Identification of solution concepts for discrete games*

Nail Kashaev† Bruno Salcedo‡

October 15, 2017

Abstract Empirical analyses of discrete games rely on behavioral assumptions that are crucial not just for estimation, but also for the validity of counterfactual exercises and policy implications. We find conditions for a general class of complete-information games under which it is possible to identify whether the behavior of economic agents satisfies some of these assumptions. For example, our results allow us to identify whether and how often firms in an entry game play Nash equilibria, and which equilibria are more likely to be selected.

Keywords Econometrics of games · Multiple equilibria · Market entry · Model selection

JEL classification C52 · C72

*We are thankful for the supervision of Joris Pinkse, Sung Jae Jun and Andrés Aradillas-López, as well as the useful comments from Marc Henry, Robert Marshall, Brendan Kline, Bulat Gafarov, Mark Roberts, Francesca Molinari, Salvador Navarro, Victor Aguiar and, especially, Paul Grieco. 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, University of Western Ontario, nkashaev@uwo.ca

‡Department of Economics, Cornell University, salcedo@cornell.edu

1. Introduction

Solution concepts are a fundamental component of economic models. Many theoretical results rely on assuming that firm behavior must be in, or converge to equilibrium. Estimation of structural parameters is often carried out under the assumption that observed data arises from equilibrium choices. Policy implications are derived from counterfactual analyses which compare equilibria under alternative policies given the estimated parameters. The current work proposes sufficient conditions to identify the validity of different solution concepts for games, in a precise sense to be defined later on.

Instead of assuming a fixed solution concept to estimate payoff parameters, we propose a framework that incorporates the solution concept as part of the parameter space. Our main result is a transparent characterization of the identified set. Under our conditions, a solution concept can rationalize the observed choices if and only if the actual behavior of the agents satisfies the corresponding behavioral assumptions almost surely (Theorem 6.1). As a corollary, we find sufficient conditions for point identification of the true solution concept from a class of solution concepts under consideration (Corollary 6.2).

Our framework can be applied to general solution concepts and general discrete complete-information games either in extensive or strategic form. We illustrate our approach by identifying whether the Nash equilibrium (NE) assumption is satisfied in the context of a two-firm entry model adapted from Bresnahan and Reiss (1990). We also show how incorrectly assuming NE play can yield misleading policy recommendations, even if the payoff parameters of the model are known. Moreover, for the particular entry model that we consider, we are able to identify whether payoffs are private information or common knowledge. As a different application, Appendix E considers an n-player coordination problem, and identifies when and how often do players coordinate on risk-dominant rather than payoff-dominant equilibria.

Our econometric framework consists of three elements. First, we assume that the econometrician observes the joint distribution of endogenous outcomes and exogenous covariates. However, she does not know the distribution of outcomes conditional on the observed and unobserved characteristics of the environment, which we call the *distribution of play*. The second element of our framework is a *structural index* which determines the players' von Neumann Morgenstern (vNM) utilities over outcomes, and depends on the characteristics of the environment through a function known up to a finite-dimensional vector of structural parameters. Finally, in order to relate these parameters to the observed data, we introduce a formal definition of *solution concept*. A solution concept specifies a set of admissible distributions over outcomes depending on the realized exogenous characteristics of the environment. For example, it could specify that entry decisions must arise from NE of the game.

Our goal is to characterize which solution concepts are satisfied by actual behavior, meaning that they contain the true distribution of play. Our first step is to establish point identification of the distribution of play (Proposition 4.1). That is, we non-parametically identify the distribution of choices conditional on *both* the observed and unobserved characteristics of the environment.

The distribution of play is interesting on its own right, and it might be enough for some applications. For instance, if a policymaker is willing to assume that the distribution of play is robust to policy interventions, then it is all the information she needs in order to derive valid policy implications.[1] We focus on identification of solution concept because of two main reasons. First, there are situations in which the external validity of a solution concept backed up by theory is more plausible than that of the distribution of play. Moreover, the distribution of play is specific to each specific context. In contrast, a solution concept such as NE can be applied and compared across different environments. Hence, focusing on the solution concept rather than the distribution of play allows one to accumulate the lessons learned from different studies.

Second, one can assume the weakest solution concept (in our setting it is rationalizabilty), build the confidence set for the payoff parameters, and then estimate robust confidence bands for the counterfactual of interest. However, as our motivating example shows, these bands can be uninformative (in our example small duopoly subsidy can lead to both increase and decrease of probability of occurrence of monopolies) since one is imposing the weakest restrictions on the behavior of the players. But if one establishes validity of the stronger solution concept (e.g., NE in pure strategies) using our methodology, then one can build more informative bounds for the counterfactual predictions. In other words, our methodology allows one to determine the strongest restrictions on players' behavior that are still consistent with observed data.

[1] Jun and Pinkse (2016) provide novel sharp bounds for the counterfactual predictions required to derive such implications.

We use two assumptions to identify the distribution of play. The first one is an exclusion restriction requiring choices to be independent of *some* of the covariates, conditional on the structural index. The second one is a richness condition requiring the family of distributions of the structural index conditional on the values of the excluded covariates to be boundedly complete. The first assumption guarantees that the true distribution of play satisfies a certain system of integral equations. The second assumption guarantees that this system has a unique solution.

The richness assumption is satisfied, for instance, if the payoffs have normally or extreme-valued distributed additive errors, as is commonly assumed in applied work. As for the exclusion restriction, one may be concerned that some covariates, which affect the payoffs, may also affect the way people choose equilibria. However, we allow equilibrium selection to depend both on the structural index and unobserved heterogeneity, and we only require independence conditional on the structural index. Hence, we still allow the excluded covariates to affect the way people choose equilibria, as long as they do so through preferences. Additionally, if there is good reason to believe that these covariates affect equilibrium selection, even after conditioning on payoffs, this hypothesis could be incorporated as part of the solution concept and tested accordingly.

We also point identify the structural parameters (Proposition 5.1). For that purpose, we use a high-level assumption that requires different parameter values to make different predictions independently of the true solution concept. This high-level assumption encompasses standard identification assumptions discussed below, and, under our exclusion restriction, it is both *sufficient and necessary* for point identification of the structural parameters. We also extend standard identification-at-infinity strategies in a way that only requires mild behavioral assumptions, which allow some forms of collusion and departures from rationality. In the context of our entry game, our identification-at-infinity approach allows us to point identify the correlation between error terms.

Identifying the distribution of play under incomplete information is more complicated, because solution concepts for incomplete information games are often inconsistent with our exclusion restriction. Our general methodology cannot be applied in such cases. In the context of our entry example, we still manage to identify whether profit functions are private information or common knowledge (Proposition 7.1). We do so by directly comparing the Aumann expectations of different solution concepts, and without requiring point identification of the distribution of play.

1.1. Related literature

For discrete games, Beresteanu et al. (2011) and Galichon and Henry (2011) characterize the sharp identified set for the payoff parameters using support functions and Choquet capacities, respectively. Henry and Mourifie (2012) provides a characterization for 2×2 games. These papers rely on assuming NE (allowing for public randomization) or correlated equilibria. In our framework, instead of working with a fixed solution concept, both the solution concept and the way players choose equilibrium are parameters of interest for which we establish point identification. Moreover, our approach can be applied to non-convex solution concepts. See Section 6 for further discussion.

Different existing methods to establish point identification of payoff distribution in discrete games use exclusion restrictions.[2] Identification-at-infinity approaches have been used elsewhere in the literature. For instance, Tamer (2003) and Ciliberto and Tamer (2009) use it in the context of entry games assuming NE in pure strategies (PNE). Bajari et al. (2010) establishes point identification for a general class of complete information games under the NE assumption. Kline (2015b) point identifies mean utility parameters assuming rationalizability. We contribute to this literature by showing how one can identify the *joint* distribution of unobservables using only mild rationality assumptions. Notably, Kline (2016) point identifies mean utility parameters using only bounded covariates. However, this approach does not suit our objective to identify the solution concept since it requires a much stronger behavioral assumption, namely, PNE.[3]

The approaches discussed in the preceding paragraph rely on imposing enough restrictions on the behavior of players so that the researcher essentially knows the probability of some outcomes for some values of the covariates. Under rationalizability or NE, this occurs at the limit when the game is dominance solvable. For entry games and assuming PNE, this occurs everywhere because duopolies and

[2] Examples include Bjorn and Vuong (1984), Bresnahan and Reiss (1991), Tamer (2003), Berry and Tamer (2006), Bajari et al. (2010), Bajari et al. (2011), Aradillas-López and Tamer (2008) and Kline (2015a). See De Paula (2013) for a review of the literature on this topic.

[3] Another attractive feature of Kline (2016) is that he does not impose parametric restrictions on the distribution of unobservables. Such flexibility, however, comes with the cost of non-identification of the distribution of unobservables.

markets with no entrants only arise when it is dominant for both firms to enter or to stay out, respectively. We contribute to this literature by showing that, in the presence of excluded covariates, such restrictions on behavior are not just sufficient, but also necessary for point identification of the distribution of payoffs.

We are not the first to exploit the power of completeness assumptions coupled with exclusion restrictions in order to identify behavior patterns of agents.[4] Berry and Haile (2014)—as well as other related papers—apply a similar strategy to a model of oligopolistic competition that allows, among other things, to discriminate between different models of competition. An important difference between their setting and ours is that they consider continuous games, while we consider discrete games. They rely crucially on having an uncountable set of outcomes in order to relax the completeness assumption to some extent.

The problem of identifying the relevance of particular solution concepts has been extensively addressed using experimental data from laboratory settings. See, for instance, Camerer (2003). However, much less is known about the behavior of firms and other economic agents outside the laboratory.

In a recent paper, Kashaev (2016) proposes a sieve likelihood-ratio-type procedure to test the NE assumption in binary games with complete information. Chiappori et al. (2002) and Palacios-Huerta (2003) analyze data from penalty kicks in professional football to investigate whether it resembles a mixed strategy equilibrium of a zero-sum game. There are two important differences between their approach and ours. First, they assume that the game is zero-sum, which eliminates the possibility of multiple equilibria. Second, instead of using a structural model, they test two reduced form implications of Nash play: indifference between alternatives, and serial independence.

Other papers follow a different approach that consists of relaxing behavioral assumptions instead of testing them. Aradillas-López and Tamer (2008) and Kline (2015a) work with k-level rationality instead of NE. Grieco (2014) and Magnolfi and Roncoroni (2016) relax the complete information assumption. In Section 2.2, we show that identifying solution concepts can be important for counterfactual analysis, regardless of whether payoff parameters are known.

[4] 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 it to achieve identification of Engel curves. Hoderlein et al. (2012) imposes bounded completeness in the context of structural models with random coefficients.

1.2. Notation

Throughout the paper, deterministic vectors and functions are denoted by lower-case Latin letters, random objects by bold letters, sets by upper-case letters, and parameters by Greek letters. Also, given a family $x = (x_k)_{k\in K}$ (e.g., a strategy profile, or a vector of covariate values) and a particular index value $k \in K$, we use the notation $x = (x_k, x_{-k})$ where $x_{-k} = (x_j)_{j\in K\setminus\{k\}}$.

2. Motivating example: an entry game

In order to illustrate our general results and suggest potential applications, we consider an entry model adapted from Bresnahan and Reiss (1990). Two firms $i \in \{1, 2\}$ must choose whether to enter a market ($y_i = 1$) or not ($y_i = 0$). Firm i's profit is given by

$$\mathbf{u}_{0i}(y) = \left(\beta_{0i}\mathbf{x}_i - \tilde{\beta}_{0i}y_{-i} - \tilde{\mathbf{e}}_i\right) y_i,$$

where y_{-i} denotes the choice of i's competitor, $\beta_{0i}, \tilde{\beta}_{0i} > 0$, $i = 1, 2$, are unknown fixed parameters, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)^\intercal$ is a vector of covariates with support $X = \mathbb{R}^2$, and $\tilde{\mathbf{e}} = (\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2)^\intercal$ is a vector of error terms observed by the firms but unobserved by the researcher.[5] The researcher observes the joint distribution of $\mathbf{x}$ and the entry decisions $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2)^\intercal$.

2.1. Solution concepts

Our main objective is to identify whether the firms' behavior satisfies assumptions that can be described by nonempty-valued correspondences mapping exogenous characteristics (both observed and unobserved) into sets of admissible distributions over endogenous outcomes. We call such correspondences solution

[5] We omit the intercept parameters for exposition purposes only. All the result are valid if one assumes that $\mathbf{u}_{0i}(y) = \left(\bar{\beta}_{0i} + \beta_{0i}\mathbf{x}_i - \tilde{\beta}_{0i}y_{-i} - \tilde{\mathbf{e}}_i\right) y_i$, $i = 1, 2$.

