{\gamma}$ are such that $\gamma_i \leq \tilde{\gamma}_i$, $i = 1, 2$; $f_{\mathbf{e}}$ is a bivariate p.d.f. satisfying Assumption 8; and $g : \mathbb{R}^2 \to [0, 1]$ is an arbitrary function. Then,

$$1 + \lim_{z_1 \to -\infty} \left. \frac{\partial_{z_1} \hat{p}(z)}{f_{\mathbf{e}_1}(z_1)} \right|_{z_2 = \tau z_1} = \lim_{z_1 \to -\infty} F_{\mathbf{e}_2|\mathbf{e}_1}(\tau z_1 | z_1),$$

for almost all $\tau \in (-1, 1)$.

Proof. First, using the change of variables $t_i = e_i - z_i$, $i = 1, 2$, we get that

$$|\hat{q}(z)| = \int_{\underline{\gamma}_1}^{\bar{\gamma}_1} \int_{\underline{\gamma}_2}^{\bar{\gamma}_2} g(-t_1, -t_2) f_{\mathbf{e}}(t_1 + z_1, t_2 + z_2)\, dt_2\, dt_1.$$

Next, since g takes values between zero and one and $f_{\mathbf{e}}$ is nonnegative,

$$\partial_{z_1} |\hat{q}(z)| = |\partial_{z_1} \hat{q}(z)| = \int_{\underline{\gamma}_1}^{\bar{\gamma}_1} \int_{\underline{\gamma}_2}^{\bar{\gamma}_2} g(-t_1, -t_2) \partial_{z_1} f_{\mathbf{e}}(t_1 + z_1, t_2 + z_2)\, dt_2\, dt_1$$

$$\leq \int_{\underline{\gamma}_1}^{\bar{\gamma}_1} \int_{\underline{\gamma}_2}^{\bar{\gamma}_2} |\partial_{z_1} f_{\mathbf{e}}(t_1 + z_1, t_2 + z_2)|\, dt_2\, dt_1 \leq \int_{z_1+\underline{\gamma}_1}^{z_1+\bar{\gamma}_1} \int_{z_2+\underline{\gamma}_2}^{z_2+\bar{\gamma}_2} |\partial_{e_1} f_{\mathbf{e}}(e_1, e_2)|\, de_1\, de_2.$$

Fix any $\tau \in (-1, 1)$ and set $z_2 = \tau z_1$. For all $e_1 \in [z_1 + \underline{\gamma}_1, z_1 + \bar{\gamma}_1]$ we have

$$\tau e_1 - \max\left\{\tau \underline{\gamma}_1, \tau \bar{\gamma}_1\right\} \leq \tau z_1 \leq \tau e_1 - \min\left\{\tau \underline{\gamma}_1, \tau \bar{\gamma}_1\right\}.$$

Moreover, for all $e_2 \in [z_2 + \underline{\gamma}_1, z_2 + \bar{\gamma}_1]$ we have

$$\tau z_1 + \underline{\gamma}_2 \leq e_2 \leq \tau z_1 + \bar{\gamma}_2.$$

Combining both sets of inequalities yields

$$\tau e_1 - \max\{\tau\underline{\gamma}_1, \tau\bar{\gamma}_1\} + \underline{\gamma}_2 \leq e_2 \leq \tau e_1 - \min\{\tau\underline{\gamma}_1, \tau\bar{\gamma}_1\} + \bar{\gamma}_2.$$

Let $\underline{e} = \underline{\gamma}_2 - \max\{\tau\underline{\gamma}_1, \tau\bar{\gamma}_1\}$ and $\bar{e} = \bar{\gamma}_2 - \min\{\tau\underline{\gamma}_1, \tau\bar{\gamma}_1\}$. Condition (i) of Assumption 8 then implies that there exists some e^* such that for all e in the integration region, if $e_1 < e^*$, then $f_{\mathbf{e}}(e) \geq 0$. Therefore, there exists some z^* such that for all $z_1 < z^*$

