Estimation of Discrete Games with Weak Assumptions on Information*

Lorenzo Magnolfi[†] and Camilla Roncoroni[‡]

This Draft: February 25, 2016
Last version available here

Abstract

We propose a method to estimate static discrete games with weak assumptions on the information available to players. In contrast to the existing literature, we do not fully specify the information structure of the game. Instead, we allow for all information structures consistent with the assumptions that players know their own payoffs and the distribution of opponents' payoffs. We make this approach tractable by adopting a weaker solution concept: Bayes Correlated Equilibrium (BCE), proposed by Bergemann and Morris (2013, 2015). We characterize the sharp identified set obtained under the assumption of BCE behavior. In simple games with modest levels of variation in observable covariates, identified sets are narrow enough to be informative, while avoiding the misspecification resulting from strong assumptions on information. In an application, we estimate a game theoretic model of entry in the Italian supermarket industry, and quantify the effect of the presence of large malls on competition. Our model yields parameter estimates and counterfactual predictions that differ from those obtained under the restrictive assumption of complete information.

Keywords: Estimation of games, informational robustness, Bayes Correlated Equilibrium, entry models, partial identification, supermarket industry.

*We are grateful to Phil Haile, Steve Berry, and Dirk Bergemann for their guidance and support. We thank Michi Igami for numerous suggestions. We thank Don Andrews, Tim Armstrong, Pierpaolo Battigalli, John Lazarev, Niccolò Lomys, Chiara Margaria, Marcelo Sant'Anna, Che-Lin Su and Ed Vytlacil for helpful discussions. We also thank several seminar participants at Johns Hopkins, NYU-Stern, Penn State, University of Wisconsin - Madison, Yale and at the Eief-Igier-Bologna Third Workshop in IO, ES NASM 2016, ES ESM 2016, EARIE 2016 and XXXI Jornadas de Economía Industrial for comments. We gratefully acknowledge financial support from the Cowles Foundation. This work was supported in part by the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center. All errors and omissions are our own.

[†]Department of Economics, University of Wisconsin-Madison, `magnolfi@wisc.edu`

[‡]Department of Economics, University of Warwick, `roncoroni@warwick.ac.uk`

1 Introduction

Empirical models of static discrete games are important tools in industrial organization, as they allow to recover the determinants of firms' behavior while accounting for the strategic nature of firms' choices. Models in this class have been applied in contexts such as entry, product or location choice, advertising, and technology adoption.[1] In discrete games the equilibrium predictions, and thus the map between the data and parameters of interest, depend crucially on the assumptions maintained on the information that players have on each other's payoffs. However, in applied contexts the nature of firms' information about their competitors is often ambiguous. Moreover, restrictive assumptions, when not satisfied in the application at hand, can result in inconsistent estimates of the payoff structure of the game.

We propose a new method to estimate the distribution of players' payoffs relying only on assumptions about the minimal information players have. In particular, we assume that players know at least (i) their own payoffs, (ii) the distribution of opponents' payoffs, and (iii) parameters and observable covariates. We admit any information structure that satisfies these assumptions. In this sense our model is incomplete, in the spirit of Manski (2003, 2009), Tamer (2003), and Haile and Tamer (2003). More precisely, we allow our model to produce any prediction that results from a Bayes Nash Equilibrium (BNE) under an admissible information structure. Our object of interest is the set of parameters that are identified given this incomplete model.

Our method nests the two main approaches used in the existing literature: complete information, frequently adopted since the pioneering work in this area (Bjorn and Vuong 1985, Jovanovic 1989, Bresnahan and Reiss 1991a, Berry 1992); and private information (see e.g. Seim 2006, and De Paula and Tang 2012). Moreover, our model generalizes the class of information structures considered by Grieco (2014), and is flexible in other dimensions. For example, agents can be informed about opponents' payoffs with different levels of accuracy, and the information structure of the game can vary across markets.

To make this approach tractable, we rely on the connection between equilibrium behavior and information, and adopt Bayes Correlated Equilibrium (BCE) as solution concept. BCE, introduced by Bergemann and Morris (2013, 2015), has the property of describing Bayes Nash Equilibrium predictions for a class of informational environments. We show that, under the assumption of BCE behavior, for every vector of parameters in the identified set there exists an admissible information structure and a BNE that deliver predictions compatible with the data. We can characterize the sharp identified set for the parameters of interest without modeling equilibrium selection, exploiting the convexity

[1] See for instance Bresnahan and Reiss (1991b), Berry (1992), Tovainen and Waterson (2005), Jia (2008), Ciliberto and Tamer (2009) for models of entry, Mazzeo (2002) and Seim (2006) for models of product choice, Sweeting (2009) for advertising, Ackerberg and Gowrisankaran (2006) for technology adoption.

of the set of equilibria. These results motivate the use of BCE to estimate the distribution of players' payoffs while being robust with respect to the informational environment, thus avoiding misspecification bias due to strong assumptions on information.

We investigate the identification power of BCE in simple entry games with linear payoffs and find that the identified sets are informative about the model's primitives. In fact, point identification is obtained under the assumption of full support variation in excluded covariates, as in Tamer (2003). More generally, however, we obtain partial identification of the payoff parameters and of the joint distribution of payoff types. We perform inference by constructing confidence intervals for the identified set using the methods of Chernozhukov, Hong and Tamer (2007).

We apply our method to the investigation of the effect of large malls[2] on the grocery retail industry in Italy. There is disagreement on the impact of the presence of these big outlets on local supermarkets, echoing the US debate on "Wal-Mart effects." Advocates of stricter regulation of large retailers claim that malls drive out existing supermarkets and leave consumers without the option of shopping at local stores.[3] However, economic theory[4] and some of the existing evidence from other countries[5] suggest that other effects might prevail. Despite the extra competition from supermarkets in malls, local stores might benefit from the agglomeration economies created by the mall.

We provide a quantitative assessment of these economic forces. We consider cross sectional data from 2013 and estimate a static entry game to capture the determinants of market structure in geographic grocery markets. We assume the entry of malls to be exogenous to the outcomes of competition among local supermarkets, conditional on market-level observables. This assumption is consistent with the presence of regulatory and geographic constraints on the choice of location for malls, and with malls' larger catchment area, determined by differences in habits for non-grocery and grocery shopping.

It is difficult to take a stance on the informational environment in this setting, especially in light of the heterogeneity among supermarket groups. Previous cross-section studies of entry have assumed that players have complete information, and interpreted the data as long-run equilibrium outcomes in which players have no ex-post regret. In our application the data provide only a snapshot of the industry at the end of a period of expansion, so it is not obvious that the configuration we observe is a stable outcome.

[2]We define large malls for the purpose of this paper as shopping centers with at least fifty independent shops and a grocery anchor.

[3]In the US, the presence of areas where consumers have limited access to fresh groceries from local stores (see for instance Bitler and Haider, 2011) has been linked to the presence of large retailers.

[4]Zhu, Singh and Dukes (2011) show that when the existing retailers offer non-overlapping product lines, they can benefit from the presence of large stores that can produce demand externalities.

[5]Ellickson and Grieco (2013) examine the effect of Wal-Mart's expansion into groceries on local supermarkets. They find evidence of a highly localized impact, which points to a significant dimension of horizontal differentiation in supermarkets possibly arising from consumers' travel costs.

Our weaker equilibrium assumption seems more appropriate than strong restrictions on information, which constrain regret in a way that does not fit this application.

We estimate the model using our robust method, and find mixed evidence on the effect of large malls on supermarkets. We do not reject values of parameters that indicate a strong negative effect of large malls on the supermarket groups that are least differentiated. We also do not reject values of parameters that indicate a beneficial effect of malls on some supermarket groups. We compare these estimates with those obtained using a model of complete information. Results differ in important ways. Under complete information, we do not reject low values (in absolute value) of competitive effects, which are rejected under weak assumptions on information. This is because the assumption of complete information imposes that players fully anticipate competitors' decisions. As a consequence, when not supported by the data, this more restrictive model may lead to underestimate how much players' profits are affected by the presence of opponents in a market.[6]

In a counterfactual, we evaluate the effect on market structure of removing large malls from markets that currently have no other supermarket. The model with complete information predicts that removing large malls results in a substantial increase in the average upper bounds of the probability of observing at least two entrants. We do not find a similar prediction with our method: the average upper bounds of the probability of observing at least two entrants may not change or may decrease. In this application, a model with restrictive assumptions on information leads us to strong conclusions, which are dispelled once more robust methods are adopted.

2 Related Literature

This paper belongs to the literature on identification and estimation of static discrete games, recently surveyed by Bajari, Hong and Nekipelov (2010) and De Paula (2013).[7] The works in this area can be classified according to the assumptions they adopt on information and on equilibrium behavior. Some authors consider the complete information game, and assume play to be Nash Equilibrium (typically in pure strategies), while others consider the incomplete information game as in Seim (2006), Sweeting (2009) and Bajari et al. (2010), and assume that data are generated by BNE play.