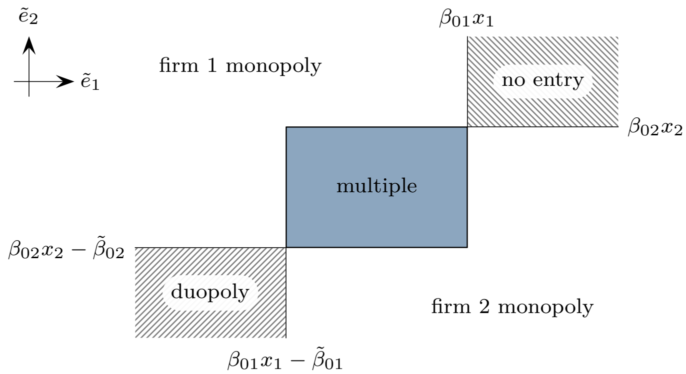

Figure 1 – q_{R} and q_{NE} correspondences for different realizations of $\tilde{\mathbf{e}}$.

concepts.[6] For instance, the assumption that each firm maximizes its profits and this fact is common knowledge among the firms (rationalizability) can be characterized by the solution concept q_{R}, illustrated in Figure 1. When $\tilde{e}_i < \beta_{01}x_i - \tilde{\beta}_{0i}$, entering the market is strictly dominant for player i. When $\tilde{e}_i > \beta_{01}x_i$, staying out of the market is strictly dominant for player i. This results in four regions of the payoff space in which the game has a unique rationalizable outcome, and q_{R} contains a unique distribution assigning full probability to it. In the remaining region—the *multiplicity region*—rationalizability imposes no restrictions on behavior, and thus q_{R} maps to the set of all distributions over Y.

As another important example, let q_{NE} be the correspondence that captures the assumptions that firms always play a NE of the corresponding simultaneous complete-information game. Since NE actions are rationalizable, q_{NE} coincides with q_{R} in the regions in which q_{R} is single-valued. However, in the multiplicity region, q_{NE} consists of only three points: a strictly mixed NE and two degenerate distributions corresponding to pure strategy NE. We will impose assumptions under which, assuming q_{R}, it is possible to identify whether the choices of the firms satisfy q_{NE}.

2.2. Policy implications

Assuming an incorrect solution concept can result in completely misleading policy implications. To see this, suppose that a policymaker wants to maximize

[6] See Section 6 for a formal definition.

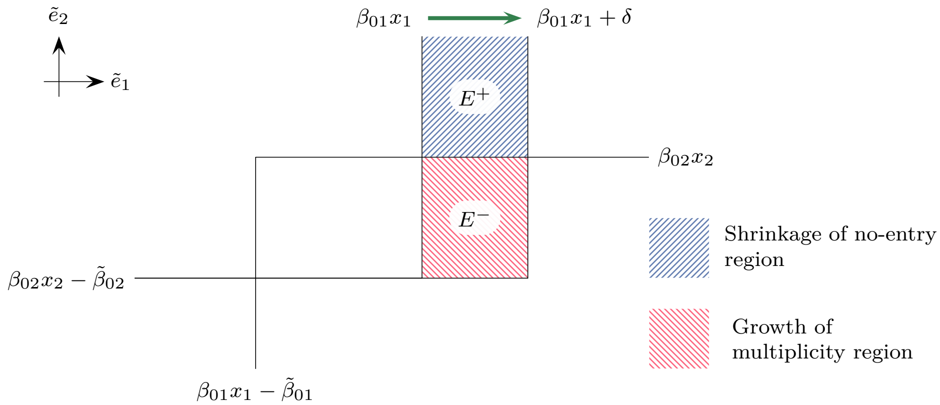

Figure 2 – Effect of the proposed policy.

the number of markets that are served by at least one firm. As a policy instrument, she can choose to offer a subsidy $\delta > 0$ to one firm, say firm 1, for entering markets in which firm 2 does not enter.[7] She actually knows the payoff parameters, and she evaluates the policy assuming that firms always play PNE.

However, suppose that the policymaker's assumption is incorrect. Firms in fact do *not* always play PNE. Instead, their behavior is rationalizable, but they never enter when their profit functions lie in the multiplicity region. This could happen, for instance, if the firms were ambiguity averse when facing strategic uncertainty, and they used maxmin strategies when the game is not dominance solvable. Let q' be the correspondence that describes this behavior.

Pure equilibria 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, given the true behavior of the firms, the effect of the policy is always smaller than under the PNE hypothesis. It can even have the opposite direction for some parameter values. This happens because the policy also increases the probability of the multiplicity region and, under the true behavior, firms never enter in this region. A firm might be willing to forego the subsidy, for fear that

[7]Entry subsidies are often used to incentivize the provision of strategic goods such as broadband internet access (Goolsbee, 2002). This section considers a specific kind of subsidy that allows for a clear and stark exposition. Appendix A discusses more realistic subsidy schemes.

another firm might also enter 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 probability of regions E^+ and E^- in Figure 2. For some parameter values and realized covariate values the adverse effect can actually dominate, and the policy may actually increase the probability that a market is not served. For example, one can easily verify that this is the case whenever the error terms are i.i.d. standard normal and $\Phi(\beta_{02}x_2 - \tilde{\beta}_{02}) < 2\Phi(\beta_{02}x_2) - 1$, where Φ is the standard normal c.d.f..

3. Econometric framework

3.1. Data generating process

Each instance of the environment is characterized by an endogenous outcome y from a finite set Y, a vector of exogenous characteristics observed by the researcher $x \in X \subseteq \mathbb{R}^{d_x}$, and a vector of error terms $e \in E \subseteq \mathbb{R}^{d_e}$.

Assumption 1 (Data generating process) Random objects $\mathbf{y} : \Omega \to Y$, $\mathbf{x} : \Omega \to X$, and $\mathbf{e} : \Omega \to E$, are defined on a probability space $(\Omega, \mathscr{F}, \Pr)$. The error terms $\mathbf{e}$ are uniformly distributed on $[0,1]^{d_e}$, and are independent of $\mathbf{x}$. The researcher observes the joint distribution of $(\mathbf{x}, \mathbf{y})$.

The restriction on the distribution of the $\mathbf{e}$ is a normalization. The unobserved characteristics of the environment are modeled as a function $\tilde{\mathbf{e}} = \tilde{e}(\mathbf{e}, \mathbf{x}, \rho_0)$, known up to the unknown parameter ρ_0. In order to distinguish $\tilde{\mathbf{e}}$ from $\mathbf{e}$, we call $\tilde{\mathbf{e}}$ *structural* errors. Since we don't impose any restrictions on $\tilde{e}$, the structural errors can have general distributions and correlation structures. For example, suppose that the structural errors $\tilde{\mathbf{e}}$ in the entry game from Section 2 are normally distributed with variances normalized to be 1, and correlation $\rho_0 \in (-1, 1)$ independently of $\mathbf{x}$. Then, it suffices to set:

$$\tilde{e}(e, x, \rho_0) = \begin{pmatrix} \Phi^{-1}(e_1) \\ \rho_0 \Phi^{-1}(e_1) + \sqrt{1-\rho_0^2}\,\Phi^{-1}(e_2) \end{pmatrix}. \tag{1}$$

With slight abuse of notation, we identify distributions over Y with vectors on the $\|Y\|$-dimensional simplex $\Delta(Y)$. We denote the observed distribution of $\mathbf{y}$ conditional on $\mathbf{x}$ by $\mu_0(\mathbf{x})$, and the unknown distribution of $\mathbf{y}$ conditional on $\mathbf{x}$ and $\mathbf{e}$ by $h_0(\mathbf{e}, \mathbf{x})$. Let H be the set of measurable functions from $E \times X$ to $\Delta(Y)$. We call each such a function a possible *distribution of play*, and h_0 the true distribution of play. A possible distribution of play $h \in H$ is *consistent with the observed data* if and only if:

$$\mu_0(\mathbf{x}) = \mathbb{E}[\, h(\mathbf{e}, \mathbf{x}) | \mathbf{x}\,] \quad \text{a.s..} \tag{2}$$

By construction, we know that h_0 is consistent with the observed data.

Example 3.1 Consider the entry model from Section 2, and suppose that agents always randomize uniformly across all NE. For (e, x) such that $(\tilde{e}(e, x, \rho_0), x)$ is in the multiplicity region, let $p^*_{-i} = (\beta_{i0} x_i - \tilde{e}_i(e, x, \rho_0))/\beta_{03}$ be the probability of firm $-i$ entering in the unique mixed NE. We suppress the explicit dependence on (e, x) for brevity of notation. Then, the true distribution of play is characterized (up to a null-set) by:

$$h_{\text{UNE}}(e, x) = \begin{cases} \delta_{(1,1)} & \text{if} \quad \tilde{e}_i(e, x, \rho_0) < \beta_{0i} x_i - \beta_{03} \text{ for } i = 1, 2 \\ \delta_{(0,0)} & \text{if} \quad \tilde{e}_i(e, x, \rho_0) > \beta_{0i} x_i \text{ for } i = 1, 2 \\ \delta_{(1,0)} & \text{if} \quad \tilde{e}(e, x, \rho_0) \in E_1(x) \\ \delta_{(0,1)} & \text{if} \quad \tilde{e}(e, x, \rho_0) \in E_2(x) \\ \psi^*(\tilde{e}(e, x, \rho_0), x) & \text{otherwise} \end{cases},$$

where δ_y denotes the distribution that assigns full probability to outcome y,

$$E_i(x) = \left\{ \tilde{e} \,\middle|\, \begin{array}{c} (\tilde{e}_i < \beta_{0i} x_i - \beta_{03} \quad \text{and} \quad \tilde{e}_{-i} > \beta_{0-i} x_{-i} - \beta_{03}) \\ \text{or} \quad (\tilde{e}_i < \beta_{0i} x_i \quad \text{and} \quad \tilde{e}_{-i} > \beta_{0-i} x_{-i}) \end{array} \right\}$$

corresponds to the region where i's monopoly is the only rationalizable outcome for $i = 1, 2$ (see Figure 1), and

$$\psi^*(\tilde{e}, x) = \frac{1}{3}\Big((1 - p_1^*)(1 - p_2^*),\ 1 + p_2^*(1 - p_1^*),\ 1 + p_1^*(1 - p_2^*),\ p_1^* p_2^*\Big)^{\mathsf{T}}$$

specifies the probability of each outcome $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, respectively, in the multiplicity region.

3.2. Structural parameters

The researcher is interested in a random structural index that takes values in a known set $U \subseteq \mathbb{R}^{d_u}$, and is given by a parametric function of the primitive exogenous characteristics of the environment. To fix ideas, let $U \subseteq \mathbb{R}^{\|I\| \times \|Y\|}$, and interpret the value of the index as representing the agents' preferences over outcomes.

Assumption 2 (Structural parameters) The structural index $\mathbf{u}_0 : \Omega \to U$ is given by $\mathbf{u}_0 = u(\beta_0, \mathbf{e}, \mathbf{x})$, where $u : B \times E \times X \to U$ is a known measurable function and β_0 is a finite-dimensional vector of parameters belonging to a known set $B \subseteq \mathbb{R}^{d_\beta}$.

Note that the structural index is written as a function of $\mathbf{e}$, rather than the structural errors $\tilde{\mathbf{e}}$. Hence, one should think of the vector of structural parameters β_0 as containing the parameters ρ_0, which map $\mathbf{e}$ to $\tilde{\mathbf{e}}$.

We have not yet introduced any link connecting β_0 to the observed data described in Assumption 1. Hence, we need additional assumptions, both in order to learn something about β_0 from the data, and to make predictions based on each candidate value of β_0.

3.3. Solution concepts

We focus on a special kind of structural assumptions which make set predictions about the distribution of endogenous outcomes as a function of the characteristics of the environment.

Definition 1 A *solution concept* is a nonempty-valued and closed-valued correspondence $q : B \times E \times X \rightrightarrows \Delta(Y)$, such that

$$\left\{\omega \in \Omega \,\middle|\, \quad \Psi \cap q(\beta, \mathbf{e}(\omega), \mathbf{x}(\omega)) \neq \emptyset\right\} \in \mathscr{F}, \tag{3}$$

for every $\beta \in B$ and every *closed* set $\Psi \subseteq \Delta(Y)$.

Equation (3) is a measurability condition requiring that $q(\beta, \mathbf{e}, \mathbf{x})$ should be a random set for every $\beta \in B$ (Molchanov, 2006). Although we define solution concepts in terms of primitives $\mathbf{x}$ and $\mathbf{e}$, and not in terms of $\mathbf{u}_0$, the fact that they depend on the structural parameters provides the missing link between the data and β_0.

Definition 2 A pair $(\beta, h) \in B \times H$ *jointly satisfies* q if:

$$h(\mathbf{e}, \mathbf{x}) \in q(\beta, \mathbf{e}, \mathbf{x}) \quad \text{a.s.}, \tag{4}$$

and q is *satisfied* if (β_0, h_0) jointly satisfies it.

Example 3.2 An example of a solution concept is the NE correspondence q_{NE} described in Section 2.1. Since q_{NE} consists of locally isolated points, it presumes that $h(e, x)$ corresponds to a particular NE given (e, x). In particular, the distribution of play h_{UNE} from Example 3.1 does not satisfy q_{NE}. However, it does satisfy the solution concept *Nash equilibrium with public randomization* q_{PRNE} defined by $q_{\text{PRNE}}(\beta, e, x) = \text{co}(q_{\text{NE}}(\beta, e, x))$, where $\text{co}(\,\cdot\,)$ denotes the convex-hull operator.