$$\begin{aligned}
\partial_{z_1} |\hat{q}(z)| \Big|_{z_2=\tau z_1} &\leq \int_{\underline{\gamma}_1}^{\bar{\gamma}_1} \int_{\underline{\gamma}_2}^{\bar{\gamma}_2} \partial_{z_1} f_{\mathbf{e}}(t_1 + z_1, t_2 + z_2)\, dt_2\, dt_1 \Bigg|_{z_2=\tau z_1} \\
&= \partial_{z_1} \int_{z_1+\underline{\gamma}_1}^{z_1+\bar{\gamma}_1} \int_{z_2+\underline{\gamma}_2}^{z_2+\bar{\gamma}_2} f_{\mathbf{e}}(e)\, de_2\, de_1 \Bigg|_{z_2=\tau z_1} \\
&= f_{\mathbf{e}_1}(z_1 + \bar{\gamma}_1) \int_{\tau z_1+\underline{\gamma}_2}^{\tau z_1+\bar{\gamma}_2} f_{\mathbf{e}_2|\mathbf{e}_1}(e_2|z_1 + \bar{\gamma}_1)\, de_2 \\
&\quad - f_{\mathbf{e}_1}(z_1 + \underline{\gamma}_1) \int_{\tau z_1+\underline{\gamma}_2}^{\tau z_1+\bar{\gamma}_2} f_{\mathbf{e}_2|\mathbf{e}_1}(e_2|z_1 + \underline{\gamma}_1)\, de_2.
\end{aligned}$$

Note that, for $\gamma_1 \in \{\bar{\gamma}_1, \underline{\gamma}_1\}$ we have that

$$\int_{\tau z_1+\underline{\gamma}_2}^{\tau z_1+\bar{\gamma}_2} f_{\mathbf{e}_2|\mathbf{e}_1}(e_2|z_1 + \gamma_1)\, de_2 = F_{\mathbf{e}_2|\mathbf{e}_1}(\tau z_1 + \bar{\gamma}_2|z_1 + \gamma_1) - F_{\mathbf{e}_2|\mathbf{e}_1}(\tau z_1 + \underline{\gamma}_2|x_1 + \gamma_1).$$

Hence, applying condition (ii) of Assumption 8, it follows that

$$\lim_{z_1 \to -\infty} \frac{\partial_{z_1} \hat{q}(z)}{f_{\mathbf{e}_1}(z_1)} \Bigg|_{z_2=\tau z_1} = 0,$$

for almost every $\tau \in (-1, 1)$. Finally, the result follows from the fact that

$$\partial_{z_1} \left(\int_{z_1}^{\infty} \int_{z_2}^{\infty} f_{\mathbf{e}}(e)\, de_2\, de_1 \right) = -f_{\mathbf{e}_1}(z_1) \left[1 - F_{\mathbf{e}_2|\mathbf{e}_1}(z_2|z_1) \right]. \qquad \blacksquare$$

C. Online Supplementary Materials

C.1. Games with multiple actions

Our results can be generalized to games with more than two actions. Suppose that firms choose actions from $Y_i = \{1, \ldots, Y_{d_Y}\}$, $d_Y < +\infty$. Player i's payoffs from outcome y are given by

$$\alpha_{i,y}(\mathbf{w}) + [\beta_{i,y_i}(\mathbf{w})\mathbf{z}_{i,y_i} - \mathbf{e}_{i,y_i}]. \tag{1}$$

Note that in contrast to the main that now we have an action-specific covariate and shock for every firm.

The following assumption is a standard location and scale normalizations of the payoffs.

Assumption 11

(i) $\alpha_{i,(0,y_{-i})}(w) = \beta_{i,0}(w) = 0$ for all i, y_{-i}, and w; $\mathbf{e}_{i,0} = 0$ a.s. for all i.

(ii) $\beta_{i,y_i}(w) \neq 0$ for all i, $y_i \neq 0$, and w.