Grieco (2014) is the first to propose a method that relaxes the standard assumptions of either complete or perfectly private information. We share with Grieco (2014) the goal of considering more flexible information structures, but instead of a parametrization of

[6]The sign of this bias is consistent with the attenuation bias documented by Bergemann and Morris (2013) and Dickstein and Morales (2015) in the context of individual decision making, and by Aguirregabiria and Magesan (2015) in the context of dynamic games.

[7]See also Borkovsky et al. (2015) for a discussion of recent advances in empirical games.

the information structure we adopt the assumption of BCE behavior. We further relax the assumptions on the correlation of private signals and symmetry of the information structure that are embedded in an informational environment with a public signal. Our model allows for private signals that are differently informative across markets and across players. Mazzeo, Seim and Varela (2014) also develop a model of entry that nests complete and incomplete information by relying on public signals. They integrate the discrete game theoretic model of product choice with a model of imperfect competition in the product market to perform merger simulations. We only consider static games in this paper, but our framework is amenable to be integrated with a model of post-entry competition, as it can allow for rich information structures, in which for instance some firms have privileged information, and offers relatively manageable computational burden.

Other papers in the literature examine the role of information in empirical game theoretic models. Aradillas-Lopez (2010) describes semiparametric inference procedures for models in which the part of players' payoffs that is unobserved to the econometrician is private information, and players might be imperfectly informed about the part of opponents' payoffs that is observable to the econometrician. We assume that everything that is observed or can be estimated by the econometrician is common knowledge among players, but allow for a richer information structure on private payoff types since we are not constrained by the analytical necessity of finding BNEs of incomplete information games. Takahashi and Navarro (2012) develop testing procedures to distinguish between information structures. They notice that if payoffs shocks are independent and players follow a deterministic equilibrium selection rule, any correlation observed in their strategies is to be traced back to the information that players might have about the opponents' payoffs. This intuition is related to the idea in De Paula and Tang (2012), who use a similar argument to test for multiplicity of equilibria. Our method allows correlation in players' strategies to come from correlation in unobservables, information and multiplicity of equilibria.[8] It does not rely on strong assumptions on information, independence of private shocks, or assumptions on equilibrium selection, and can be used when more restrictive assumptions on information are rejected by the data.[9]

The literature offers also different ways of dealing with the equilibrium multiplicity inherent in game theoretic models.[10] We follow Tamer (2003), Berry and Tamer (2006) and Ciliberto and Tamer (2009), who abandon strong assumptions on equilibrium selection and allow for set identification of parameters. Our model also relaxes assumptions

[8]See also Xu (2014) for a model that allows for correlated payoff types.

[9]While we do not pursue testing in this paper, it is possible to recast the problem in an inferential framework that naturally allows for testing such as the semiparametric likelihood methods of Chen, Torgovitsky and Tamer (2011) used by Grieco (2014).

[10]Previous work considered aggregation of outcomes (Bresnahan and Reiss 1991b), assumptions on the order of entry (Berry 1992), and parametric equilibrium selection rules (Bajari, Hong and Ryan 2010) to address multiplicity.

on the information available to players, an important dimension of unobserved heterogeneity. The focus of this paper on partial identification is in line with the broader research agenda summarized in Manski (2003, 2007) and Tamer (2010). In particular, we rely on some ideas presented in Beresteanu, Molinari and Molchanov (2011), who provide a useful characterizations of the sharp identified set for models with convex predictions.[11]

We build on the theoretical work of Bergemann and Morris (2013, 2015). These authors define the equilibrium concept used in this paper and highlight the link between behavioral assumptions and information structures. They discuss robust prediction using BCE as a way of capturing the implications of equilibrium behavior for a given payoff environment, allowing for all possible information structures. We focus instead on identification, recovering elements of the payoff structure when the outcomes of the game are observed but the information structure is unknown. Bergemann and Morris (2013) discuss both identification and prediction in a linear-quadratic common value environment under BCE behavior. This paper considers identification in private value models of discrete games widely used in the applied literature, and discusses techniques that allow employing BCE in estimation of empirical models of games.

We are not the first to consider the empirical content of weaker assumptions on players' behavior. Aradillas-Lopez and Tamer (2008) consider identification and inference in static models of discrete games under the assumption of rationalizable behavior. We adopt BCE as a solution concept to pursue robustness to unobserved informational environments, and leverage on the tractability and empirical content of BCE. Yang (2009) examines estimation of discrete games of complete information under Nash behavior, using the non-sharp restrictions imposed by Correlated Equilibrium in order to simplify computation. Bayes Correlated Equilibrium is similarly convenient to compute, and at the same time allows us to obtain sharp restrictions for Bayes Nash behavior in a class of games that allows for a rich range of information structures. Aguirregabiria and Magesan (2015) develop a method for estimation of dynamic discrete games allowing for players' beliefs not to be in equilibrium, while maintaining the assumption of independent private payoff shocks. We assume equilibrium behavior, but under our weak assumptions we also avoid strong restrictions on beliefs.[12]

Our emphasis on identification and estimation under weak assumptions on information is similar to the spirit of Dickstein and Morales (2015), who examine a non strategic model of firms' export decisions, and develop a method to estimate payoff parameters without fully specifying firms' information on their expected revenues. They show that

[11] Galichon and Henry (2011) also provide a characterization of the identified set for discrete games.

[12] While their model is motivated by strategic uncertainty of players, we focus instead on removing strong assumptions on the information structure, a primitive that determines crucially players' beliefs in equilibrium.

restrictive assumptions on information result on bias on parameters, and rely on moment inequalities to set-identify parameters under weak assumptions. We consider instead game-theoretic models, and obtain sharp bounds by explicitly capturing the implications of a rich class of information structures and relying on the BCE assumption.

Our study of the effect of the presence of large malls on local supermarkets is related to several papers that examine the effect of entry of large store formats, especially Wal-Mart in the US, on other retailers.[13] Among the papers using game theoretic entry models, Jia (2008) estimates a structural model with chain effects in which first Wal-Mart and Kmart make simultaneous moves and then small retailers decide whether to open, and finds that the entry of Wal-Mart impacts negatively the number of small retailers. Grieco (2014) uses his model of entry with flexible information to study the impact of Wal-Mart's Supercenters on small grocery stores in rural US counties, and reports a mild but largely negative effects of the presence of Wal-Mart on small stores' profits. The impact of large retailers on smaller stores has also been investigated outside the US. Igami (2011) examines the supermarket industry in Tokyo, studying the effect of the entry of large supermarkets on competitors. We examine instead the Italian supermarket industry, and focus on the impact that large grocery stores that anchor regional malls have on local supermarkets.[14] In a companion paper, Magnolfi and Roncoroni (2015), we study the role of political connections in shaping market structure in this industry.

3 Model

We consider a class of static games, indexed by realizations of covariates $x \in X \subseteq \mathbb{R}^{d_X}$. Games with different levels of x can be interpreted as different markets where firms interact. Players are in a finite set N. Each player $i \in N$ chooses an action $y_i \in Y_i$. The action space $Y = \times_{i \in N} Y_i$ and N do not depend on x. The econometrician observes cross-sectional data on discrete outcomes $y \in Y \subset \mathbb{R}^{d_Y}$ and covariates x, and wants to recover the determinants of behavior. The setup is summarized in Assumption 1 below.

Assumption 1. The econometrician observes the distribution $P_{x,y}$ of the random vector (x, y). This joint distribution induces a set of conditional probability measures

$$\left\{ P_{y|x} \in \Delta(Y) : x \in X \right\}$$

on Y. The finite set of players N is also observable.

[13] See Basker (2007) for a survey of this literature.

[14] For the Italian supermarket industry, Viviano (2008) studies the impact of entry deregulation on employment in the retail sector, finding a positive effect of large store entry on employment in small retail firms in part of her sample.

To identify the determinants of behavior, the first step is to assume that the data are generated by a true structure in a well-defined class. We outline the primitives of this structure in the next subsections, describing separately the payoff environment and the informational environment that players face. All features of the true structure generating observed behavior are common knowledge among players.

3.1 Payoff Environment

Every player i has a payoff type $\varepsilon_i \in \mathcal{E}_i \subseteq \mathbb{R}$. Payoff types $\varepsilon = (\varepsilon_i)_{i \in N}$ are distributed according to the cdf $F(\cdot;\theta_\varepsilon)$, parametrized by the finite dimensional parameter $\theta_\varepsilon \in \Theta_\varepsilon \subseteq \mathbb{R}^{d_{\Theta_\varepsilon}}$. Payoffs to player i, denoted by π_i, depend on action profiles and realizations of payoff types. Payoffs are also affected by observable covariates $x \in X$, and finite dimensional payoff parameters $\theta_\pi \in \Theta_\pi \subseteq \mathbb{R}^{d_{\Theta_\pi}}$, so that for every player i and every pair (x, θ_π) there is a map:

$$\pi_i^{x,\theta_\pi} : Y \times \mathcal{E}_i \to \mathbb{R}.$$

We assume that ε is independent of the vector x. A realization of x and a vector of parameters $\theta = (\theta_\pi, \theta_\varepsilon) \in \Theta = \Theta_\pi \times \Theta_\varepsilon$ pins down a payoff structure. We want to identify, from data on behavior and market observable characteristics x, the vector of parameters θ. We present a model with ε independent of x and finite dimensional parameters θ, but these restrictions are not necessary for our general discussion of robust identification.[15]

We introduce now an example that we will use throughout the description of the model: a two-player entry game with payoffs linear in covariates, and independent uniformly distributed types.