3.4. Distribution of play and selection mechanism

The distribution of play is well-defined independently of any notion of solution concept. However, when a pair (β, h) jointly satisfies some q, then the distribution of play is closely related to the selection mechanism.

Definition 3 For given q, β, e, and x, a selection mechanism is a *random* probability measure over $q(\beta, e, x)$. That is, $\mathbf{sel} : q(\beta, e, x) \to \Delta(q(\beta, e, x))$.

Note that the selection mechanism is allowed to be random. That is, in two different markets with the same realization of observed and unobserved payoff shifters, players can use two different selection mechanisms.

Using the above definition of the selection mechanism, one can show that if a pair (β, h) jointly satisfies some q, then the distribution of play is an average of the composite of the selection mechanism and the solution concept. Indeed, suppose that $q(\beta, \mathbf{x}, \mathbf{e})$ is a finite set with probability 1, and let l be an arbitrary element of $q(\beta, \mathbf{x}, \mathbf{e})$. Then

$$h(\mathbf{e}, \mathbf{x}) = \mathbb{E}\left[\sum_{l \in q(\beta, \mathbf{x}, \mathbf{e})} \mathbf{sel}(l) l \middle| \mathbf{x}, \mathbf{e} \right] = \sum_{l \in q(\beta, \mathbf{x}, \mathbf{e})} \mathbb{E}\left[\mathbf{sel}(l) | \mathbf{x}, \mathbf{e} \right] l,$$

where $\mathbf{sel}(l)$ is a probability that the equilibrium l is chosen.

3.5. Identification

The main objective of the paper is to determine which solution concepts accurately characterize the behavior of the agents, that is, which solution concepts are satisfied by (β_0, h_0) in the sense of Definition 2. To address this issue, we have to determine which distributions of play and structural parameters are consistent with the data. To narrow these sets, we restrict the parameter space through a series of structural assumptions, each of them stating that (β_0, h_0) belongs to some known set $\Theta \subseteq B \times H$.

Definition 4 The *sharp identified set* for (β_0, h_0) under assumption Θ consists of the set of pairs $(\beta, h) \in \Theta$ such that h is consistent with the observed data, i.e., it satisfies Equation (2).

Definition 5 A solution concept q is *consistent with the data* under Θ if it is jointly satisfied by some (β, h) belonging to the sharp identified set under Θ.

We have thus defined two different properties for any given solution concept. We say that it is satisfied if the actual distribution of play belongs to it almost surely, and we say that it is consistent if the observed data could be generated by a choice pattern that satisfies it. In general, only consistency can be directly tested. However, the researcher may be interested in whether a solution concept accurately characterizes behavior, as this may be crucial for the validity of counterfactual analyses. Our main result (Theorem 6.1) establishes conditions under which these two properties are equivalent, thus enabling to test which solution concepts are satisfied.

4. Identifying the distribution of play

In this section, we make two assumptions that allow us to solve (2), in order to recover the true distribution of play from the observed distribution μ_0. The first assumption is an exclusion restriction requiring that some of the observed covariates only affect choices through the structural index. The second assumption is a richness condition requiring that these covariates have sufficient heterogeneity and generate sufficient variation in the distribution of the structural index.

Suppose the observable covariates can be written as $\mathbf{x} = (\mathbf{w}^\intercal, \mathbf{z}^\intercal)^\intercal$, where $\mathbf{z}$ represents those excluded covariates that only affect choices through the structural index. Let W and Z denote the supports of $\mathbf{w}$ and $\mathbf{z}$, respectively. The following assumption requires that $h_0(\mathbf{e}, \mathbf{x})$ does not depend on $\mathbf{z}$, given $\mathbf{u}_0$ and $\mathbf{w}$.

Assumption 3 (Exclusion restriction) The distribution of play $h_0(\mathbf{e}, \mathbf{x})$ is measurable with respect to the σ-algebra generated by $(\mathbf{u}_0, \mathbf{w})$, i.e., there exists a measurable function $\tilde{h}_0 : U \times W \to \Delta(Y)$ such that $h_0(\mathbf{e}, \mathbf{x}) = \tilde{h}_0(\mathbf{u}_0, \mathbf{w})$ a.s..

In terms of the underlying game, and having a fixed solution concept in mind, Assumption 3 imposes some restrictions on the way agents choose between different equilibria. Namely, it requires the *average* of the selected equilibria to be independent of $\mathbf{e}$ and $\mathbf{z}$ conditional of $\mathbf{w}$ and $\mathbf{u}_0$. This is the only assumption that we impose on selection criteria. Also, by only assuming conditional independence, we allow for the possibility that $\mathbf{z}$ could affect the selection, but only through $\mathbf{u}_0$ or $\mathbf{w}$. This assumption is implied by commonly used assumptions.[8] For instance, Bajari et al. (2010) allows equilibrium selection to be random, but the probability of choosing each equilibrium is fully parametric and depends only on payoffs. In turn, other papers make the much stronger assumption that players always choose the same equilibrium.

The next assumption is that the collection of distributions of $u(\beta, \mathbf{e}, \mathbf{x})$ conditional on different realizations of $\mathbf{z}$ constitutes a boundedly complete family. See Andrews (2011) for more details on boundedly complete families.

[8] In fact, Henry and Mourifie (2012) shows that non-parametric point identification of payoffs assuming Nash equilibrium is not possible in 2×2 games without imposing restrictions on the selection criteria.

Assumption 4 (Bounded completeness) For every measurable and bounded function $g : U \times W \to [-1, 1]$, if $\mathbb{E}[\, g(\mathbf{u}_0, \mathbf{w})|\mathbf{w}, \mathbf{z}\,] = 0$ a.s., then $g(\mathbf{u}_0, \mathbf{w}) = 0$ a.s..

Assumption 4 is implied by Assumption 4' below, which is in turn satisfied by all models with additive errors with normal or extreme-valued distributions, and also by more general specifications commonly used both in applied work (e.g., Ciliberto and Tamer (2009)), and in numerical simulations from theoretical work (e.g., Kline (2016)).

Assumption 4' There exists a random variable $\mathbf{v}$ with support $V \subseteq \mathbb{R}^{d_v}$, and known functions $\tilde{u} : B \times V \times W \to U$ and $m : Z \to \mathbb{R}^{d_v}$ such that (i) $\mathbf{u}_0 = \tilde{u}(\beta_0, \mathbf{v}, \mathbf{w})$ a.s., (ii) the distribution of $\mathbf{v}$ conditional on $\mathbf{x}$ belongs to the exponential family with $m(\mathbf{z})$ as a parameter, and (iii) the support of $m(\mathbf{z})$ conditional on $\mathbf{w}$ contains an open set a.s..

Condition (i) means that $\mathbf{v}$ is a sufficient statistic for $\mathbf{u}_0$ given $\mathbf{w}$. Conditions (ii) and (iii) are standard assumptions which imply that, given β and $\mathbf{w}$, the collection of distributions of $\mathbf{v}$ conditional on different realizations of $\mathbf{z}$ constitutes a complete family of distributions (see Theorem 2.12 in Brown (1986)). Since $\mathbf{u}_0$ is measurable with respect to $\mathbf{v}$, the collection of conditional distributions of $\mathbf{u}_0$ is also complete. Therefore, Assumption 4' implies Assumption 4.

There are two crucial restrictions imposed by Assumption 4'. First, the distribution of the structural errors should belong to the exponential family (or another boundedly complete family). Second, there must be continuous error-specific covariates that shift the effect of each error term. It is very flexible in terms of the functional forms that satisfy it. It is satisfied by our entry game (see Section 4.1), and by pricing games both with linear and logistic demand systems as the ones studied in Berry et al. (1995). It is also satisfied whenever $\mathbf{v}$ is a linear function of $\mathbf{e}$ and $\mathbf{x}$, and the covariate space has at least the same dimension as the error space. Therefore, it incorporates a large class of multiple-index models. Consider for instance the following example based on the classical teamwork model from Hölmstrom (1982).

Example 4.1 Partners $i \in I = \{1, \ldots, n_i\}$ participate in a joint-venture. Each partner provides a level of non-contractible effort $y_i \in \{0, 1, \ldots, n_y\}$. The private cost $\mathbf{c}_i$ and productivity $\mathbf{a}_i$ for agent i are given by $\mathbf{c}_i = c(\beta_0, \mathbf{w}, \beta_{0ic}\mathbf{z}_{ic} + \tilde{\mathbf{e}}_{ic})$ and

$\mathbf{a}_i = a(\beta_0, \mathbf{w}, \beta_{0ia}\mathbf{z}_{ia} + \tilde{\mathbf{e}}_{ia})$, where $(\tilde{\mathbf{e}}_{ic}, \tilde{\mathbf{e}}_{ia})$ are agent specific shocks, $(\mathbf{z}_{ic}, \mathbf{z}_{ia})$ are excluded covariates, and c and a are known functions. *Per capita* output is given by a known function π of $\mathbf{a} = (\mathbf{a}_i)_{i\in I}$ and $y = (y_i)_{i\in I}$. Output is shared equally so that the realized utility function for player i is $\mathbf{u}_{0i}(y) = \pi(\mathbf{a}, y) - y_i\mathbf{c}_i$.

Assume that the structural agent-specific shocks $(\tilde{\mathbf{e}}_{ic}, \tilde{\mathbf{e}}_{ia})_{i\in I}$ are i.i.d. according to the Gumbell distribution with parameter $(0, 1)$, that is, the c.d.f. of each $\tilde{\mathbf{e}}_{ij}$ is $F_{\tilde{\mathbf{e}}_{ij}}(\tilde{e}_{ij}) = \exp(-\exp(-\tilde{e}_{ij}))$. Let $\mathbf{v} = (\beta_{0ic}\mathbf{z}_{ic} + \tilde{\mathbf{e}}_{ic}, \beta_{0ia}\mathbf{z}_{ia} + \tilde{\mathbf{e}}_{ia})_{i\in I}$ and $\mathbf{z} = (\mathbf{z}_{ic}, \mathbf{z}_{ia})_{i\in I}$. Clearly, $\mathbf{u}_0$ can be expressed as a known function of $(\mathbf{w}, \mathbf{v})$. Furthermore, the family of distributions of $\mathbf{v}$ conditional on $\mathbf{x} = (\mathbf{w}, \mathbf{z})$ belongs to the exponential family with $m(\mathbf{z}) = (\beta_{0ic}\mathbf{z}_{ic}, \beta_{0ia}\mathbf{z}_{ia})_{i\in I}$ as a parameter.[9] Hence, Assumption 4' is satisfied as long as the support of $\mathbf{z}$ conditional on $\mathbf{w}$ contains an open ball and $\beta_{0ic}, \beta_{0ia} \neq 0$ for all i.

Assumptions 3 and 4 are sufficient to guarantee that the true distribution of play is point identified up to the vector of unknown structural parameters β_0.

Proposition 4.1 *Under assumptions 1–4, if (β, h) and (β, h') belong to the sharp identified set, then $h(\mathbf{e}, \mathbf{x}) = h'(\mathbf{e}, \mathbf{x})$ a.s..*

Proof. Let (β, h) and (β, h') be consistent with the data and satisfy the assumptions of the proposition. The consistency condition (2) implies that

$$\mathbb{E}[\, h(\mathbf{e}, \mathbf{x}) - h'(\mathbf{e}, \mathbf{x})|\mathbf{x}\,] = \mu_0(\mathbf{x}) - \mu_0(\mathbf{x}) = 0 \quad \text{a.s..}$$

By Assumption 3, there exist functions $\tilde{h}, \tilde{h}' : U \times W \to \Delta(Y)$ such that $h(\mathbf{e}, \mathbf{x}) = \tilde{h}(\mathbf{u}_\beta, \mathbf{w})$ and $h'(\mathbf{e}, \mathbf{x}) = \tilde{h}'(\mathbf{u}_\beta, \mathbf{w})$ a.s., where $\mathbf{u}_\beta = u(\beta, \mathbf{e}, \mathbf{x})$. Therefore,

$$\mathbb{E}\left[\, \tilde{h}(\mathbf{u}_\beta, \mathbf{w}) - \tilde{h}'(\mathbf{u}_\beta, \mathbf{w})|\mathbf{x} \right] = \mathbb{E}[\, \mathbf{h} - \mathbf{h}'|\mathbf{x}\,] = 0 \quad \text{a.s..}$$

[9] Let $\tilde{v}_i = \beta_{0ic}z_{ic} + \tilde{e}_{ic}$, $\tilde{v}_{2i} = \beta_{0ia}z_{ia} + \tilde{e}_{ic}$, $\tilde{m}_i = \beta_{0ic}z_{ic}$ and $\tilde{m}_{2i} = \beta_{0ia}z_{ia}$ for $i = 1, \ldots, \iota$, so that $v = (\tilde{v}_i, \tilde{v}_{2i})_{i\in I}$ and $m(z) = (\tilde{m}_i, \tilde{m}_{2i})_{i\in I}$. Then, the p.d.f. of $\mathbf{v}$ conditional on $\mathbf{x}$ is:

$$\begin{aligned} f_{\mathbf{v}|\mathbf{x}}(v|x) &= \prod_{i=1}^{\iota} f_{\tilde{\mathbf{e}}_{ic}}(v_{ic} - \beta_{0ic}z_{ic}) f_{\tilde{\mathbf{e}}_{ia}}(v_{ia} - \beta_{0ia}z_{ia}) \\ &= \Psi(v) \exp\left(\sum_{i=1}^{2\iota} \eta_i(m(z))T_i(v) - \phi(m(z))\right), \end{aligned}$$

where $\Psi(v) = \exp\left(-\sum_{i=1}^{\iota}(\tilde{v}_i + \tilde{v}_{2i})\right)$, $\phi(m(z)) = \sum_{i=1}^{\iota}(\tilde{m}_i + \tilde{m}_{2i})$, and $T_i(v) = \exp(-\tilde{v}_i)$, $T_{2i}(v) = \exp(-\tilde{v}_{2i})$, $\eta_i(m(z)) = \tilde{m}_i$ and $\eta_{2i}(m(z)) = \tilde{m}_{2i}$, for $i = 1, \ldots, \iota$.