Let $\mathbf{z} = (\mathbf{z}_{i,y_i})_{i \in I, y_i \in Y_i \setminus \{0\}}$ be a $d_Z = \sum_i d_{Y_i}$-dimensional vector of payoff relevant action-specific covariates; $x = (z^\top, w^\top)^\top$ be the vector of all observed covariates; and $\mathbf{e} = (\mathbf{e}_{i,y_i})_{i \in I, y_i \in Y_i \setminus \{0\}}$ be a vector of payoff shocks. We allow shocks to be correlated and we impose no restrictions on the sign of $\alpha_{i,y}(\,\cdot\,)$.

We group all payoff parameters and $\Sigma(\,\cdot\,)$ into a single parameter $\theta \in \Theta$.

Proposition C.1 *Under assumptions 1–4, and 11, both θ_0 and h_0 are identified, and any solution concept S nested into $S_R(\theta)$ is discernible relative to the set of parameters that satisfy assumptions 1–3, and 11.*

Proof. Similar to the proof of Proposition 4.1 we can turn a game with many actions to a game with two actions by sending z_{i,y_i} to $+\infty$ or $-\infty$, and then apply Proposition A.2 to identify the payoff parameters. Then similarly to the proof of Proposition 4.2, identification of h_0 follows from completeness of the exponential family of distributions. The latter automatically implies discernibility of Nash solution concept in rationalizability. ■

C.2. Proof of Nondiscernibility of PNE and SAA

Let $f_{\mathbf{e}}$ denote the p.d.f. of $\mathbf{e}$. Our assumptions imply that $f_{\mathbf{e}}(e_1, e_2) > 0$ and $f_{\mathbf{e}}(e_1, e_2) = f_{\mathbf{e}}(e_2, e_1)$ almost everywhere on $\mathbb{R}^2$. For each possible outcome $y \in \{0,1\}^2$, let $p_{\text{PNE}}(y; \eta')$ and $p_{\text{SAA}}(y; \eta)$ denote the probabilities of the outcome according to each of the two solution concepts under consideration.

Fix any parameter value $\eta \geq 0$. We will show that there exists some $\eta' \geq 0$ such that $p_{\text{PNE}}(y; \eta') = p_{\text{SAA}}(y; \eta)$ for every possible outcome y. If $\eta = 0$, then we can simply set $\eta' = 0$. Hence, for the rest of the proof, we assume that $\eta > 0$.

On one hand, if $\eta' = \eta$, then

$$\begin{aligned} p_{\text{PNE}}((0,0); \eta') &= \int_{\eta'}^{\infty} \int_{\eta'}^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 \\ &< \int_{\eta}^{\infty} \int_{\eta}^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 + \int_0^{\eta} \int_0^{\eta} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 \\ &= p_{\text{SAA}}((0,0); \eta). \end{aligned}$$

(See Figure 1). On the other hand, if $\eta' = 0$, then

$$\begin{aligned} p_{\text{PNE}}((0,0); \eta') &= \int_0^{\infty} \int_0^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 \\ &> \int_{\eta}^{\infty} \int_{\eta}^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 + \int_0^{\eta} \int_0^{\eta} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 \\ &= p_{\text{SAA}}((0,0); \eta). \end{aligned}$$

Since $p_{\text{PNE}}((0,0); \eta')$ is continuous in η', there exists some $\eta' \in (0, \eta)$ such that $p_{\text{PNE}}((0,0); \eta') = p_{\text{SAA}}((0,0); \eta)$. Fix such η'.

Since $p_{\text{PNE}}((1,1); \eta') = p_{\text{SAA}}((1,1); \eta)$ and there are only four possible outcomes it follows that

$$p_{\text{PNE}}((1,0); \eta') + p_{\text{PNE}}((0,1); \eta') = p_{\text{SAA}}((1,0); \eta) + p_{\text{SAA}}((0,1); \eta).$$