Example 1. *(Two Player Entry Game)* Consider a model of a two player, binary action game. Players are in the set $N = \{1, 2\}$. Actions are "out" or "enter" , represented as $Y_i = \{0, 1\}$. Payoffs are:

$$\pi_i^{x,\theta_\pi} = y_i \left(x_i' \beta_i + \Delta_{-i} y_{-i} + \varepsilon_i \right),$$

so that the payoff parameter vector is $\theta_\pi = (\beta_i, \Delta_i)_{i=1,2}$. Payoff types ε_i are distributed iid according to a uniform distribution on $[-1, 1]$. Payoffs can be visualized in the following payoff matrix:

[15] We do not pursue nonparametric identification of the payoff structure in this paper, as it is not directly related to our main goal of achieving robustness with respect to assumptions on information. We present instead a simple parametric setup to maintain the link with the previous literature and applied work. See Lewbel and Tang (2015) for an example of nonparametric identification and estimation of the payoff structure in models of games with incomplete information, and Tang (2010) for a model that relaxes the independence between ε and x.

Player 2:		Out	Enter
Player 1:		0	1
Out	0	$(0, 0)$	$\left(0,\, x_2'\beta_2 + \varepsilon_2\right)$
Enter	1	$\left(x_1'\beta_1 + \varepsilon_1,\, 0\right)$	$\begin{pmatrix} x_1'\beta_1 + \Delta_2 + \varepsilon_1, \\ x_2'\beta_2 + \Delta_1 + \varepsilon_2 \end{pmatrix}$

3.2 Informational Environment

We assume that every player i knows the realization of her payoff type ε_i. In addition, every player receives a private random signal τ_i^x, which may be informative about the full vector of payoff types ε. An information structure S specifies, for every value of x, the set of signals a player can receive and the probability of receiving them, depending on the realization of the vector of payoff types. Formally:

$$S = \left(T^x, \left\{\mathbb{P}_{\tau|\varepsilon,x} : \varepsilon \in \mathcal{E}\right\}\right)_{x \in X},$$

where T^x is a subset of a separable metric space and represents the set of realizations of the vector of signals $\tau^x = \left(\tau_i^x\right)_{i \in N}$. The probability kernel $\left\{\mathbb{P}_{\tau|\varepsilon,x} : \varepsilon \in \mathcal{E}\right\}$ is the collection of probability distributions of τ^x conditional on every realization of ε. The sets of signals and the distribution of signal vectors depend on x, since we allow the informational environment to change with observed characteristics of the payoff environment.

We denote $\mathcal{S}_0$ the collection of all possible information structures S. More formally, $\mathcal{S}_0$ is a general nonparametric class of information structures:

$$\mathcal{S}_0 := \left\{S \mid \forall x \in X,\ T^x \text{ is separable metric space, } \mathbb{P}_{\tau|\varepsilon,x} \text{ is probability measure on } (T^x, \mathcal{B}(T^x))\right\},$$

where $\mathcal{B}$ denotes the Borel $\sigma-$algebra.

Two extreme examples of information structures are *complete information*, denoted by $\overline{S}$, and *minimal information*, denoted by $\underline{S}$, which corresponds to the private value environment. Most prior work on estimation of discrete games assumes one of these extremes, both nested by our framework. The structure $\overline{S}$ features $T_i^x = \mathcal{E}$ for every $x \in X$, and provides players with perfectly informative signals: $\mathbb{P}_{\tau^x|\varepsilon}\{\varepsilon\} = 1$ for all $\varepsilon \in \mathcal{E}$, $x \in X$. Instead, in the minimal information structure $\underline{S}$ signals τ^x are uninformative: $\mathbb{P}_{\tau^x|\varepsilon} = \mathbb{P}_{\tau^x}$ for all $\varepsilon \in \mathcal{E}$, $x \in X$.

Our framework can also accommodate many more examples of information structures, such as *privileged information* S^P in which some players know the type of some other players.[16] In this case, the signal spaces for all players are $T_i^x = \mathcal{E}$. For an informed

[16]This information structure resembles the ones that characterize the proprietary information model of

player i, signals τ_i^x are distributed according to $\mathbb{P}_{\tau_i^x|\varepsilon}\{\varepsilon\} = 1$ for all $\varepsilon \in \mathcal{E}$, $x \in X$, while for an uninformed player j signals τ_j^x are distributed according to $\mathbb{P}_{\tau_j^x|\varepsilon} = \mathbb{P}_{\tau_j^x}$.

In some models, the information structure can depend on the correlation among payoff types, as in the following example of *informative Normal public signals*.

Example 2. *(Informative Normal Public Signals)* Consider a model of a two player, binary action game, with $N = \{1,2\}$ and $Y_i = \{0,1\}$. Payoffs are:

$$\pi_i^{x,\theta_\pi} = y_i\left(\theta_\pi y_{-i} + \varepsilon_i\right),$$

and payoff types ε_i are jointly Normal with zero means, unit variances, and correlation ρ. Moreover, player i receives signals $\tau_i = \left(\tau_i^1, \tau_i^2\right) \in \mathbb{R}^2$, distributed according to:

$$\mathbb{P}_{\tau_1|\varepsilon} = N\left(\left(\begin{array}{c} \varepsilon_2 \\ \varepsilon_1 \end{array}\right), \left(\begin{array}{cc} 1-\frac{\rho}{\sigma^2} & 0 \\ 0 & 1-\frac{\rho}{\sigma^2} \end{array}\right)\right),$$

and $\left(\tau_1^1, \tau_1^2\right) = \left(\tau_2^2, \tau_2^1\right)$. One interpretation for this model is that the payoff type ε_i can be decomposed in two parts, $\varepsilon_i = \eta_i^1 + \eta_i^2$, and $\tau_i = \left(\eta_{-i}^1, \eta_i^1\right)$. The vector $\left(\eta_{-i}^1, \eta_i^1\right)$ represents a publicly observed component of the payoff type that is correlated across players, while η_i^2 is an idiosyncratic and privately known component of the payoff type. The class of information structures defined in this example is the same that is considered in Grieco (2014). See Appendix C for more discussion.

3.3 Equilibrium

The parameter vector $\theta = (\theta_\pi, \theta_\varepsilon)$ and the information structure S summarize the elements of the structure that are unknown to the econometrician; a pair (θ, S) pins down an incomplete information game $\Gamma^x(\theta, S)$ for every x. We assume that players' behavior is described by a profile of strategies that are a Bayes Nash Equilibrium (BNE) of this game. We denote the equilibrium strategy profile as:

$$s = (s_i)_{i\in N} \in \left(\times_{i\in N}\left(\Delta\left(Y_i\right)\right)^{\mathcal{E}_i \times T_i^x}\right).$$

We also denote the set of all BNE for the game $\Gamma^x(\theta, S)$ as $E_{\theta,S,x}^{BNE}$. In general an incomplete information game $\Gamma^x(\theta, S)$ can have multiple equilibria, so that the set of equilibria $E_{\theta,S,x}^{BNE}$ may not be a singleton.

The informational environment of the game has an important impact on equilibrium behavior. When players receive informative signals on their opponents' payoff types,

Engelbrecht-Wiggans, Milgrom and Weber (1983) for common value auctions, and the model of Kim and Che (2004) for independent private value auctions.

their beliefs and hence their equilibrium behavior will reflect this information. The more informative the signals that player i receives about ε_{-i}, the more we expect player i's equilibrium behavior to vary with the realizations of ε_{-i}. Conversely, players who receive uninformative signals will only base their equilibrium behavior on their own payoff type.[17]

Example. *(1, Cont.)* We illustrate these points in Figure 1, which depicts equilibrium outcomes in the space of payoff types for a two-player entry game with no covariates x, competitive effects $\Delta_1 = \Delta_2 = -\frac{1}{2}$, and payoff types iid uniform over the interval $[-1,1]$. The three panels correspond, respectively, to games with information structure $\overline{S}$, $\underline{S}$ and S^P, and show how different informational environments result in radically different equilibrium behavior.

For each equilibrium strategy $s \in E^{BNE}_{\theta,S,x}$ we can formulate the following prediction on behavior:

Definition 1. *(BNE Prediction)* A BNE s of the game $\Gamma^x(\theta, S)$ induces a distribution over outcomes p_s :

$$p_s(y) = \int_{\varepsilon\in\mathcal{E}} \left[\int_{\tau\in T} \prod_{i\in N} \{s_i(y_i|\varepsilon_i,\tau_i)\}\, d\mathbb{P}_{\tau|\varepsilon}\{\tau\} \right] dF(\varepsilon;\theta_\varepsilon),$$

for all $y \in Y$, where $s_i(y_i|\varepsilon_i,\tau_i)$ indicates the probability of y_i as specified by $s_i(\varepsilon_i,\tau_i)$.