Because $\mathbf{u}_\beta$ satisfies Assumption 4, it follows that $\tilde{h}(\mathbf{u}_\beta, \mathbf{w}) = \tilde{h}'(\mathbf{u}_\beta, \mathbf{w})$ a.s., and, consequently, $h(\mathbf{e}, \mathbf{x}) = h'(\mathbf{e}, \mathbf{x})$ a.s.. ■

4.1. Distribution of play in the entry game

Now let us consider the problem of identifying the distribution of play in our entry model with normally distributed errors. For now, assume that profit functions are common knowledge, and that the true solution concept is measurable with respect to $\mathbf{u}_0$. Under these conditions, Assumption 3 is satisfied as long as the way firms choose between equilibria depends only on the firms' preferences.

There could be some observable covariates that affect the selection criteria beyond their direct effect on payoffs. These may include, for example, the relative size of the firms, the relationship of the firms with the market (domestic vs. foreign), or the experience of the firms' managers. Fortunately, Assumption 3 does not require that the distribution of play should be unaffected by *all* covariates. It could still be satisfied as long as there exist *some* excluded covariates which generate enough variation in $\mathbf{u}_0$.

As for Assumption 4, $\mathbf{u}_0$ can be written as $\mathbf{u}_{0i}(y) = (\mathbf{v}_i - \beta_{03} y_{-i}) y_i$, where $\mathbf{v}_i = \beta_{0i}\mathbf{x}_i - \tilde{\mathbf{e}}_i$. Note that

$$\mathbf{v}|\mathbf{x} = x \sim N\left(\begin{pmatrix}\beta_{10}x_1 \\ \beta_{20}x_2\end{pmatrix}, \begin{pmatrix}1 & \rho_0 \\ \rho_0 & 1\end{pmatrix}\right).$$

Since normal distributions belong to the exponential family, Assumption 4' is satisfied. Hence, the true distribution of play is point identified up to β_0.

5. Point identification of structural parameters

We impose a high-level assumption that is sufficient for point identification of β_0. The condition is also necessary for point identification of β_0 under our exclusion restriction. It encompasses standard approaches used elsewhere in the literature, including, for instance, the identification-at-infinity approaches from

Tamer (2003) and Bajari et al. (2010).

Assumption 5 (Identifying single predictions) There exists a known family of pairs $(h_k, (X_k(t))_{t\geq 0})_{k\in[0,1]}$, with $h_k : B \times E \times X \to \Delta(Y)$ measurable, and $X_k(t) \subseteq X$ with $\Pr(X_k(t)) > 0$, such that for every k:

$$\lim_{t\to\infty} \Pr\Big(h_0(\mathbf{e},\mathbf{x}) = h_k(\beta_0,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big) = 1, \tag{5}$$

and for each $\beta \in B$ with $\beta \neq \beta_0$ there exists some k^* such that:

$$\lim_{t\to\infty} \mathbb{E}\Big[h_{k^*}(\beta_0,\mathbf{e},\mathbf{x}) - h_{k^*}(\beta,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_{k^*}(t)\Big] \neq 0. \tag{6}$$

Condition (5) requires that, for some (limiting) regions of the covariates' support, the conditional probabilities of outcomes are known up the structural parameter. Condition (6) then requires that these probabilities identify β_0.

Proposition 5.1 *Under assumptions 1–3, Assumption 5 is sufficient and necessary to point identify β_0.*

Proof. Sufficiency — If (β, h) belongs to the identified set under Assumption 5, Condition (5) implies that for every k:

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}\Big[h_k(\beta,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big] &= \lim_{t\to\infty} \mathbb{E}\Big[h(\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big] \\
&= \lim_{t\to\infty} \mathbb{E}\Big[h_0(\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big] \\
&= \lim_{t\to\infty} \mathbb{E}\Big[h_k(\beta_0,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big],
\end{aligned}$$

where the second equality follows from (2). Moreover, if Assumption (5) is satisfied and $\beta \neq \beta_0$, then, by Condition (6), there would exist some k^* such that

$$\lim_{t\to\infty} \mathbb{E}\Big[h_{k^*}(\beta_0,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big] \neq \lim_{t\to\infty} \mathbb{E}\Big[h_{k^*}(\beta,\mathbf{e},\mathbf{x}) \Big| \mathbf{x} \in X_k(t)\Big].$$

Therefore, if (h, β) belongs to the sharp identified set under assumptions 1–3 and 5, then $\beta = \beta_0$.

Necessity — Suppose that β_0 is point identified under assumptions 1–3, and take any $\beta \in B$, $\beta \neq \beta_0$. By Assumption 3, there exists some $\tilde{h}_0 : U \times W \to \Delta(Y)$

such that

$$\mu_0(\mathbf{x}) = \mathbb{E}\left[\tilde{h}_0(\mathbf{u}_0, \mathbf{w}) \,\middle|\, \mathbf{x} \right] \text{a.s..}$$

Since we are not assuming bounded completeness (Assumption 4), there might be more than one function satisfying this condition. Any such function suffices for our purposes. Moreover, since β_0 is point identified and μ_0 is observed, we can treat $\tilde{h}_0$ as known.

Let $h_{k^*}(\beta, e, x) = \tilde{h}_0(u(\beta, e, x), w)$. Because β does not belong to the sharp identified set under 1–3, it must be the case that

$$\mu_0(\mathbf{x}) \neq \mathbb{E}\left[\tilde{h}_0(u(\beta, \mathbf{e}, \mathbf{x}), \mathbf{w}) \,\middle|\, \mathbf{x} \right]$$

with positive probability. Hence, there exists some set $X_{k^*} \subseteq X$ with $\Pr(X_{k^*}) > 0$ and such that

$$\begin{aligned} \mathbb{E}\left[h_{k^*}(\beta, \mathbf{e}, \mathbf{x}) \,\middle|\, \mathbf{x} \in X_{k^*} \right] &= \mathbb{E}\left[\tilde{h}_0(u(\beta, \mathbf{e}, \mathbf{x}), \mathbf{w}) \,\middle|\, \mathbf{x} \in X_{k^*} \right] \\ &\neq \mathbb{E}\left[\tilde{h}_0(\mathbf{u}_0, \mathbf{w}) \,\middle|\, \mathbf{x} \in X_{k^*} \right] \\ &= \mathbb{E}\left[h_{k^*}(\beta_0, \mathbf{e}, \mathbf{x}) \,\middle|\, \mathbf{x} \in X_{k^*} \right]. \end{aligned}$$

■

Note that we do not rely on bounded completeness to identify structural parameters. Also, the proof uses our exclusion restriction for establishing necessity, but not for sufficiency. Assumptions 1, 2 and 5 are sufficient to point identify β_0. The rest of this section shows how payoff parameters can be point identified in different specific settings.

5.1. Identification at infinity in the entry game

Consider the entry game from Section 2 with normally distributed structural errors. Assuming rationalizable behavior and a smoothness condition, the structural parameters are identified at infinity. To the best of our knowledge, this is the first formal result establishing identification of the correlation parameter under rationalizability.

Proposition 5.2 *Under the following assumptions*

(i) The structural error satisfies (1), where the correlation $\rho_0 \in (-1, 1)$ is an unknown structural parameter;

(ii) Firms make rationalizable choices, i.e., $h_0(\mathbf{e}, \mathbf{x}) \in q_{\mathrm{R}}(\beta_0, \mathbf{e}, \mathbf{x})$ a.s.;

(iii) h_0 is continuously differentiable with respect to x in the multiplicity region;

the vector of structural parameters for the entry game $\beta_0 = (\beta_{01}, \beta_{02}, \tilde{\beta}_{01}, \tilde{\beta}_{02}, \rho_0)^{\mathsf{T}}$ is point identified.

As usual, the mean utility parameters β_{0i} and $\tilde{\beta}_{0i}$ are identified under rationalizability by the limit of i's entry probabilities when $\mathbf{x}_{-i} \to \infty$ and $\mathbf{x}_{-i} \to -\infty$. See, for instance, Tamer (2003). The correlation parameter ρ_0 requires additional work. We identify it by analyzing the limit of the derivative $\partial[\mu_0(x)](0,0)/\partial x_1$, when x diverges along a specific direction. The details are in Appendix B. The normality assumption can be relaxed to accommodate more general parametric distributions. The result can be easily generalized to entry games with more than two players.

5.2. Cournot oligopoly

The restriction to two firms, two actions, and linear payoff functions is not crucial to identify the structural parameters. To see this, consider the following Cournot competition example. Suppose there are n_I firms, $i \in \{1, \ldots, n_I\}$. Each firm chooses a quantity $y_i \in \{1, 2, \ldots\}$ of an indivisible good to supply to the market. Profits for firm i are given by

$$\mathbf{u}_{0i}(y) = \left(\bar{\beta}_{01} + \beta_{0i}\mathbf{x}_i - \tilde{\beta}_{0i} \sum_{j=1}^{n_I} y_j - \tilde{\mathbf{e}}_i \right) y_i,$$

where $\bar{\beta}_{0i}, \beta_{0i}, \tilde{\beta}_{0i} > 0$, and $\tilde{\mathbf{e}}_i = F_i^{-1}(\mathbf{e}_i)$ for some known c.d.f. $F_i(\,\cdot\,)$. Note that, when $\mathbf{x}_i \to -\infty$, there is a high probability that the only rational choice for firm i is to choose $y_i = 0$. When this is the case for all firms $i \neq 1$, firm 1 faces a single-agent decision problem. Using this fact, we can show that

$$\lim_{x_{-i} \to -\infty} \Pr\left(\mathbf{y}_1 = 0 \,\middle|\, \mathbf{x} = x\right) = 1 - F_1\left(\bar{\beta}_{01} + \beta_{01}x_1 - \tilde{\beta}_{01}\right)$$

$$\lim_{x_{-i}\to-\infty} \Pr\left(\mathbf{y}_1 = 1 \,\middle|\, \mathbf{x} = x\right) = F_1\left(\bar{\beta}_{01} + \beta_{01}x_1 - \tilde{\beta}_{01}\right) - F_1\left(\bar{\beta}_{01} + \beta_{01}x_1 - 3\tilde{\beta}_{01}\right)$$

These equations identify $\bar{\beta}_{01}$, β_{01} and $\tilde{\beta}_{01}$. Similarly, one can identify the rest of the parameters.

5.3. Weaker behavioral assumptions

Since the main objective of this work is to identify patterns of behavior, it is desirable to make as few behavioral assumptions as possible. Our two previous examples rely on assuming that choices are rationalizable in each instance of the game. This assumption might be systematically violated in different settings. One possibility is that firms can enforce collusive agreements. In such settings, a firm could agree to forego a profitable entry opportunity to increase the profit of other firms, in exchange for reciprocal behavior in other markets. Another possibility is that economic agents exhibit ambiguity aversion and utilize maxmin strategies that violate subjective-expected-utility rationality. Our general approach can accommodate these departures from rationalizability. Appendix C shows that payoff parameters can be identified under mild behavior assumptions. In particular, we only assume two rounds of elimination of *absolutely dominated* strategies in the sense of Salcedo (2012) and Halpern and Pass (2012).

6. Identification of solution concepts

Recall that a solution concept is consistent with the data if it is satisfied by some (β, h) belonging to the sharp identified set. Under the assumptions from sections 4 and 5, the sharp identified set collapses to (β_0, h_0). Hence, under these assumptions, a solution concept is consistent with the data if and only if it is satisfied by *the true* structural parameters and distribution of play.

Theorem 6.1 *Under assumptions 1–5,* (β_0, h_0) *are point identified, and a solution concept is consistent with the data if and only if it is satisfied by the players' behavior almost surely, i.e., if and only if* $h_0(\mathbf{e}, \mathbf{x}) \in q(\beta_0, \mathbf{e}, \mathbf{x})$ a.s..