Now, we will show that $p_{\text{PNE}}((1,0); \eta') = p_{\text{PNE}}((0,1); \eta')$ and $p_{\text{SAA}}((1,0); \eta) = p_{\text{SAA}}((0,1); \eta)$. This implies that the probabilities of all outcomes are the same under both solution concepts. For PNE, we have that

$$p_{\text{PNE}}((1,0); \eta') = \int_{-\infty}^{0} \int_0^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1 + \int_0^{\eta'} \int_{e_1}^{\eta'} f_{\mathbf{e}}(e_1, e_2)\, de_2\, de_1$$

$$
\begin{aligned}
&= \int_{-\infty}^{0}\int_{0}^{\infty} f_{\mathbf{e}}(e_2, e_1)\, de_1\, de_2 + \int_{0}^{\eta'}\int_{e_2}^{\eta'} f_{\mathbf{e}}(e_2, e_1)\, de_1\, de_2 \\
&= \int_{-\infty}^{0}\int_{0}^{\infty} f_{\mathbf{e}}(e_1, e_2)\, de_1\, de_2 + \int_{0}^{\eta'}\int_{e_2}^{\eta'} f_{\mathbf{e}}(e_1, e_2)\, de_1\, de_2 \\
&= p_{\text{PNE}}((0,1);\eta'),
\end{aligned}
$$

where the second equality follows from using the change of variables $(e_1, e_2) \to (e_2, e_1)$, and the third one from the symmetry fo $f_{\mathbf{e}}$. The argument for SAA is completely analogous. ■

C.3. Alternative Entry Subsidy for the Motivating Example

A form of subsidy that is more common in practice consists of giving a lump sum subsidy $\hat{\tau} > 0$ to any firm that enters a market with some observable characteristics (see, e.g., Goolsbee (2002)). Under the PNE assumption, every market that would be served without the policy would also be served with the policy. Hence, the policy has an unambiguously positive effect (abstracting from the cost). However, this need not be the case under SAA.

Proposition C.2 *Suppose that firms profits are given by*

$$
\boldsymbol{\pi}_i(y) = y_i \cdot \Big[\alpha + \eta(1 - y_{-i}) - \mathbf{e}_i\Big],
$$

firms make entry decisions in accordance with the SAA model, and $\mathbf{e}$ *is normally distributed with zero mean and the identity matrix as a covariance matrix. There exists an open set* $\Xi \subseteq \mathbb{R}^2$ *and a threshold* $\bar{\tau}$ *such that if* $(\alpha, \eta) \in \Xi$ *and* $\hat{\tau} < \bar{\tau}$*, then the probability that a market is not served is increasing in the size of the subsidy.*

Proof. Under strategic ambiguity there is no entry if either $\mathbf{e}_i > \alpha + \eta + \hat{\tau}$ for $i = 1, 2$, or $\alpha + \tau < \mathbf{e}_i < \alpha + \eta + \hat{\tau}$ for $i = 1, 2$ (See Figure 1). Hence, the probability that a market is not served as a function of $\hat{\tau}$ is given by

$$
P(\hat{\tau}) = \Big[1 - \Phi(\alpha + \eta + \hat{\tau})\Big]^2 + \Big[\Phi(\alpha + \eta + \hat{\tau}) - \Phi(\alpha + \hat{\tau})\Big]^2. \qquad (2)
$$

Taking derivatives

$$P'(\hat{\tau}) = -2\phi(\alpha+\eta+\hat{\tau})\Big[1-\Phi(\alpha+\eta+\hat{\tau})\Big]\dots$$
$$\dots\ +2\Big[\phi(\alpha+\eta+\hat{\tau})-\phi(\alpha+\hat{\tau})\Big]\cdot\Big[\Phi(\alpha+\eta+\hat{\tau})-\Phi(\alpha+\hat{\tau})\Big]. \tag{3}$$

Evaluating when $\hat{\tau}=0$, $\alpha<0$, and $\eta=-\alpha+\sqrt{-\alpha}$ yields

$$\frac{P'(0)}{2\phi(-\alpha)} = \Big[\Phi\left(\sqrt{-\alpha}\right)-1\Big] + \left[1-\frac{\phi(\alpha)}{\phi\left(\sqrt{-\alpha}\right)}\right]\cdot\Big[\Phi\left(\sqrt{-\alpha}\right)-\Phi(\alpha)\Big] \tag{4}$$

When $\alpha \longrightarrow -\infty$, the right-hand side converges to 1. Hence, we must have $P'(0)>0$ when $-\alpha$ is sufficiently large. Since P is continuous, this must also be true in an open set. ■