The set $E^{BNE}_{\theta,S,x}$ of equilibria might not be a singleton, and we do not make any specific assumption on equilibrium selection. For any couple (θ, S), we define a prediction correspondence $Q^{BNE}_{\theta,S} : X \rightrightarrows \Delta^{|Y|-1}$

$$Q^{BNE}_{\theta,S}(x) := \text{co}\left[\left\{p \in \Delta^{|Y|-1} : \exists\, s \in E^{BNE}_{\theta,S,x} \text{ such that } p = p_s\right\}\right],$$

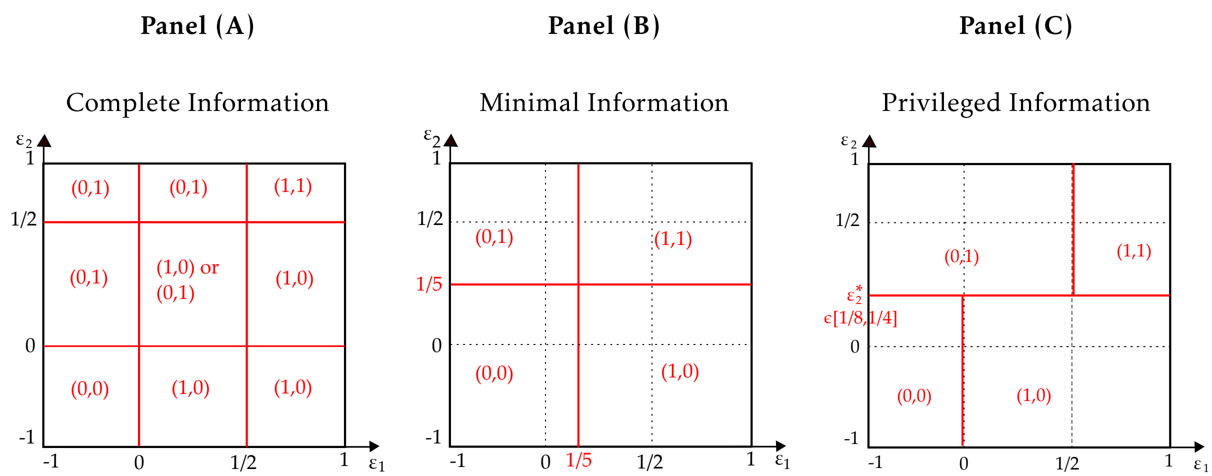
where co$[\cdot]$ takes the convex hull of a set. The prediction correspondence describes the set of distributions over actions y that can be obtained in a game $\Gamma^x(\theta_0, S_0)$ under the assumption of BNE play. The convex hull operator takes care of the multiplicity of equilibria, considering all possible distributions over equilibria. In the next section, we consider identification in this model.

4 Identification

We maintain that for each level of market characteristics x, observed behavior is compatible with a BNE in a game $\Gamma^x(\theta_0, S_0)$ in the class described in Section 3. We are interested

[17]This results in different levels of ex-post regret that players experience: when not informed about their opponents' type, players might optimally choose actions that will result suboptimal ex post, when the equilibrium strategy profile is realized.

Figure 1: Equilibrium Behavior and Information



This figure represents Bayes Nash Equilibrium outcomes in the space $(\varepsilon_1, \varepsilon_2)$ for the two-player entry game described in Example 1, with payoffs $\pi_i(y, \varepsilon) = y_i\left(-\frac{1}{2}y_j + \varepsilon_i\right)$ for $i = 1, 2$ and $\varepsilon_i \overset{iid}{\sim} U[-1, 1]$.

Panel (A) represents outcomes for a game with complete information, in which, for $(\varepsilon_1, \varepsilon_2) \in \left[0, \frac{1}{2}\right]^2$ there are multiple equilibria.

Panel (B) represents minimal information outcomes, generated by equilibrium strategies that prescribe $y_i = 1$ whenever $\varepsilon_i \geq \frac{1}{5}$.

Panel (C) represents privileged information outcomes, generated by equilibria in which player 2 does $y_2 = 1$ whenever ε_2 is greater or equal than a threshold $\varepsilon_2^* \in \left[\frac{1}{8}, \frac{1}{4}\right]$, and player 1 responds optimally to the realization of ε_2, which he observes, and player 2's equilibrium strategy.

in recovering θ_0, but we do not know the true information structure S_0. We first link the game theoretic structure to observables, and then describe the identified set we can obtain for the parameters of interest if we allow for any $S \in \mathcal{S}_0$, and for any Bayes Nash Equilibrium constructed given any S.

Given the formal characterization of the implications of equilibrium behavior provided by Definition 1, we summarize below our assumptions on the data generating process.

Assumption 2. For all $x \in X$, the outcomes y are generated by equilibrium play of the game $\Gamma^x(\theta_0, S_0)$, so that $P_{y|x} \in Q^{BNE}_{\theta_0,S_0}(x)$.

Under Assumptions 1 and 2, the sharp identified set of parameters is defined as:

$$\Theta_I^{BNE}(\mathcal{S}_0) = \left\{\theta \in \Theta | \exists\ S \in \mathcal{S}_0 \text{ such that } P_{y|x} \in Q^{BNE}_{\theta,S}(x)\ P_x - a.s.\right\}. \tag{4.1}$$

This is the set of parameters whose implications, without restrictions on equilibrium selection or information structure, are compatible with the observables. All parameters $\theta \in \Theta_I^{BNE}(\mathcal{S}_0)$ are *observationally equivalent*, since for each of them there exists an information structure $S \in \mathcal{S}$ that generates a correspondence $Q^{BNE}_{\theta,S}$ rationalizing the observables. The set $\Theta_I^{BNE}(\mathcal{S}_0)$ is our object of interest. It captures the restrictions on the parameters that we obtain under weak assumptions on the information structure.

However, definition (4.1) seems hardly useful in practice, since computing correspondences $Q^{BNE}_{\theta,S}$ for all S in the large class $\mathcal{S}_0$ is an analytical challenge. A brute force approach would in fact require specifying all possible information structures, and finding the corresponding sets of predictions. In the following subsections, we propose a method to identify the set $\Theta_I^{BNE}(\mathcal{S}_0)$ that sidesteps the analytical difficulties inherent in a direct approach by relying on the connection between equilibrium behavior and robustness to assumptions on information.

4.1 Bayes Correlated Equilibrium

In this subsection, we show how the adoption of Bayes Correlated Equilibrium as solution concept solves the problem of characterizing the robust identified set $\Theta_I^{BNE}(\mathcal{S}_0)$. We start with the definition of BCE, which follows Bergemann and Morris (2015).

Definition 2. *(BCE)* A Bayes Correlated Equilibrium (BCE) for the game $\Gamma^x(\theta, \underline{S})$ is a probability measure ν over $(Y \times \mathcal{E})$ that is consistent with the prior:

$$\sum_{y \in Y}\left[\int_{\{\tilde{\varepsilon} \leq \varepsilon\}} d\nu\{y, \tilde{\varepsilon}\}\right] = F(\varepsilon; \theta_\varepsilon),$$

for all $\varepsilon \in \mathcal{E}$, and incentive compatible:

$$\forall i, \varepsilon_i, y_i \text{ such that if } \nu\{y_i|\varepsilon_i\} > 0, \quad \sum_{y_{-i}\in Y_{-i}} \left[\int_{\mathcal{E}_{-i}} \pi_i^{x,\theta_\pi}(y, \varepsilon_i)\, d\nu\{y_{-i}, \varepsilon_{-i}|y_i, \varepsilon_i\} \right] \geq$$
$$\sum_{y_{-i}\in Y_{-i}} \left[\int_{\mathcal{E}_{-i}} \pi_i^{x,\theta_\pi}\left(\left(y_i', y_{-i}\right), \varepsilon_i\right) d\nu\{y_{-i}, \varepsilon_{-i}|y_i, \varepsilon_i\} \right] \qquad \forall y_i' \in Y_i.$$

The BCE concept is an extension of Correlated Equilibrium to an incomplete information setup, under the assumptions that players have a common prior on the distribution of payoff types, and can observe additional signals. BCE behavior is not represented by strategy functions, but rather by a joint distribution of observable actions and payoff types. This distribution needs to be consistent with the common prior that players maintain, so that its marginal over payoff types reflects the common knowledge of the underlying distribution of ε. Moreover, players are best responding to equilibrium beliefs, as summarized by the BCE distribution.

Notice that we define BCE for the game of minimal information $\Gamma^x(\theta, \underline{S})$ in which players only know their own payoff type ε_i. While in principle BCE can be defined for any incomplete information game, we use Definition 2 in what follows, and denote as $E^{BCE}_{\theta,x}$ the set of BCE for the game $\Gamma^x(\theta, \underline{S})$. The set $E^{BCE}_{\theta,x}$ is convex, since it is defined by equalities and inequalities that are linear in the equilibrium distribution.

In order to capture the BCE predictions on observed behavior, we consider the marginal with respect to players' actions of a BCE distribution ν.