Proof. The result is a direct consequence of propositions 4.1 and 5.1. Suppose that q is consistent with the data, i.e., there exist (β, h) in the identified set which satisfy q. By Proposition 5.1, we know that we have $\beta = \beta_0$. Since (β_0, h_0) always belongs to the identified set, it follows from Proposition 4.1 that $h(\mathbf{e}, \mathbf{x}) = h_0(\mathbf{e}, \mathbf{x})$ a.s.. Therefore, $h_0(\mathbf{e}, \mathbf{x}) \in q(\beta_0, \mathbf{e}, \mathbf{x})$ a.s.. ∎

Let us contrast Theorem 6.1 with the characterization from Beresteanu et al. (2011), hereafter BMM. In the context of games, BMM deal with the identification of payoff parameters assuming equilibrium behavior. However, one could add a *convex* solution concept q_0 as part of their parameter space. Then, assuming that β_0 is point identified, Theorem 2.1 in BMM would imply that q_0 is consistent with the data if and only if

$$\mu_0(\mathbf{x}) \in \mathbb{E}\left[\, q_0(\beta_0, \mathbf{e}, \mathbf{x}) | \mathbf{x} \,\right] \quad \text{a.s.,} \tag{7}$$

where $\mathbb{E}[\, q_0(\beta, \mathbf{e}, \mathbf{x}) | \mathbf{x} \,]$ denotes the conditional Aumann expectation of $q_0(\beta, \mathbf{e}, \mathbf{x})$.

The expectation in (7) is taken over the unobserved heterogeneity embedded in $\mathbf{e}$. Hence, according to BMM's characterization, a solution concept is consistent as long as its restrictions are satisfied *on average*, conditioning only on $\mathbf{x}$. In contrast, under the assumptions of Theorem 6.1, the actual distribution of play is point identified and the restrictions should be satisfied *ex post*, conditioning on both $\mathbf{x}$ and $\mathbf{e}$. This means that a solution concept is consistent with the data if and only if the agents choices satisfy it *almost surely*.

Moreover, since condition (7) is expressed in terms of Aumann expectations, it cannot discriminate between solution concepts whose Aumann expectations coincide. For example, it cannot discriminate between q_{NE} and q_{PRNE}, as defined in Example 3.2. Suppose for instance that firms in the entry game behave according to the distribution of play h_{UNE} from Example 3.1. Since h_{UNE} is consistent with q_{PRNE}, and q_{PRNE} and q_{NE} have the same Aumann expectation, it follows that μ_0, β_0 and q_{NE} jointly satisfy (7). However, Theorem 6.1 implies that, in this case, q_{NE} would not be consistent with the observed data. For non-convex solution concepts, condition (7) is necessary but not sufficient for consistency with the observed data.

6.1. Point identification of the solution concept

Theorem 6.1 can be used to establish point identification of the true solution concept from a set of competing alternatives.

Assumption 6 (β_0, h_0) jointly satisfy at least one "true" solution concept q_0 from a known set Q.

Say that the solution concept is point identified if q_0 is the only solution concept in Q which is consistent with the data. By Theorem 6.1, q is consistent with the data only if $q(\beta_0, \mathbf{e}, \mathbf{x})$ contains the distribution of play a.s.. By Assumption 6, the distribution of play is in $q_0(\beta_0, \mathbf{e}, \mathbf{x})$ a.s.. Hence, as long as $q(\beta_0, \mathbf{e}, \mathbf{x})$ and $q_0(\beta_0, \mathbf{e}, \mathbf{x})$ are disjoint with positive probability, q cannot be consistent with the data. Point identification is thus attained whenever every pair of competing solution concepts make disjoint predictions with positive probability.

Corollary 6.2 *Under assumptions 1–6, if for every $q \in Q \setminus \{q_0\}$ there exists some $F \subseteq E \times X$ such that $\Pr((\mathbf{e}, \mathbf{x}) \in F) > 0$ and $q(\beta_0, e, x) \cap q_0(\beta_0, e, x) = \emptyset$ for almost all $(e, x) \in F$, then (β_0, h_0, q_0) is point identified.*

Without imposing assumptions 3–5, two solution concepts q and q' could be consistent with the observed data, even if they make disjoint predictions with positive probability. There could exists pairs (h, β) and (h', β') which yield the same distribution over observables, and are consistent with q and q', respectively. The content of the corollary is that, in that case, at least one of these pairs would not be consistent with assumptions 3–5.

The condition from the corollary cannot help with nested solution concepts q, q' with $q(e, x) \subseteq q'(e, x)$ for all $(e, x) \in E \times X$. However, the condition is sufficient but not necessary. Theorem 6.1 implies that the smaller solution concept could be ruled out if $h_0(\mathbf{e}, \mathbf{x}) \in q'(\beta_0, \mathbf{e}, \mathbf{x}) \setminus q(\beta_0, \mathbf{e}, \mathbf{x})$ with positive probability. Similarly, it is easy to construct examples in which it is possible to rule out the difference $q' \setminus q$.

6.2. Identification of solution concepts in the entry game

One of the most commonly used behavior assumptions is that of Nash equilibrium. It is tantamount to assuming that (*i*) each firm maximizes its profits given its beliefs, and (*ii*) each firm holds correct beliefs about its opponents (Aumann and Brandenburger, 2016). Our analysis implies that, assuming q_{R} and an exclusion restriction in the entry game, it is possible to test whether q_{NE} is satisfied. That is, assuming common knowledge profit maximization, it is possible to test whether firms actions are best responses to *correct* equilibrium beliefs within the framework of our model. A testing procedure for this hypothesis can be found in Kashaev (2016).

Moreover, point identifying the distribution of play helps to answer a wide range of questions such as which equilbria are more likely to be played, whether the players use mixed strategies, or whether choices are made sequentially or simultaneously.

Example 6.1 Suppose that firms in the entry game play subgame perfect Nash equilibria (SPNE), but the researcher does not know whether the firms choose their actions simultaneously or sequentially, or which firm moves first. To keep things simple, we consider only three solution concepts: q_{PRNE} corresponds to the simultaneous move game, and q_i restricts behavior to be SPNE of the game in which firm i always moves first, for $i = 1, 2$.

Our previous arguments to establish identification of β_0 and h_0 still apply. Hence, in order to point identify $q_0 \in \{q_{\mathrm{PRNE}}, \bar{q}_1, \bar{q}_2\}$, it suffices to establish the sufficient condition from Corollary 6.2. Clearly, q_1 and q_2 satisfy the condition because q_i predicts that i is the only entrant in the multiplicity region from Figure 1. However, without further assumptions, Corollary 6.2 does not help to distinguish between q_{PRNE} and q_i because they are nested for this particular game. This point has been raised, for instance, by Bresnahan and Reiss (1990).

We overcome this issue by assuming that there is a different covariate $\mathbf{z}_1$ which takes both positive and negative values with positive probability and such that:

$$\mathbf{u}_{01}(y) = (\beta_{01}\mathbf{x}_1 - \beta_{03}\mathbf{z}_1 y_2 - \tilde{\mathbf{e}}_1) y_1.$$

That is, firm 1 sometimes benefits from having firm 2 in the market. This assumption can be justified considering asymmetric retailers. Suppose that firm 2 is a large departmental store with a well renowned brand, and firm 1 is a small local firm. Firm 1 may benefit from having firm 2 nearby, as firm 2 may attract a large costumer flow, while firm 2 may still prefer to be a monopolist.

With this new covariate, it suffices to consider the set:

$$\left\{(x, \tilde{e}, z_1) \in \mathbb{R}^5 \,\middle|\, 0 < \beta_2 x_2 - \tilde{e}_2 < \beta_3 \quad \wedge \quad \beta_3 z_1 < \beta_1 x_1 - \tilde{e}_1 < 0\right\}.$$

In this region, the simultaneous move game only admits an asymmetric mixed strategy equilibrium, while the sequential games only admit pure strategy equilibria, and we thus have $q_{\text{PRNE}} \cap q_i = \emptyset$. Therefore, by Corollary 6.2, all the equilibrium concepts are distinguishable, and q_0 is point identified.

7. Incomplete information games

In this section, we identify one of two different information structures for our entry model. We assume that errors are normally distributed and independent across firms and that the firms' entry choices satisfy static equilibrium conditions, and we allow for public randomization. We discriminate the assumption that payoffs are common knowledge, versus the alternative hypothesis that each firm i only observes $\mathbf{x}$ and $\mathbf{e}_i$, but not $\mathbf{e}_{-i}$.[10]

Under the incomplete-information hypothesis, a pure strategy for i is a function $s_i : E_i \to \{0, 1\}$. The game only admits pure strategy BNE, and a strategy profile s^* is a BNE if and only if it satisfies

$$s_i^*(e_i) = \begin{cases} 0 & \text{if} \quad e_i > \bar{e}_i(x) \\ 1 & \text{if} \quad e_i < \bar{e}_i(x) \end{cases}, \tag{8}$$

for some $\bar{e}(x) = (\bar{e}_1, \bar{e}_2)$ such that

$$\bar{e}_i(x) = \beta_{0i} x_i - \beta_{03} \Phi(\bar{e}_{-i}(x)), \tag{9}$$

[10] This application is closely related to the work of Grieco (2014) and Magnolfi and Roncoroni (2016), who construct estimates for structural parameters in an entry environment that are robust to different informational assumptions. In contrast, we focus on the possibility to discern between different information structures.

for $i = 1, 2$. There always exists at least one such equilibrium, and there are multiple equilibria for some values of x. See Appendix D for more details. Let q_{BNE} denote the (convex hull) of the set of Bayes Nash Equilibria (BNE) of the incomplete information game.

Our identification at-infinity-approach works under both q_{PRNE} and q_{BNE}, and thus β_0 is point identified. However, Theorem 6.1 cannot be directly applied to this problem, because of our exclusion restriction (Assumption 3). For a solution concept q to be consistent with this restriction, it has to admit a selection that is measurable with respect to $(\mathbf{u}_0, \mathbf{w})$. Our incomplete information entry game does not admit any such selection, because q_{BNE} changes depending on the public information contained in $\mathbf{x}$. To see this, let (x, e) and (x', e') be such that $x'_1 > x_1$, and:

$$e'_i = e_i + \frac{x_1 - x'_1}{\beta_{10}}.$$

Then, the realized vNM indexes functions are the same, but $\bar{e}(x)$ and $\bar{e}(x')$ differ. Despite this difficulty, we can still discriminate between q_{PRNE} and q_{BNE}.

Proposition 7.1 *(β_0, q_0) is point identified in the entry game under the assumptions that the structural error satisfies (1) with $\rho = 0$, and (β_0, h_0) satisfies either q_{BNE} or q_{PRNE}.*

Proof. Lemma D.1 in Appendix D shows that, as x diverges to infinity along a specific path, the maximum probability of a duopoly under any NE of the complete information game is strictly less than the minimum probability under any BNE of the incomplete information game. This implies that there exists a set of $X' \subseteq X$ with $\Pr(X') > 0$ and such that

$$\mathbb{E}\left[q_{\text{PRNE}}(\beta_0, \mathbf{e}, \mathbf{x}) \middle| X' \right] \cap \mathbb{E}\left[q_{\text{BNE}}(\beta_0, \mathbf{e}, \mathbf{x}) \middle| X' \right] = \emptyset. \tag{10}$$

Consequently, condition (7) can either be satisfied by q_{PRNE} or q_{BNE}, but not by both. Since (7) is a necessary condition for a solution concept to be consistent with the data, this implies that we can discriminate between q_{PRNE} and q_{BNE}. ■

While the equilibrium concepts can be identified in this case, the derivation is far from trivial even when considering only two competing information structures for a simple 2×2 game. In particular, it requires a full characterization of the set of equilibria for all possible realizations of the residuals, and integrating over a selection of extreme points of these sets. Part of the value of Proposition 4.1, is that it allows one to establish point identification in a straightforward way that does not involve computing Aumann expectations. This suggests the need to further investigate a more practical approach for models with incomplete information.

8. Conclusion

The current work analyzes discrete games using a general framework that incorporates the solution concept as a parameter of interest, and only requires mild assumptions about behavior of players. We establish point identification of the distribution over the outcomes conditional on both observed and unobserved payoff shifters. This allows one to conduct counterfactual analysis to derive valid policy implications. Also, it allows one to identify whether a solution concept—convex or not—is consistent with the observed data. For example, it is possible to test whether firms' entry choices satisfy NE conditions, or which equilibria are more likely to arise in coordination problems. We also generalize the identification-at-infinity-approach to point identify payoff parameters. Moreover, we show that existence of known-single-predictions regions is not only sufficient, but also necessary for point identification of the distribution of payoffs under our exclusion restriction.