C.4. Discernibility in Entry Example

Let us consider a modified version of the entry example. Suppose that everything is as in Section 2, except that firm i's profit is given by

$$y_i\cdot\Big[\eta_0(1-y_{-i})+\mathbf{z}-\mathbf{e}_i\Big],$$

where $\mathbf{z}$ is a covariate supported on the whole real line, independent of $\mathbf{e}$, and such that Assumption 3 holds. The researcher observes the joint distribution of outcomes $(\mathbf{y}_1,\mathbf{y}_2,\mathbf{z})$. We claim that, with this added covariate, PNE and SAA are no longer observationally equivalent. Now, if the firm behavior corresponds to SAA, then it cannot be explained by PNE, and vice versa.

Proposition C.3 *In the entry example with a covariate, for every $\eta,\eta'>0$ there exists $z\in\mathbb{R}$ such that SSA with $\eta_0=\eta$ and PNE with $\eta_0=\eta'$ imply different outcome distributions conditional on $\mathbf{z}=z$.*

Proof. It suffices to consider the probability of no entry, i.e., $\mathbf{y}=(0,0)$. Fix any $\eta,\eta'>0$. If $\eta'\geq\eta$, then PNE implies a higher probability of no entry than SAA regardless of the realization of $\mathbf{z}$ (see Figure 1). Hence, we assume for the rest of the proof that $\eta>\eta'$.

The probability of no entry conditional on $\mathbf{z}=z$ is $[1-\Phi(z+\eta)]^2$ under PNE and $[\Phi(z+\eta') - \Phi(z)]^2 + [1 - \Phi(z+\eta')]^2$ under SAA. Let $\chi : \mathbb{R} \to \mathbb{R}$ give the difference between these probabilities as a function of z, i.e.,

$$\chi(z) = [\Phi(z+\eta') - \Phi(z)]^2 + [1 - \Phi(z+\eta')]^2 - [1 - \Phi(z+\eta)]^2.$$

We will show that there exist numbers z such that $\chi(z) \neq 0$.

Note that χ is differentiable and

$$\begin{aligned}\chi'(z) = {} & 2[\Phi(z+\eta') - \Phi(z)][\phi(z+\eta') - \phi(z)] \\ & - 2[1 - \Phi(z+\eta')]\phi(z+\eta') + 2[1 - \Phi(z+\eta)]\phi(z+\eta).\end{aligned}$$

Let $z^* := -\eta'/2 < 0$, so that $z^* = -(z^* + \eta')$ and $z^* + \eta' > 0$. Since $\phi(\,\cdot\,)$ is symmetric around 0, this implies that $\phi(z^*) = \phi(z^* + \eta')$. Therefore the first term of $\chi'(z^*)$ is equal to zero and we have

$$\begin{aligned}\chi'(z^*) &= -2[1 - \Phi(z^*+\eta')]\phi(z^*+\eta') + 2[1 - \Phi(z^*+\eta)]\phi(z^*+\eta) \\ &= 2[\phi(z^*+\eta) - \phi(z^*+\eta')] + 2\Phi(z^*+\eta')\phi(z^*+\eta') - 2\Phi(z^*+\eta)\phi(z^*+\eta) \\ &< 2\Phi(z^*+\eta')\phi(z^*+\eta) - 2\Phi(z^*+\eta)\phi(z^*+\eta) \\ &= -2\phi(z^*+\eta)[\Phi(z^*+\eta) - \Phi(z^*+\eta')] < 0,\end{aligned}$$

where the first inequality follows because ϕ is decreasing on the positive real line, and thus $\phi(z^*+\eta') > \phi(z^*+\eta)$. Since $\chi'(z^*) \neq 0$, there exists an open set Z' such that $\chi(z) \neq 0$ for all $z \in Z'$. ■

Note that the proof of Proposition C.3 considers only the probability of $\mathbf{y} = (0, 0)$. The probability of this outcome in the multiplicity region is zero under any PNE. Hence, the conclusion of the proposition does not depend on any assumptions about equilibrium selection.