Definition 3. *(BCE Prediction)* The BCE distribution ν induces a distribution over outcomes p_ν defined as:

$$p_\nu(y) = \int_{\varepsilon\in\mathcal{E}} d\nu(y, \varepsilon).$$

The observable implications of BCE behavior in a structure characterized by $(\theta, \underline{S})$ are described by the prediction correspondence $Q^{BCE}_\theta : X \rightrightarrows \Delta^{|Y|-1}$, defined as:

$$Q^{BCE}_\theta(x) = \left\{ p \in \Delta^{|Y|-1} : \exists \nu_x \in E^{BCE}_{\theta,x} \text{ such that } p = p_{\nu_x} \right\}.$$

Since the set $E^{BCE}_{\theta,x}$ is convex, any convex combination of BCE distributions is also a BCE distribution. Therefore, $Q^{BCE}_\theta(x)$ captures equilibrium predictions with no restrictions on equilibrium selection.

Figure 2 shows the set of BCE outcomes for the setup in Example 1. Panel (A) shows that BCE imposes weaker restrictions on equilibrium behavior: the sets of BNE predictions obtained under a specific assumption on information are all contained in the set of BCE predictions. Panel (B) illustrates instead that BCE predictions are still a relatively small subset of all possible outcomes, represented by the simplex.

Figure 2: BCE and BNE Predictions
Comp

BCE Predictions Complete Information Minimal Information Predictions Privileged Information

Figure 2: Panel (A)

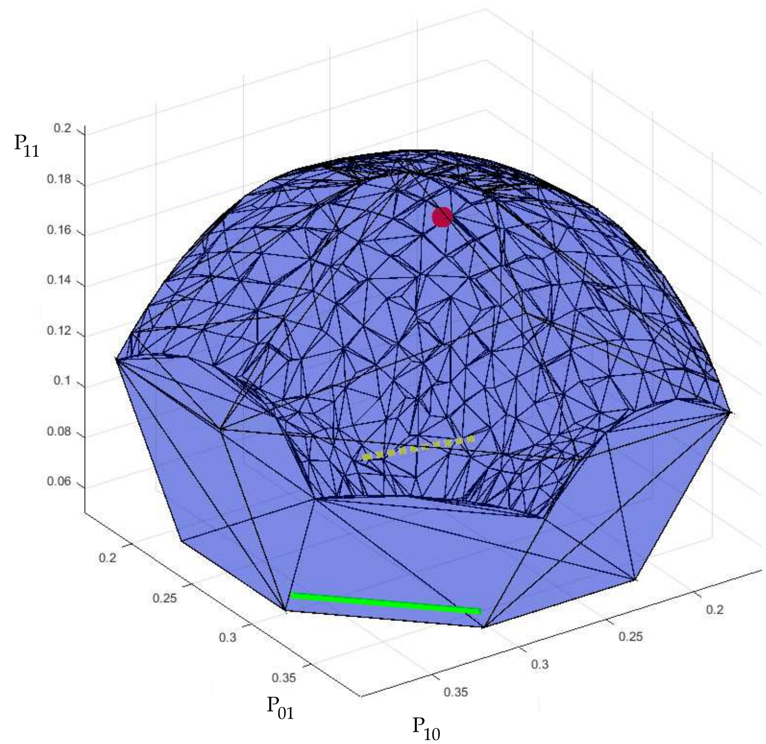
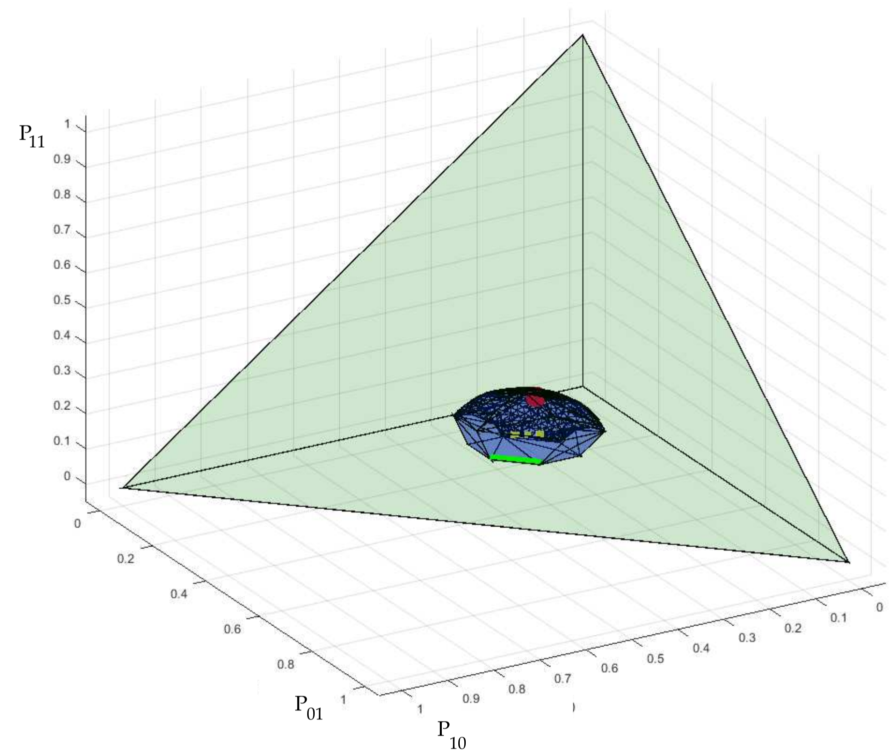


Figure 2: Panel (B)



In this figure we compare BCE predictions Q_θ^{BCE} with the BNE predictions $Q_{\theta,S}^{BNE}$ obtained under different information structures S for the two-player entry game described in Example 1. The axes represent probabilities of different outcomes P_y. The solid green line represents the pure strategy Nash predictions of the complete information game $Q_{\theta,\overline{S}}^{BNE}$. The red dot represents the BNE prediction of the minimal information game $Q_{\theta,\underline{S}}^{BNE}$. The yellow dotted line represents the privileged information predictions Q_{θ,S^P}^{BNE}.

In panel (A), we represent the set of BCE predictions, containing the BNE predictions under different restrictions on information.

In panel (B), we represent the unit simplex, with the set of BCE predictions inside.

We are most interested in the implications of adopting BCE behavior for identification. Under Assumptions 1 and 2 the behavioral assumption of BCE, the identified set of parameters in this class of games is defined by:

$$\Theta_I^{BCE} = \left\{\theta \in \Theta \text{ such that } P_{y|x} \in Q_\theta^{BCE}(x)\ P_x - a.s.\right\}. \tag{4.2}$$

4.2 BCE Identification

Bergemann and Morris (2013, 2015) establish the robust prediction property of BCE. In our setup, this property translates into the equivalence, for any given θ, of the BCE predictions Q_θ^{BCE} and the union of BNE equilibrium predictions $Q_{\theta,S}^{BNE}$ taken over $S \in \mathcal{S}_0$. Figure 2 illustrates this result by representing the polytope Q_θ^{BCE} as well as the sets of BNE predictions $Q_{\theta,\mathrm{S}}^{BNE}$ for the three information structures described in Example 2, and for the payoff structure described in Example 1. We show that a robustness result holds also for identification. In particular, the identified set we obtain under BCE is equivalent to our object of interest, $\Theta_I^{BNE}(\mathcal{S}_0)$. For any restriction on information $\mathcal{S} \subseteq \mathcal{S}_0$, let:

$$\Theta_I^{BNE}(\mathcal{S}) = \left\{\theta \in \Theta | \exists\ S \in \mathcal{S} \text{ such that } P_{y|x} \in Q_{\theta,S}^{BNE}(x)\ P_x - a.s.\right\}$$

be the identified set of parameters consistent with BNE behavior. We have then:

Proposition 1. *(Robust Identification) Let Assumptions 1 and 2 hold. Then,*

1. *the identified set under BCE behavior contains the true parameter value, $\theta_0 \in \Theta_I^{BCE}$, and*
2. $\Theta_I^{BCE} = \Theta_I^{BNE}(\mathcal{S}_0)$.

Proof. See Appendix B. □

Proposition 1 offers a foundation for the use of the BCE behavioral assumption for identification on robustness grounds. The adoption of BCE allows for the characterization of a set of parameters consistent with equilibrium behavior and a common prior, with minimal assumptions on information. The object $\Theta_I^{BNE}(\mathcal{S}_0)$, impossible to characterize when relying on BNE behavior, is easily defined by relying on the weaker solution concept of BCE. Although the proposition shows that for all parameters $\theta \in \Theta_I^{BCE}$ there must be an information structure S such that $\theta \in \Theta_I^{BNE}(\{S\})$, it is not necessarily true that every restriction on information $\mathcal{S}$ results in implications on parameters that will be considered in the set Θ_I^{BCE}. In fact, we show in Section 6 that the restriction $\mathcal{S}$ could be falsified, so that $\Theta_I^{BNE}(\mathcal{S}) = \emptyset$. In the next subsection we present a computable characterization of the BCE identified set.