Our results apply to a general class of discrete complete-information games both in strategic and extensive form. In its current state, our methodology cannot be directly applied to general incomplete-information games, and relies on semi-parametric assumptions. Incomplete information could be dealt with, if the researcher observes covariates that were private information at the moment the game was played, but became public after choices where made. As for the parametric restrictions, it is hard to conceive a full non-parametric approach, but it may be feasible to allow for much greater flexibility by introducing random coefficients. We leave these as open problems for future research.

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.

Aumann, R. J. and Brandenburger, A. (2016). Epistemic conditions for nash equilibrium. In Arló-Costa, H., Hendricks, V. F., and Benthem, J. V., editors, *Readings in Formal Epistemology*, volume 1 of *Graduate Texts in Philosophy*, pages 863–894. Springer.

Bajari, P., Hahn, J., Hong, H., and Ridder, G. (2011). A note on semiparametric estimation of finite mixtures of discrete choice models with application to game theoretic models. *International Economic Review*, 52(3):807–824.

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

Battalio, R., Samuelson, L., and Huyck, J. V. (2001). Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica*, 69(3):749–764.

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

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

Berry, S. T., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Berry, S. T. and Tamer, E. (2006). Identification in models of oligopoly entry. In Blundell, R., Newey, W. K., and Persson, T., editors, *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress, Volume 2*, volume 42 of *Econometric Society Monographs*, pages 46–85. Cambridge University Press.

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.

Bresnahan, T. F. and Reiss, P. C. (1991). Empirical models of discrete games. *Journal of Econometrics*, 48(1–2):57–81.

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.

Camerer, C. (2003). *Behavioral Game Theory: experiments in strategic interaction*. Princeton University Press.

Carlsson, H. and Van Damme, E. (1993). Global games and equilibrium selection. *Econometrica*, 61(5):989–1018.

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. 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.

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

Fox, J. T. (2007). Semiparametric estimation of multinomial discrete-choice models using a subset of choices. *RAND Journal of Economics*, 38(4):1002–1019.

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

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.

Grieco, P. L. E. (2014). Discrete games with flexible information structures: an application to local grocery markets. *RAND Journal of Economics*, 45(2):303–340.

Halpern, J. Y. and Pass, R. (2012). Game theory with translucent players. Cornell Univeristy, Mimeo.

Harsanyi, J. C. and Selten, R. (1988). *A general theory of equilibrium selection in games*. The MIT Press.

Henry, M. and Mourifie, I. (2012). Nonparametric sharp bounds for payoff functions in 2x2 games. Mimeo.

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

Hölmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, 13(2):324–340.

Jun, S. J. and Pinkse, J. (2016). Counterfactual prediction in complete information games: point prediction under partial identification. Mimeo.

Kashaev, N. (2016). Testing for nash behavior in binary games of complete information. Mimeo.

Kline, B. (2015a). An empirical model of non-equilibrium behavior in games. Mimeo.

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

Kline, B. (2016). The empirical content of games with bounded regressors. *Quantitative Economics*, 7(1):37–81.

Magnolfi, L. and Roncoroni, C. (2016). Estimation of discrete games with weak assumptions on information. mimeo.

Molchanov, I. (2006). *Theory of random sets*. Springer.

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.

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

Palacios-Huerta, I. (2003). Professionals play minimax. *The Review of Economic Studies*, 70(2):395–415.

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

Salcedo, B. (2012). Implementation without commitment in moral hazard environments. Mimeo.

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

A. Alternative entry subsidy

The policy instrument discussed in Section 2.2 provides a very simple example of how assuming an incorrect solution concept might yield completely misleading policy implications. Similar conclusions can be drawn with respect to different forms of subsidy, but the analysis of the model becomes more cumbersome. For example, a form of subsidy that is more common in real life consists of giving a lump sum subsidy $\delta^* > 0$ to *any* firm that enters a market with some observable characteristics (see, e.g., Goolsbee (2002)).

Again, 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, once again, the story is very different under the true behavior of the firms.

Proposition A.1 *Suppose that firms make entry decisions in accordance with q' as defined in Section 2.2. If $\delta^* < \tilde{\beta}_{0i}$, $i = 1, 2$, then for every realization of the covariates x, there exist an open set $E(x)$ such that, if $\tilde{e} \in E(x)$, then the market would be served only if there is no subsidy δ^*. Moreover, if $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$ are i.i.d. with c.d.f. F and p.d.f. f such that $F(0) \geq 1/2$ and $f(0) > 0$, then there exist thresholds $\bar{x}, \bar{\delta}^*, \underline{\tilde{\beta}}_{0i} > 0$, such that if $\|x\| \leq \bar{x}$, $\delta^* \leq \bar{\delta}^*$ and $\tilde{\beta}_{0i} \geq \underline{\tilde{\beta}}_{0i}$ for $i = 1, 2$, then the subsidy reduces the probability that a market is served conditional on $\mathbf{x} = x$.*

Proof. Let $E_i(x)$ be the set given by

$$E_i(x) = \left\{ \tilde{e} \;\middle|\; \begin{array}{c} \beta_{0i} x_i \leq \tilde{e}_i \leq \beta_{0i} x_i + \delta^* \\ \text{and} \quad \beta_{0-i} x_{-i} - \tilde{\beta}_{0-i} + \delta^* \leq \tilde{e}_{-i} \leq \beta_{0-i} x_{-i} \end{array} \right\},$$

for $i = 1, 2$. The lower bound for $\tilde{e}_i$ and the upper bound on $\tilde{e}_{-i}$ require that, without the subsidy, $(\tilde{e}, x)$ belongs to the region where a firm-i monopoly is the only rationalizable outcome, and hence the market would be served under q'. The remaining inequalities require that, with the subsidy, $(\tilde{e}, x)$ belongs to the multiplicity region and hence the market will *not* be served under q'. The set $E(x) = E_1(x) \cup E_2(x)$ is exactly the set of markets for which there would be served without the subsidy but not with the subsidy.

Let $g(\beta, \delta, x)$ denote the probability that a market is not served given the subsidy δ^* and parameters β, conditional on $\mathbf{x} = x$. Under q', g is given by the probability of the no-entry region plus the probability of the multiplicity region as shown in Figure 1, that is

$$g(\beta, \delta^*, x) = \Big(1 - F(\beta_1 x_1 + \delta^*)\Big)\Big(1 - F(\beta_2 x_2 + \delta^*)\Big)$$

$$+\Big(F(\beta_1 x_1+\delta^*)-F(\beta_1 x_1+\delta^*-\tilde\beta_1)\Big)$$
$$\cdot\Big(F(\beta_2 x_2+\delta^*)-F(\beta_2 x_2+\delta^*-\tilde\beta_2)\Big).$$

Differentiating with respect to δ^* and evaluating at $x=(0,0)^\mathsf{T}$ yields

$$\frac{\partial g}{\partial\delta^*}\bigg|_{x=(0,0)^\mathsf{T}} = -2\Big(1-F(\delta^*)\Big)f(\delta^*)+\Big(F(\delta^*)-F(\delta^*-\tilde\beta_1)\Big)\Big(f(\delta^*)-f(\delta^*-\tilde\beta_2)\Big)$$
$$+\Big(F(\delta^*)-F(\delta^*-\tilde\beta_2)\Big)\Big(f(\delta^*)-f(\delta^*-\tilde\beta_1)\Big).$$

Taking limits as $\tilde\beta_i\to\infty$, $i=1,2$, yields

$$\lim_{\substack{\tilde\beta_1\to\infty\\ \tilde\beta_2\to\infty}}\frac{\partial g}{\partial\delta^*}\bigg|_{x=(0,0)^\mathsf{T}} = 2f(\delta^*)(2F(\delta^*)-1).$$

The assumptions that $F(0)\geq 1/2$ and $f(0)>0$ imply that there exists some $\bar\delta^*$ such that, if $\delta<\bar\delta^*$, this limit is strictly positive. By continuity of g, this implies that for large values of $\tilde\beta_i$ and small values of $\|x\|$, g is strictly increasing in δ^* around $\delta^*=0$. That is, a small subsidy would increase the probability that the market is *not* served. ■

B. Proof of Proposition 5.2

Fix any (ρ,h) satisfying assumptions (ii) and (iii). Let $G,P:\mathbb{R}^2\to[0,1]$ be the functions given by

$$G(x)=\int_{x_1-\beta_3}^{x_1}\int_{x_2-\beta_3}^{x_2} f_{\tilde{\mathbf{e}}}\left(\tilde e_1,\tilde e_2;\rho\right)[h_0(\beta_0,x,\tilde e)](0,0)\,de_2\,de_1,$$

and

$$P(x):=[\mu(x)](0,0)=G(x)+\int_{x_1}^{\infty}\int_{x_2}^{\infty} f_{\tilde{\mathbf{e}}}\left(\tilde e_1,\tilde e_2;\rho\right)\,de_2\,de_1,$$

where $f_{\tilde{\mathbf{e}}}(\,\cdot\,;\rho)$ denotes the joint density of the structural error terms, given the correlation parameter ρ. In words, $G(x)$ is the probability that $\mathbf{u}$ falls in the multiplicity region *and* firms do not enter the market, and $P(x)$ is the probability that firms do not enter the market, conditional on $\mathbf{x} = x$ and given (ρ, h).

Condition (*iii*) guarantees that G and P are differentiable and

$$\begin{aligned}
\frac{\partial P}{\partial x_1} &= \frac{\partial}{\partial x_1}\left(G(x) + \int_{x_1}^{\infty}\int_{x_2}^{\infty}\frac{1}{\sqrt{1-\rho^2}}\phi(\tilde{e}_1)\phi\left(\frac{\tilde{e}_2 - \rho\tilde{e}_1}{\sqrt{1-\rho^2}}\right)\,d\tilde{e}_2\,d\tilde{e}_1\right)\\
&= \frac{\partial}{\partial x_1}\left(G(x) + \int_{x_1}^{\infty}\phi(\tilde{e}_1)\left[1 - \Phi\left(\frac{x_2 - \rho\tilde{e}_1}{\sqrt{1-\rho^2}}\right)\right]\,d\tilde{e}_1\right)\\
&= \frac{\partial G}{\partial x_1} - \phi(x_1)\left[1 - \Phi\left(\frac{x_2 - \rho x_1}{\sqrt{1-\rho^2}}\right)\right].
\end{aligned}$$

Therefore

$$\frac{1}{\phi(x_1)}\frac{\partial P}{\partial x_1} = \Phi\left(\frac{x_2 - \rho x_1}{\sqrt{1-\rho^2}}\right) - 1 + \frac{1}{\phi(x_1)}\frac{\partial G}{\partial x_1}. \tag{11}$$

This is the equation that we will use to identify ρ_0. In order to do so, we will first find a bound for $|\partial G/\partial x_1|$ that dissipates when x_1 goes to $-\infty$.

Note that

$$\begin{aligned}
\frac{1}{\phi(x_1)}\left|\frac{\partial G}{\partial x_1}\right| = \frac{1}{\phi(x_1)}\Bigg| &\int_{x_2-\beta_3}^{x_2}[h(x, x_1, e_2)](0,0) f_{\tilde{\mathbf{e}}}(x_1, \tilde{e}_2; \rho)\,d\tilde{e}_2\\
&- \int_{x_2-\beta_3}^{x_2}[h(x, x_1 - \beta_3, e_2)](0,0) f_{\tilde{\mathbf{e}}}(x_1 - \beta_3, \tilde{e}_2; \rho)\,d\tilde{e}_2\\
&+ \int_{x_1-\beta_3}^{x_1}\int_{x_2-\beta_3}^{x_2}\frac{\partial}{\partial x_1}[h(x, e)](0,0) f_{\tilde{\mathbf{e}}}(\tilde{e}_1, \tilde{e}_2; \rho)\,d\tilde{e}_2\,d\tilde{e}_1\Bigg|\\
\leq \frac{1}{\phi(x_1)}&\left|\int_{x_2-\beta_3}^{x_2}\phi(x_1) f_{\tilde{\mathbf{e}}_2|\tilde{\mathbf{e}}_1}(\tilde{e}_2|\tilde{e}_1 = x_1; \rho)\,d\tilde{e}_2\right|\\
&+ \frac{1}{\phi(x_1)}\left|\int_{x_2-\beta_3}^{x_2}\phi(x_1 - b_3) f_{\tilde{\mathbf{e}}_2|\tilde{\mathbf{e}}_1}(\tilde{e}_2|\tilde{e}_1 = x_1 - \beta_3; \rho)\,d\tilde{e}_2\right|
\end{aligned}$$

$$+\frac{C}{\phi(x_1)}\left|\int\limits_{x_1-\beta_3}^{x_1}\int\limits_{x_2-\beta_3}^{x_2} f_{\tilde{\mathbf{e}}}(\tilde{e}_1,\tilde{e}_2;\rho)\,d\tilde{e}_2\,d\tilde{e}_1\right|,$$

for some constant $C > 0$. For the inequality we use the triangle inequality, the fact that $h \in [0,1]$, and the fact that $|\partial h/\partial x_1|$ is uniformly bounded in the multiplicity region by some C, because h is continuously differentiable and the multiplicity region is compact. Now we look at each of the terms on the right-hand side separately.

The first term is equal to

$$\int\limits_{x_2-\beta_3}^{x_2} \frac{1}{\sqrt{1-\rho^2}}\,\phi\left(\frac{\tilde{e}_2-\rho x_1}{\sqrt{1-\rho^2}}\right) d\tilde{e}_2 = \Phi\left(\frac{x_2-\rho x_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{x_2-\rho x_1-\beta_3}{\sqrt{1-\rho^2}}\right).$$

The second term can be written as

$$\begin{aligned}\frac{\phi(x_1-\beta_3)}{\phi(x_1)}\,\Pr\left(x_2-\beta_3 < \tilde{\mathbf{e}}_2 < x_2 \,\middle|\, \tilde{\mathbf{e}}_1 = x_1-\beta_3;\rho\right) &< \frac{\phi(x_1-\beta_3)}{\phi(x_1)}\\ &= \exp\left(\beta_3 x_1 - \frac{1}{2}\beta_3^2\right).\end{aligned}$$

And, finally, the third term is

$$\begin{aligned}\frac{C}{\phi(x_1)}\,\Pr\left(x_i-\beta_3 < \tilde{\mathbf{e}}_i < x_i \text{ for } i=1,2;\,\rho\right) &< \frac{C}{\phi(x_1)}\Pr(x_1-\beta_3 < \tilde{\mathbf{e}}_1 < x_1;\rho)\\ &= C\left(\frac{\Phi(x_1)-\Phi(x_1-\beta_3)}{\phi(x_1)}\right).\end{aligned}$$

Hence, we have that

$$\begin{aligned}\frac{1}{\phi(x_1)}\left|\frac{\partial G}{\partial x_1}\right| &< \Phi\left(\frac{x_2-\rho x_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{x_2-\rho x_1-\beta_3}{\sqrt{1-\rho^2}}\right)\\ &\quad \ldots \quad + \exp\left(\beta_3 x_1 - \frac{1}{2}\beta_3^2\right) + C\left(\frac{\Phi(x_1)-\Phi(x_1-\beta_3)}{\phi(x_1)}\right).\end{aligned}$$

Now we take any $\tau \in (-1,1)$, and take limits as $x_1 \to -\infty$ along the line $x_2 = \tau x_1$.

$$\lim_{\substack{x_1\to-\infty\\ x_2=\tau x_2}} \frac{1}{\phi(x_1)}\left|\frac{\partial G}{\partial x_1}\right| < \lim_{x_1\to-\infty}\left[\Phi\left(\frac{(\tau-\rho)x_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{(\tau-\rho)x_1-\beta_3}{\sqrt{1-\rho^2}}\right)\right.$$

$$
\begin{aligned}
\dots \quad &+ \exp\left(\beta_3 x_1 - \frac{1}{2}\beta_3^2\right) + C\left(\frac{\Phi(x_1) - \Phi(x_1 - \beta_3)}{\phi(x_1)}\right)\Bigg] \\
&= \left[\frac{1}{2} - \Phi\left(\frac{-\beta_3}{\sqrt{1-\rho^2}}\right)\right] \mathbb{1}(\tau = \rho) < \frac{1}{2}, \qquad (12)
\end{aligned}
$$

where the limit of the last term disappears because, using L'Hôpital's rule,

$$
\begin{aligned}
\lim_{x_1 \to -\infty} \frac{\Phi(x_1) - \Phi(x_1 - \beta_3)}{\phi(x_1)} &= \lim_{x_1 \to -\infty} \frac{\phi(x_1) - \phi(x_1 - \beta_3)}{-x_1 \phi(x_1)} \\
&= \lim_{x_1 \to -\infty} \frac{1}{x_1}\left(-1 + \exp\left(\beta_3 x_1 - \frac{1}{2}\beta_3^2\right)\right) = 0.
\end{aligned}
$$

Using (11) and (12) together, we have that

$$
\lim_{\substack{x_1 \to -\infty \\ x_2 = \tau x_2}} \frac{1}{\phi(x_1)} \frac{\partial P}{\partial x_1} = \begin{cases} -1 & \text{if } \tau > \rho \\ p \in (-1, 0) & \text{if } \tau = \rho \\ 0 & \text{if } \tau < \rho \end{cases} . \qquad (13)
$$

This condition point identifies ρ_0. To see this, note that if (ρ, h) rationalizes the data, we must have $\mu = \mu_0$ a.s., and therefore $P = P_0$ a.s.. Then, if we take limits as in (13), we must converge to a number in $(-1, 0)$ if and only if $\tau = \rho_0$, which implies that we must have $\rho = \rho_0$.

■

C. Identification at infinity

This appendix provides sufficient conditions for point identification of β_0 using an identification-at-infinity approach. We assume that, for some values of some covariates, some players make choices to solve a given optimization problem with probability approaching one (Assumption 8), and that the corresponding choice models are identified (Assumption 9). Formulating these assumptions requires some additional structure. First, outcomes should correspond to profiles of players' actions.

Assumption 7 Observable outcomes belong to a product space $Y = \times_{i \in I} Y_i$.

Each Y_i corresponds to the set of outcome characteristics that player i controls. In our entry model, we have $Y_i = \{0, 1\}$. Since we are neither specifying a particular game tree nor a particular information structure, elements of Y_i should be interpreted as actions rather than strategies. However, we still can define best response correspondences, *as if* Y_i was a set of strategies in a simultaneous move game. Let

$$\mathrm{BR}_i(y_{-i}; \beta, e, x) = \underset{y_i \in Y_i}{\arg\max} \left\{ [g_i(\beta, e, x)](y_i, y_{-i}) \right\}$$

be the set of actions $y_i \in Y_i$ that would maximize i's preferences given e and x if i's opponents chose y_{-i}, and the value of true structural parameter was β.

Assumption 8 For every player i and each $y_{-i} \in Y_{-i}$, there exists a covariate $\mathbf{x}_k$ such that (i) $\mathbf{u}_{0i}$ and $\mathbf{x}_k$ are independent conditional on $\mathbf{x}_{-k} = (\mathbf{x}_j)_{j \neq k}$, (ii) the support of $\mathbf{x}_{-k}$ is independent of $\mathbf{x}_k$, (iii) the support of $\mathbf{x}_k$ is unbounded from above, and (iv):

$$\lim_{x_k \to \infty} \mathbb{E} \left[[h_0(\mathbf{e}, \mathbf{x})] \Big(\mathrm{BR}_i(y_{-i}; \beta_0, \mathbf{e}, \mathbf{x}) \Big) \,\Big|\, \mathbf{x}_k \geq x_k \right] = 1. \tag{14}$$

Assumption 8 is a joint assumption on the distribution of play and the payoff distribution. The key condition is condition (iv) which requires that, for large values of $\mathbf{x}_k$, player i faces a single-agent decision problem without uncertainty, and makes choices which maximize his utility. It is satisfied in environments satisfying the following two conditions. The first condition is that we can write:

$$\mathbf{u}_{\beta,-i}(y) = \mathbf{v}_{-i}(y; \beta) + \mathbf{x}_k \cdot \mathbb{1}(y_{-i} = y^*_{-i}),$$

where $\mathbf{v}_{-i}$ is independent of $\mathbf{x}_k$. This ensures that, when the value of $\mathbf{x}_k$ is high, the payoffs that players $-i$ obtain from playing y_{-i} become arbitrarily high.

The second condition is that players perform two rounds of eleimination of absolutely dominated strategies in the sense of Halpern and Pass (2012) and Salcedo (2012, §4).[11] This mild rationality requirement is satisfied by all solution concepts

[11] An action y_i^0 is absolutely dominated by y_i^* with respect to Y_{-i} if $\min_{y_{-i} \in Y_{-i}} u_i(y_i^*, i_{-i}) > \max_{y_{-i} \in Y_{-i}} u_i(y_i^0, i_{-i})$. That is, whenever the worst payoff that player i can get from playing

which imply two rounds of (myopic) rationality, and by those which imply individual rationality. In particular, can accommodate different forms of collusion and standard departures from Bayesian rationality. In any case, our assumptions are implied by standard assumptions used elsewhere in the literature, for instance, in Bajari et al. (2010).

Assumption 8 guarantees that, asymptotically, players faces single-agent decision problems. The next step is to assume that the observed distribution of optimal choices in these problems identifies the parameters that govern the *marginal* distribution of i's payoffs.

Assumption 9 For every $\beta \in B \setminus \{\beta_0\}$ there exists a player i, a strategy profile y^*, and a set $X' \subseteq X$, such that $\Pr(X') > 0$, $\mathrm{BR}_i(y^*_{-i}; b_0\mathbf{e}, \mathbf{x})$ is single-valued a.s. conditional on $\mathbf{x} \in X'$, and

$$\Pr\left(y_i^* \in \mathbf{BR}_i(y^*_{-i}; \beta_0) \,\middle|\, \mathbf{x} \in X'\right) \neq \Pr\left(y_i^* \in \mathbf{BR}_i(y^*_{-i}; \beta) \,\middle|\, \mathbf{x} \in X'\right). \tag{15}$$

Different standard sets of conditions imply Assumption 9. For example, one could assume independent action specific additive residuals with a known distribution as Bajari et al. (2010), or semi-parametric single index models as Fox (2007). In any case, these assumptions are sufficient to point identify β_0 without imposing more restrictive solution concepts.

Proposition C.1 *Under assumptions 1–2 and 7–9, β_0 is point identified.*

Proof. When best response correspondences are single-valued almost surely, Assumptions 7, 8, and 9 imply Assumption 5. To see this, simply take $Y_k = \{y_i^k\} \times Y_{-i}$, $X_k^n = \{x_k \geq n\} \times X_{-k}$, and $h_k(\beta_0, e, x) = \mathbb{1}(\mathrm{BR}_i(y_{-i}; \beta_0, e, x), y_{-i})$. The proof of sufficiency for Proposition 5.1 does not require Assumption 3. Hence, the exact same argument can applies in this case. ■

Our restriction to single-valued best response correspondences is not crucial. The same result holds true without it, but the proof is significantly longer.

---

y_i^* is strictly greater than the best payoff she can get from playing y_i^0.

D. Omitted details from section 7

First we need to characterize the set of BNE. Note that, interim preferences over pure strategies are represented by the expected utility function:

$$v_i(a; x, e_i) = \Big[\beta_{0i}x_i - e_i - \beta_{03}\Pr(a_{-i}(\mathbf{e}_{-i}) = 1 \,|\, x)\Big]a_i(e_i). \tag{16}$$

Since (16) is linear in a_i, and the coefficient is monotone in e_i, it follows that best response correspondence is monotone in e_i and the agent is generically not indifferent. Hence, there can only be pure strategy BNE and they need to be threshold strategies as in (8). Furthermore, such strategies constitute a BNE if and only if each agent i is indifferent between entering and not entering, which is exactly (9).

To establish existence, note that (9) can be rewritten as:

$$\bar{e}_1 = f_1(\bar{e}_2) \equiv \beta_{01}x_1 - \beta_{03}\Phi(\bar{e}_2) \qquad \wedge \qquad \bar{e}_1 = f_2(\bar{e}_2) \equiv \Phi^{-1}\left(\frac{\beta_{02}x_2 - \bar{e}_2}{\beta_{03}}\right),$$

and let $f_3(e_2) = f_1(e_2) - f_2(e_2)$, for $e_2 \in (\beta_{02}x_2 - \beta_{03},\ \beta_{02}x_2)$. Note that f_3 is continuous, and converges to $-\infty$ and ∞ at the extremes of these interval. Hence, the mean value theorem implies that there exist some e_2 such that $f_3(e_2) = 0$, i.e., $f_1(e_2) = f_2(e_2)$. There exists a set of parameters for which there are multiple equilibria, but all equilibria are of this form.

It remains to establish the missing lemma for the proof of Proposition 7.1.