4.3 Support Function Characterization of the Identified Set

We have argued in Proposition 1 that Θ_I^{BCE} is the set of all parameters compatible with the observable and the nonparametric class of information structures $\mathcal{S}_0$, and that focusing on this set can provide informative bounds and prevent misspecification. To estimate and compute Θ_I^{BCE}, however, we need to provide a tractable characterization, since it's not immediately obvious how to compute the set as defined in equation (4.2).

For every $x \in X$, the set of BCE predictions $Q_\theta^{BCE}(x)$ is a convex set. Convexity of the set of predictions follows directly from the definition of BCE. Hence, we can represent $Q_\theta^{BCE}(x)$ through its support function. This is similar to results in Beresteanu, Molinari and Molchanov (2011).[18]

Let $h\left(Q_\theta^{BCE}(x);\cdot\right): \mathbb{R}^{|Y|} \to \mathbb{R}$ denote the support function of the set $Q_\theta^{BCE}(x)$:

$$h\left(Q_\theta^{BCE}(x);b\right) = \sup_{p\in Q_\theta^{BCE}(x)} b'p.$$

The support function provides a representation of the set of predictions:

$$p \in Q_\theta^{BCE}(x) \Longleftrightarrow \left\{b'p \leq h\left(Q_\theta^{BCE}(x);b\right) \forall\, b \in B\right\},$$

where B is the unit sphere in $\mathbb{R}^{|Y|}$. We have then:

$$\begin{aligned}\Theta_I^{BCE} &= \left\{\theta \in \Theta \middle| P'_{y|x}b \leq h\left(Q_\theta^{BCE}(x);b\right) \forall\, b \in B\, P_x - a.s.\right\}\\ &= \left\{\theta \in \Theta \middle| \max_{b\in B} \min_{p\in Q_\theta^{BCE}(x)} \left[b'P_{y|x} - b'p\right] = 0\, P_x - a.s.\right\}.\end{aligned} \tag{4.3}$$

The computation of this object is simplified by the fact that the optimization problem is linear in the variable p. Appendix A provides computational details.

4.4 Inference

Suppose now that we observe an iid sample of size t of players choices and covariates $\left\{y_j, x_j\right\}_{j=1}^n$. To apply existing inferential methods, we assume that the set of covariates X is discrete.[19] We perform inference following an extremum estimation approach by re-defining the identified set characterized in (4.3) as the set of minimizers of a non-negative criterion function G, or

$$\Theta_I^{BCE} = \{\theta \in \Theta | G(\theta) = 0\},$$

[18] Appendix D describes how our characterization of the identified set maps into their framework.

[19] While several recent methods for inference in partially identified models such as Andrews and Shi (2013) do not require discrete covariates, they prove to be too computationally costly for the estimation of our model.

for

$$G(\theta) = \int_X \sup_{b \in B} \left[b' P_{y|x} - h\left(Q_\theta^{BCE}(x); b \right) \right] dP_x\{x\}.$$

The sample analogue of the population criterion function is:

$$G_n(\theta) = \frac{1}{n} \sum_{j=1}^{n} \sup_{b \in B} \left[b' \hat{P}_{y|x_j} - h\left(Q_\theta^{BCE}\left(x_j\right); b \right) \right],$$

where $\hat{P}_{y|x_j}$ is the empirical frequency of strategy profile y in observations with covariates $x = x_j$. The population criterion function inherits a smoothness property from the continuity of the payoff function and the upper hemi-continuity of the equilibrium correspondence, so that we can obtain a consistent estimator of the identified set as in Chernozhukov, Hong and Tamer (2007):

Proposition 2. *Assume that:*

(i) the map $\theta_\pi \to \pi_i^{x,\theta_\pi}(y, \varepsilon_i)$ is continuous for all i, x, y and ε_i, the quantity

$$|\pi_i^{x,\theta_\pi}(y_i, y_{-i}, \varepsilon_i) - \pi_i^{x,\theta_\pi}\left(y_i', y_{-i}, \varepsilon_i\right)|$$

is bounded above, and the map $\theta_\varepsilon \to F(\cdot; \theta_\varepsilon)$ is continuous for all ε;

(ii) the following uniform convergence condition holds: $\sup_\Theta \sqrt{n}|G_n(\theta) - G(\theta)| = O_p(1)$;

(iii) the sample criterion function G_n is stochastically bounded over Θ_I at rate $1/n$.

Then, the set $\hat{\Theta}_I = \{\theta \in \Theta | \, nG_n(\theta) \leq \log n\}$ is a consistent estimator of Θ_I^{BCE}.

Proof. See Appendix B. □

The previous proposition shows that our setup satisfies condition C.1 in Chernozhukov, Hong and Tamer (2007), and we proceed to apply their methods. We are interested in building confidence regions $\mathcal{L}_n$ for the identified set Θ_I^{BCE}, with the property that:

$$\liminf_{n \to \infty} P\{\Theta_I \subseteq \mathcal{L}_n\} \geq 1 - \alpha.$$

We base these regions on level sets of the sample criterion function:

$$\mathcal{L}_n(c) = \{\theta \in \Theta | \, G_n(\theta) \leq c\}.$$

The confidence set can be characterized as in Ciliberto and Tamer (2009) with a likelihood-ratio type statistic:

$$L_n := \sup_{\theta \in \Theta_I} G_n(\theta),$$

so that :

$$\{\Theta_I \subseteq \mathcal{L}_n(c)\} \iff \{L_n \leq c\},$$

and we can obtain the desired confidence region as $\mathcal{L}_n(\hat{c})$, where $\hat{c}$ set at the value of the alpha-quantile of the asymptotic distribution of L_n. Appendix A describes the details of how we obtain $\hat{c}$ and compute $\mathcal{L}_n(\hat{c})$.

5 Identifying Power of BCE

We address in this section the issue of the informativeness of Θ_I^{BCE}, the set identified under BCE behavior. Whenever variation in covariates allows the econometrician to observe games in which strategic considerations are negligible, several features of the model can be point identified. This identification strategy was first proposed by Tamer (2003) for games of complete information under the assumption of pure Nash Equilibrium behavior, but it still applies without restrictions on information and equilibrium selection.[20] To simplify our discussion, we restrict our attention to games with two actions and two players, and with payoffs linear in covariates. These assumptions are summarized in the following:

Assumption 3. *Let $|N| = 2$ and $Y = \{0,1\}^2$; let payoffs be:*

$$\pi_i^{x,\theta_\pi}(y,\varepsilon_i) = y_i\left(z'\beta^C + x_i'\beta_i^E + \Delta_{-i}y_{-i} + \varepsilon_i\right).$$

1. *Vectors of covariates are partitioned as $x = (x_1, x_2, z) \in X_1 \times X_2 \times \mathcal{Z} = X$, and the distribution P_x is such that x_i has everywhere positive Lebesgue density conditional on z, x_{-i}, for $i = 1,2$, and there exists no linear subspace E of $X_i \times Z$ such that $P_x(E) = 1$.*

2. *Payoff types $(\varepsilon_1, \varepsilon_2)$ are independent of covariates x, and distributed according to an absolutely continuous cdf $F(\cdot;\theta_\varepsilon)$, defined on $\mathcal{E} = \mathbb{R}^2$.*

Proposition 3. *Suppose the econometrician observes the distribution of the data $\left\{P_{y|x} : x \in X\right\}$, generated by BCE play of a game. Then, under Assumption 3,*

1. *payoff parameters β^C, β^E and Δ are point identified as in single-agent threshold crossing models; and*

2. *the structure implies bounds on the payoff type parameter θ_ε.*

Proof. See Appendix B. □

[20] For an alternative identification strategy in games of complete information, see Kline (2015).

The proposition shows that, whenever there are covariates' values for which one player has a dominant strategy, identification of payoffs proceeds as in single-agent binary choice models.[21] Since dominance only requires that players be rational and know their own payoffs, it applies to discrete games in which players have information on at least their own payoffs. While we do not expect the assumptions of the theorem to hold literally in applications, Proposition 3 indicates at least a source of variation that helps identification.

In Table 1 we show projections of Θ_I^{BCE} for different information structures S_0 and different levels of variation in covariates. We consider in particular the following discrete sets of covariates with increasing variance, with $X^{'} = \{-1, 0, 1\}$ and $X^{''} = \{-3, 0, 3\}$ to show variation in covariates at work in reducing the size of the identified set.

Our Proposition 3 implies that the model has identifying power on the joint distribution of payoff types. Consider now entry game with same payoffs as in the previous example, marginal distributions of payoff types F_i logit with unit variance for $i = 1, 2$, and joint distribution of payoff types:

$$F(\cdot; \theta_\varepsilon) = \mathscr{C}(F_1, F_2; \rho),$$

where $\mathscr{C}$ denotes the Frank copula function. We set values of exogenous covariates at $X^{'}$. Table 2 shows projections of the identified sets.

6 Assumptions on Information and Identification

While we pursue identification under BCE behavior, the current approach of the literature is to restrict the class of admissible information structures. This is done by choosing $\mathcal{S} \subseteq \mathcal{S}_0$ such that $Q_{\theta,S}^{BNE}$ is analytically tractable for $S \in \mathcal{S}$, and focusing the analysis on the set $\Theta_I^{BNE}(\mathcal{S})$. For instance, seminal papers in the literature such as Bresnahan and Reiss (1991a), Berry (1992) and Tamer (2003) assume complete information, which corresponds to the restriction $\mathcal{S} = \left\{\overline{S}\right\}$. Conversely, other authors such as Seim (2006), Sweeting (2009) and Bajari et al. (2010), restrict $\mathcal{S}$ to the minimal information structure $\underline{S}$, in which signals τ^x are uninformative.

Ideally, the restriction imposed on the information structure $\mathcal{S}$ is *pointwise correct*, that is $\mathcal{S} = \{S_0\}$, or at least *correct* i.e. $S_0 \in \mathcal{S}$. In this case,

$$\Theta_I^{BCE} \supseteq \Theta_I^{BNE}(\mathcal{S}) \supseteq \Theta_I^{BNE}(\{S_0\}) \neq \emptyset,$$

where the first inclusion follows from Proposition 1. However, in typical applications there is little evidence on the nature of S_0. If instead $S_0 \notin \mathcal{S}$, the model is *misspecified*, and

[21] For results on the identification of single-agent threshold crossing models, see Manski (1988).

Table 1: Variation in Covariates and Θ_I^{BCE}

Panel (A): $X^{'}$	θ_0 (1)	$S_0 = \bar{S}$ (2)	$S_0 = \underline{S}$ (3)	$S_0 = S^P$ (4)
β^C	1	[.90,1.02]	[.81,1.11]	[.76,1.04]
β_1	1	[.92,1.09]	[.90,1.21]	[.91,1.17]
β_2	1	[.92,1.09]	[.90,1.21]	[.83,1.17]
Δ_1	-1	[-2.36,-.82]	[-1.48,-.78]	[-1.99,-.84]
Δ_2	-1	[-2.36,-.82]	[-1.48,-.78]	[-2.13,-.83]
Panel (B): $X^{''}$				
β^C	1	[.97,1.08]	[.87,1.08]	[.86,1.08]
β_1	1	[.96,1.05]	[.96,1.09]	[.96,1.06]
β_2	1	[.96,1.05]	[.96,1.09]	[.95,1.08]
Δ_1	-1	[-1.30,-.90]	[-1.23,-.90]	[-1.26,-.90]
Δ_2	-1	[-1.30,-.90]	[-1.23,-.90]	[-1.29,-.89]

This table reports projections of the identified sets for the parameters of the two players entry model with payoffs $\pi_i^{x,\theta_\pi}(y,\varepsilon_i) = y_i\left(z'\beta^C + x_i'\beta_i^E + \Delta_{-i}y_{-i} + \varepsilon_i\right)$ for $i = 1,2$ and ε_i iid standard Normal across players and markets. The identified sets are obtained under weak assumptions on information, and computed as projections of Θ_I^{BCE}. Column (1) reports the values of the true parameter vector θ_0, while columns (2), (3) and (4) report projections of Θ_I^{BCE} for different assumptions on the information structure S_0 that characterizes the game that generates the data. Panel (A) reports identified sets for data generated with x uniformly distributed in the set $X^{'}$, while Panel (B) reports identified sets with x uniformly distributed in the set $X^{''}$. Computational details are in Appendix A.

Table 2: Θ_I^{BCE} with Correlated Payoff Types

	θ_0 (1)	Proj. of Θ_I^{BCE} (2)
β^C	1	[0.67,1]
β_1	1	[0.75,1]
β_2	1	[0.87,1.17]
Δ_1	-1	[-1.02,-0.12]
Δ_2	-1	[-1.50,-0.65]
ρ	0.8	[0.56,0.88]

This table reports projections of the identified sets for the parameters of the two players entry model with payoffs $\pi_i^{x,\theta_\pi}(y,\varepsilon_i) = y_i\left(z'\beta^C + x_i'\beta_i^E + \Delta_{-i}y_{-i} + \varepsilon_i\right)$ for $i = 1,2$.
In this case, ε are distributed according to $F(\cdot;\theta_\varepsilon) = \mathscr{C}(F_1, F_2; \rho)$, where F_i is a Logit distribution, $\mathscr{C}$ denotes the Frank copula function, and ρ is a correlation parameter. The identified sets are obtained under the assumption that the information structure is complete information, and computed as projections of Θ_I^{BCE} for each parameter of the model. The information structure of the game that generates the data is complete information so $S_0 = \overline{S}$, and covariates x are uniformly distributed in the set $X^{'}$. Column (3) reports values of the true parameter vector θ_0, while column (4) reports projections of the identified set.

Computational details are in Appendix A.

Table 3: Identified set when $\mathcal{S} = \left\{\overline{S}\right\} = \{S_0\}$

	θ_0 (1)	Proj. of $\Theta_I^{BNE}\left(\left\{\overline{S}\right\}\right)$ (2)
β^C	1	[0.96,1.02]
β_1	1	[0.96,1.04]
β_2	1	[0.96,1.04]
Δ_1	-1	[-1.11,-0.97]
Δ_2	-1	[-1.11,-0.97]

This table reports projections of the identified sets for the parameters of the two players entry model with payoffs $\pi_i^{x,\theta_\pi}(y,\varepsilon_i) = y_i\left(z'\beta^C + x_i'\beta_i^E + \Delta_{-i}y_{-i} + \varepsilon_i\right)$ for $i = 1,2$.

Payoff types ε_i are iid standard Normal across players and markets, and the identified sets are obtained under the assumption that the information structure is complete information, and computed as projections of $\Theta_I^{BNE}\left(\left\{\overline{S}\right\}\right)$ for each parameter of the model. The information structure of the game that generates the data is complete information so $S_0 = \overline{S}$, and covariates x are uniformly distributed in the set X'. Column (1) reports values of the true parameter vector θ_0, while column (2) reports projections of the identified set.

Computational details are in Appendix A.

one of the following three scenarios will occur. Either (i) the misspecification has benign consequences, that is $\theta_0 \in \Theta_I^{BNE}(\mathcal{S})$, or (ii) $\theta_0 \notin \Theta_I^{BNE}(\mathcal{S}) \neq \emptyset$, that is misspecification results in a nonempty identified set, selecting arbitrarily a region of $\Theta_I^{BNE}(\mathcal{S}_0)$ that does not contain θ_0, or (iii) the model is *falsified* by the data, that is $\Theta_I^{BNE}(\mathcal{S}) = \emptyset$. In this case, no parameter θ can rationalize the observables given the restriction on information $\mathcal{S}$.

We can compute the set $\Theta_I^{BNE}(\mathcal{S})$ identified under restrictions on information using the same model as in section 5.1 with covariates uniformly distributed in the set X'. When the restrictive assumptions are pointwise correct, that is they correspond to the true information structure in the data generating process S_0, there are gains in identifying power. For instance, if $\mathcal{S} = \{\underline{S}\} = \{S_0\}$ the model is point identified. If $\mathcal{S} = \left\{\bar{S}\right\} = \{S_0\}$, we obtain results in Table 3.

However, if the model is misspecified, that is we assume $\mathcal{S} = \left\{\overline{S}\right\}$ but $S_0 = \underline{S}$, or $\mathcal{S} = \{\underline{S}\}$ but $\overline{S} = S_0$, in this simple model the set $\Theta_I^{BNE}(\mathcal{S})$ is empty. No parameter values can reconcile the data with the structure of the model, which makes the model falsified by the data.[22] In Proposition 1 we establish that the identified set under BCE contains only those parameters for which there exists an information structure and a corresponding BNE that generate predictions that match the data. If no such values exist for $\mathcal{S} \subseteq \mathcal{S}_0$, then all the information structures $S \in \mathcal{S}$ are falsified. When estimation is performed under a misspecified assumption, we can expect biased estimates. We show in the following subsection how assumptions about information can affect identification.

6.1 Impact of Strong Assumptions on Identification

We consider a binary, 2-player entry game as described in Example 1, with one payoff parameter and no covariates. In this game, $|N| = 2$ and $Y = \{0,1\}^2$, and payoffs are:

$$\pi_i(y, \varepsilon_i) = y_i(\Delta y_{-i} + \varepsilon_i),$$

with F_i distribution of ε_i uniform on the interval [-1,1]. The parameter Δ belongs to the interval $\Theta = [-1,0]$, restricted not to exceed in absolute value the maximum payoff type.