Lemma D.1 *Let $x(t) \in X$ be given by $x_i(t) = \beta_{03}t/\beta_{0i}$. There exists some $t_0 \in \mathbb{R}$ such that $M_{\mathrm{NE}}(t) < m_{\mathrm{BNE}}(t)$ for every $t \geq t_0$, where:*

$$M_{\mathrm{NE}}(t) \equiv \max\left\{\psi(1,1) \,\middle|\, \psi \in \mathbb{E}[\,\bar{\mathbf{q}}_{\mathrm{NE}}(\beta_0)|x(t)\,]\right\},$$
$$m_{\mathrm{BNE}}(t) \equiv \min\left\{\psi(1,1) \,\middle|\, \psi \in \mathbb{E}[\,\bar{\mathbf{q}}_{\mathrm{BNE}}(\beta_0)|x(t)\,]\right\}.$$

Proof. Note that $x(t)$ makes the incomplete information game symmetric, so that $\bar{e}_1(x(t)) = \bar{e}_2(x(t))$ for all t. Let $\hat{e}(t) \in (\beta_{03}(t-1), \beta_{03}t)$ be the minimum solution to (9), so that the probability of $(1,1)$ in any BNE is bounded below by the probability of of $(1,1)$ in the BNE with thresholds $\hat{e}(t)$, i.e.,

$$m_{\mathrm{BNE}}(t) = \Phi^2(\hat{e}(t)). \tag{17}$$

On the other hand, for the complete information game, there are three regions of realizations of e_i. If $e_i > \beta_{03}t$ then not entering is dominant for i, if $e_i < \beta_{03}(t-1)$ then entering is dominant for i, and, if $e_i \in (\beta_{03}(t-1), \beta_{03}t)$ then i wants to enter if $-i$ is not entering, and wants to stay out if $-i$ is entering. Hence, the outcome $(1,1)$ can occur in equilibrium only if $e_i < \beta_{03}(t-1)$ for $i = 1,2$ or if $e_i \in (\beta_{03}(t-1), \beta_{03}t)$ for $i = 1,2$. In the first case, $(1,1)$ is the only Nash equilibrium. In the former case, the game has two pure equilibria $(0,1)$ and $(1,0)$, and a strictly mixed equilibrium in which $(1,1)$ occurs with some probability $q \in (0,1)$. This implies that

$$\begin{aligned} M_{\mathrm{NE}}(t) &= \Phi^2(\beta_{03}(t-1)) + \Big[\Phi^2(\beta_{03}t) - \Phi(\beta_{03}(t-1))\Big]^2 q^2 \\ &< \Phi^2(\beta_{03}(t-1)) + \Big[\Phi^2(\beta_{03}t) - \Phi(\beta_{03}(t-1))\Big]^2 \end{aligned} \tag{18}$$

Now let $G : \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by:

$$G(t) = \frac{\Phi^2(e^*(t)) - \Phi^2(\beta_{03}(t-1))}{[\Phi(\beta_{03}t) - \Phi(\beta_{03}(t-1))]^2}$$

Combining (17) and (18), we can establish the desired result if we can show that there exists some t_0 such that $G(t) > 1$ for all $t \geq t_0$. We will actually show the stronger result $\lim_{t\to\infty} G(t) = \infty$. To keep the notation simple, we assume from here on that $\beta_{03} = 1$, the proof can be easily adapted to the general case with $\beta_{03} > 0$.

First note that:

$$\begin{aligned} \lim_{t\to\infty} G(t) &= \lim_{t\to\infty} \frac{\Phi(\hat{e}(t)) - \Phi(t-1)}{[\Phi(t) - \Phi(t-1)]^2} \cdot \Big(\Phi(\hat{e}(t)) + \Phi(t-1)\Big) \\ &= 2 \cdot \lim_{t\to\infty} \frac{\Phi(\hat{e}(t)) - \Phi(t-1)}{[\Phi(t) - \Phi(t-1)]^2} \end{aligned}$$

Both the numerator and the denominator converge to 0, hence we would like to apply L'Hôpital's rule (further ahead we show that this can indeed be done, since all the limits are well defined). For that purpose, recall that $\hat{e}(t)$ is a (symmetric) solution to (9) with $x = x(t)$, i.e.,

$$\hat{e}(t) = t - \Phi(\hat{e}(t)).$$

Therefore, taking implicit derivatives with respect to t, it follows that $\hat{e}$ is differentiable and:

$$\hat{e}'(t) = \frac{1}{1 + \phi(\hat{e}(t))},$$

where ϕ denotes the standard normal p.d.f.. Hence, applying L'Hôpital's rule, and using the fact that $\phi'(x) = -x \cdot \phi(x)$, it follows that:

$$\begin{aligned}
\lim_{t\to\infty} G(t) &= 2 \cdot \lim_{t\to\infty} \frac{\phi(\hat{e}(t))\hat{e}'(t) - \phi(t-1)}{2[\Phi(t) - \Phi(t-1)][\phi(t) - \phi(t-1)]} \\
&= \lim_{t\to\infty} \left(\frac{\phi(t-1) - \phi(\hat{e}(t))\hat{e}'(t)}{\phi^2(t-1)} \right) \left(\frac{\phi(t-1)}{\Phi(t) - \Phi(t-1)} \right) \left(\frac{\phi(t-1)}{\phi(t-1) - \phi(t)} \right) \\
&\geq \liminf_{t\to\infty} \frac{\phi(t-1) - \phi(\hat{e}(t))\hat{e}'(t)}{\phi^2(t-1)} \times \lim_{t\to\infty} \frac{\phi(t-1)}{\Phi(t) - \Phi(t-1)} \cdots \\
&\qquad \cdots \times \lim_{t\to\infty} \frac{\phi(t-1)}{\phi(t-1) - \phi(t)}.
\end{aligned}$$

We will show that the first term is strictly positive, the second limit equals $+\infty$, and the third one equals 1. This implies that $\lim_{t\to\infty} G(t) = +\infty$, thus completing the proof.

We begin with the second and third limits. First note that:

$$\frac{\phi(t)}{\phi(t-1)} = \exp\left(-\frac{1}{2}t^2\right) \exp\left(\frac{1}{2}(t-1)^2\right) = \exp\left(\frac{1}{2} - t\right) \xrightarrow[t\to\infty]{} 0.$$

This implies that:

$$\lim_{t\to\infty} \frac{\phi(t-1)}{\phi(t-1) - \phi(t)} = \lim_{t\to\infty} \frac{1}{1 - \phi(t)/\phi(t-1)} = 1. \tag{19}$$

For the second limit, L'Hôpital's rule and (19) imply that:

$$\lim_{t\to\infty} \frac{\phi(t-1)}{\Phi(t) - \Phi(t-1)} = \lim_{t\to\infty} (t-1) \times \frac{\phi(t-1)}{\phi(t-1) - \phi(t)} = +\infty.$$

The first limit is slightly more complicated because we do not have a closed form expression for $\hat{e}(t)$. Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by:

$$g(t) = \frac{\phi(t-1) - \phi(\hat{e}(t))\hat{e}'(t)}{\phi^2(t-1)} = \frac{1 - \frac{\phi(\hat{e}(t))}{\phi(t-1)}\hat{e}'(t)}{\phi(t-1)}$$

It only remains to show that $\liminf_{t\to\infty} g(t) > 0$. Suppose towards a contradiction that this is false. It is straightforward to see that $g(t) \geq 0$, and thus this is equivalent to supposing that $\liminf_{t\to\infty} g(t) = 0$.

Under this supposition, there would exist an increasing sequence (t_n) such that $\lim t_n = \infty$ and $\lim g(t_n) = 0$. Now consider the sequence (r_n) given by:

$$r_n = \frac{\phi(\hat{e}(t_n))}{\phi(t_n - 1)}$$

Since $\hat{e}(t)$ belongs to $(t-1, t)$ for all t, it follows that $r_n \in (0,1)$ for all n. Further suppose towards a contradiction that $\liminf r_n < 1$. This would imply that there exists a subsequence (t_{m_n}) such that $r_{m_n} \longrightarrow r^* \in [0,1)$. This however would imply the contradiction:

$$\lim g(t_{m_n}) = \lim \frac{1 - r_{m_n}\hat{e}'(t_{m_n})}{\phi(t_{m_n} - 1)} = +\infty > 0,$$

where we used the facts that $\lim \hat{e}'(t_{m_n}) = 1$ and $\lim \phi(t_{m_n} - 1) = 0$. Therefore, if $\lim g(t_n) = 0$, then it must be the case that $\liminf r_n \geq 1$, and thus, since $r_n \in (0,1)$, it would follow that $\lim r_n = 1$. However, this would in turn imply the contradiction:

$$\begin{aligned}\lim g(t_n) &= \lim \frac{1 - r_n\hat{e}'(t_n)}{\phi(t_n - 1)} \geq \lim \frac{1 - \hat{e}'(t_n)}{\phi(t_n - 1)} = \lim \frac{1 - \frac{1}{1+\phi(\hat{e}(t_n))}}{\phi(t_n - 1)} \\ &= \lim \frac{\phi(\hat{e}(t_n))}{\phi(t_n - 1)} \times \frac{1}{1 + \phi(\hat{e}(t_n))} = \lim r_n \times \lim \frac{1}{1 + \phi(\hat{e}(t_n))} = 1 > 0.\end{aligned}$$

Therefore it must be the case that $\liminf_{t\to\infty} g(t) > 0$, and the proof is thus complete.

■

E. Equilibrium selection in coordination games

By point identifying the distribution of play, our approach allows one to identify the probability that each equilibrium is played. The role of this selection criteria is specially relevant in coordination games with multiple equilibria that are ranked in the Pareto sense. Different theoretical literatures suggest that agents are likely to coordinate on risk-dominant rather than efficient equilibria (e.g., Harsanyi and Selten (1988) and Carlsson and Van Damme (1993)). This has important welfare implications as it opens the possibility of welfare-improving policies. However, experimental evidence suggests that agents may coordinate on different equilibria depending on the specific payoffs (Battalio et al., 2001). Our methodology might bring some light to this issue by identifying the likelihood of different equilibria for different real life situations, and how these likelihoods depend on the characteristics of the environment.

To fix ideas, we analyze a particular n-player coordination game modelling a regime-change environment. However, similar payoff structures can be used to model various situations including coordinated attack problems (Rubinstein, 1989), bank-runs and currency attacks (Morris and Shin, 2003), or tacit collusion in oligopolistic markets Green et al. (2014).

Consider a small village with citizens $i \in I = \{1, 2, \ldots, \iota\}$. At a given time, each citizen chooses whether to manifest discontent towards the current regime (revolt) or not, $y_i \in \{0, 1\}$. The regime is changed if and only if the proportion of citizens revolting is greater than some threshold given by

$$\mathbf{t} = \Phi(\eta_0^\mathsf{T}\mathbf{w} + \eta_{0I}\mathbf{z}_I + \mathbf{e}_I),$$

where Φ is the standard normal p.d.f.. We normalize the payoff from not revolting to 0, and assume that the payoff from revolting is given by

$$\mathbf{u}_{0i}(1, y_{-i}) = \begin{cases} \lambda_0^\mathsf{T}\mathbf{w} + \lambda_{0i}\mathbf{z}_i + \lambda_{0i}^*\mathbf{z}_i^* + \mathbf{e}_i^* & \text{if} \quad \sum_i y_i > \mathbf{t} \cdot \iota \\ \gamma_0^\mathsf{T}\mathbf{w} + \gamma_{0i}\mathbf{z}_i + \gamma_{0i}^*\mathbf{z}_i^* + \mathbf{e}_i & \text{otherwise} \end{cases}.$$

In our specification $\mathbf{z} = (\mathbf{z}_I, (\mathbf{z}_i, \mathbf{z}_i^*)_{i\in I})$ is the vector of excluded covariates satisfying Assumption 3, $\mathbf{x} = (\mathbf{w}, \mathbf{z})$ is the vector of observed covariates, $\mathbf{e} = (\mathbf{e}_I, (\mathbf{e}_i, \mathbf{e}_i^*)_{i\in I})$ is the vector of error terms, and $\beta_0 = (\eta_0, \lambda_0, \gamma_0, (\lambda_{0i}, \lambda_{0i}^*, \gamma_{0i}, \gamma_{0i}^*)_{i\in I})$

is the vector of unknown structural parameters.

We maintain the assumption that payoffs are common knowledge and the citizens play NE, but we remain agnostic about which equilibria are more likely to arise. A coordination problem arises when citizens want to manifest discontent if and only if the revolt is successful, and $\mathbf{t} > 1/\iota$, meaning that more than one player is required for a successful revolt. In that case, there is a NE in which no one revolts, and it is Pareto dominated by a different NE in which everyone revolts. However, depending on the realized utility functions, the inefficient NE might be risk-dominant.

We do not require a lot of structure on $\mathbf{w}$ nor the signs of the parameters. We simply assume that (*i*) the distribution of the error terms belongs to the exponential family, (*ii*) the excluded covariates $\mathbf{z}$ are continuous and their coefficients are different from zero, and (*iii*) the support of $\mathbf{z}_I$ and $\mathbf{z}_i$ for $i \in I$ is the entire real line.[12] Covariate $\mathbf{z}_I$ should be something that affects the strength or resiliency of the current regime. For instance, it could be an indicator of the financial health of the regime. Each $\mathbf{z}_i$ and $\mathbf{z}_i^*$ should affect how much citizen i cares about changing the regime or about being on good terms with the current regime (in case the revolt fails). For instance, one could use household income or, better yet, an proxy for the proportion of i's business that depend on the current regime.

Our identification-at-infinity approach can be used in this setting to point identify β_0. However, even if β_0 was not point identified, our assumptions guarantee that Proposition 4.1 applies. Therefore h_0 is point identified, at least up to β_0. This makes it possible to recover from the data the average distribution of choices for each possible realization of the vNM indexes. Hence, we can measure the exact welfare loss arising from miss-coordination or coordination on inefficient equilibria, and the probability that each kind of equilibrium is selected as a function of the characteristics of the environment.

[12] We use the same covariates to identify payoff parameters and the distribution of play, but we do so only to keep the notation simple. In general, we do need some excluded covariates and some unbounded covariates, but they need not be the same.