Restrictive assumptions on information have substantial impact on identification in this game. To see this more clearly, consider the non-sharp identified set:

$$\tilde{\Theta}_I^{BNE}(\mathcal{S}) = \left\{\Delta \in \Theta | \exists\, S \in \mathcal{S}_0 \text{ such that } P_y(1,1) \in Q_{\theta,S}^{BNE}\right\},$$

obtained by using only the observable probability of the outcome $(1,1)$. Under the assumption of complete information, that is $\mathcal{S} = \left\{\overline{S}\right\}$, if we restrict for analytical convenience the equilibrium correspondence to only allow for pure strategy equilibria, we can immediately recover the parameters $\Delta \in \tilde{\Theta}_I^{BNE}$ by solving:

$$\begin{aligned} P_y(1,1) &= (1 - F_i(-\Delta))^2 \\ &= \left(\frac{1+\Delta}{2}\right)^2. \end{aligned}$$

If instead we adopt the restriction of private signals, that is $\mathcal{S} = \{\underline{S}\}$, we have that the symmetric BNE characterized by $s_i(y_i = 1) = \int s_i(y_i = 1|\varepsilon_i)\, dF(\varepsilon_i)$ solves the equation:

$$\begin{aligned} s_i(y_i = 1) &= 1 - F_i(\overline{\varepsilon}), \\ &= \frac{1 - \overline{\varepsilon}}{2}, \end{aligned}$$

[22] As pointed out by Ponomareva and Tamer (2011), estimating a misspecified model may result in tight bounds, which however can be far from the true value. Moreover, in this case, we do not expect that the estimated parameter sets under the falsified restriction on information will be contained in the confidence set estimated under BCE.

where the threshold level $\overline{\varepsilon}$ is pinned down by the equation:

$$\overline{\varepsilon} + \Delta\left(\frac{1-\overline{\varepsilon}}{2}\right) = 0,$$

so that $\overline{\varepsilon} = -\frac{\Delta}{2-\Delta}$, and

$$s_i\left(y_i = 1\right) = \frac{1}{2-\Delta},$$

and one of the corresponding implications of equilibrium on observable behavior is:

$$P_y(1,1) = \left(\frac{1}{2-\Delta}\right)^2.$$

Under the assumption that $\mathcal{S} = \left\{S^P\right\}$, in equilibrium player 1 knows when player 2 will enter. There are in this case a multiplicity of equilibria, in which player 2 has a threshold strategy characterized by the value $\overline{\varepsilon}_2$, and player 1 will always enter if $\varepsilon_1 > -\Delta$ and enter only if $\varepsilon_2 < \overline{\varepsilon}_2$ whenever $0 < \varepsilon_1 < -\Delta$. For such a strategy to be an equilibrium, it must be that:

$$\varepsilon_2 + \Delta\left(1 - F_i\left(-\Delta\right)\right) \geq 0,$$

for all $\varepsilon_2 \geq \overline{\varepsilon}_2$, and

$$\varepsilon_2 + \Delta\left(1 - F_i\left(0\right)\right) \leq 0,$$

for all $\varepsilon_2 \leq \overline{\varepsilon}_2$, which in turn implies that $\overline{\varepsilon}_2 \in \left[-\frac{\Delta(1+\Delta)}{2}, -\frac{\Delta}{2}\right]$, and:

$$P_y\left\{(1,1)\right\} \in \left[\frac{(2-\Delta)(1+\Delta)}{4}, \frac{(2-\Delta(1+\Delta))(1+\Delta)}{4}\right].$$

Notice that for player 2 equilibrium behavior need not depend on the uninformative signal τ_2.

Suppose now that the true parameter is $\Delta_0 = -0.5$. For a certain value of $P_y(1,1)$ observed in the data, different restrictions on the information structure will yield different identified sets. Table 4 summarizes the identified set $\tilde{\Theta}_I^{BNE}\left(\mathcal{S}\right)$ under different combinations of $\mathcal{S}$ and S_0. When $S_0 = S^P$, we assume that data are generated by the threshold strategy that lies in the middle of the continuum of equilibria, that is with $\overline{\varepsilon}_2 = 3/16$.

From Table 4, it appears immediately how overstating the amount of information available to players leads to an identified parameter that is lower, in absolute value than the true parameter value. Intuitively this is because the probability that both players enter as predicted by the model depends on Δ and on players' degree of certainty that their opponent will also enter. Overstating the amount of information leads the econometrician to impose on players beliefs about their opponent's actions that are too precise. For this level of precision in beliefs, in turn, the model rationalizes the observed $P_y\left\{(1,1)\right\}$

Table 4: Identified Set $\Theta_I^{BNE}(\mathcal{S})$ under different $\mathcal{S}$ and S_0 and $\Delta_0 = -0.5$

S_0:		$\overline{S}$	S^P	$\underline{S}$
$\mathcal{S}$:	$\overline{S}$	$\{-0.50\}$	$\{-0.36\}$	$\{-0.2\}$
	S^P	$[-0.82, -0.72]$	$[-.54, -0.47]$	$[-0.29, -0.26]$
	$\underline{S}$	$\{-2\}$	$\{-1.14\}$	$\{-0.50\}$

This table reports the identified sets for the parameters of the two players entry model with payoffs $\pi_i(y, \varepsilon_i) = y_i(\Delta y_{-i} + \varepsilon_i)$ for $i = 1, 2$ and ε_i iid $U[-1, 1]$ across players and markets. The identified sets $\tilde{\Theta}_I^{BNE}(\mathcal{S})$ are non sharp, and obtained under a restrictive assumption on information $\mathcal{S}$ (corresponding to different row) and a true information structures S_0 (corresponding to a different column). The true value of the parameter θ in the data generating process is $\Delta_0 = -0.5$. When $\mathcal{S}_0 = S^P$, we assume that data are generated by the threshold strategy that lies in the middle of the continuum of equilibria, that is with $\overline{\varepsilon}_2 = 3/16$.

with a value of Δ that is lower, in absolute value, than Δ_0. This type of attenuation bias has already been recognized in the literature on single agents decisions in Bergemann and Morris (2013), and in the context of dynamic games by Aguirregabiria and Magesan (2015).

This example shows that misspecification of the information structure can result in significant bias in the identified parameters. While in this case the direction of the bias is intuitive, in more complex games it will not be as simple to determine. Estimation of games under the assumption of BCE allows to avoid this bias.

7 Application: the Impact of Large Malls on Local Supermarkets

Regional and super-regional malls, typically built around a grocery "anchor,"[23] are a relatively new phenomenon in Italy, having gained popularity mostly in the last fifteen years. The emergence of large malls has sparked a debate on their economic impact on local retailers. We focus on the impact of malls on local supermarkets. In the view of their critics, large supermarkets in malls represent a formidable competitor to local supermarkets, and end up decreasing local stores' profits.[24] The entry of large shopping centers might gen-

[23]While these anchor supermarkets are not regarded by industry experts as very successful in their own right, they receive rent subsidies from mall operators, since they are believed to attract consumers that shop at other stores in the mall.

[24]A recent survey of retailers finds that shop owners rank the emergence of large malls as the second factor that most affected their business in the last five years. See Confesercenti press

erate a market structure with fewer supermarkets in operation or even a local monopoly, thus hurting those consumers that would most benefit from the availability of local stores and from the presence of smaller store formats. Advocates of this view propose tighter restrictions on entry by large malls, especially in markets that cannot support other stores. Others contend that format differentiation results in little competition between local supermarkets and large mall anchors, and the economic activity linked to large malls might generate spillovers that strengthen local demand. According to this view, regulation that restricts the number of malls in operation would be ultimately harmful to consumers.

We quantify the effect of malls on supermarket industry groups by estimating a game theoretic model in which industry players decide strategically whether to operate stores in local grocery markets, and entry of large malls is exogenous to the dynamics of competition in grocery markets. This calls for the use of a model that explicitly recognizes the strategic nature of supermarket industry players, as well as the heterogeneity in the underlying determinants of their entry behavior.

It is not obvious which informational environment prevails in this setting. In principle, local players might possess a superior knowledge of grocery markets and thus be in a good position to evaluate the profitability of entry by their rivals. Alternatively, foreign firms might leverage the long history of their competitors to have a more precise outlook on their profitability. While the no-regret feature of pure Nash equilibria in games of complete information is often viewed as a plausible characteristic of the long-run industry snapshot that is captured with a static model, this argument is not particularly strong for this industry in the period we study. The year of our cross-sectional data, 2013, comes at the end of a 15-years long period of growth in the industry, which was sparked by a overhaul of regulation in 1998. However, both accounting data and trade press sources indicate that many firms and individual outlets have been unprofitable for several years in the period, so that regret for not having anticipated competition in local markets cannot be ruled out.

We estimate the model and conduct counterfactual analysis relying on weak assumptions on information, using the method we have developed in the previous sections of this paper. We also estimate the game under the assumption of complete information, and evaluate the consequences on estimates of imposing more restrictive assumptions. To reduce complexity, we abstract from the multi-store nature of national players' decisions, from geographical location and store format choice within a market.[25]

release at http://www.confesercenti.it/blog/imprese-dei-centri-storici-sondaggio-confesercenti-swg-fisco-ha-inciso-negativamente-per-8-su-10/.

[25]For studies of empirical entry games in retail industries that model these aspects, see for instance Jia (2008), Ellickson, Houghton and Timmins (2013) and Nishida (2014) for multi-store firms and economies of density, Mazzeo (2002), Seim (2006) and Datta and Sudhir (2013) for endogenous product and location choice. See Aguirregabiria and Suzuki (2015) for a recent survey of structural models of competition in retail.

7.1 Data and Institutional Details

We have data on store presence and characteristics for all supermarkets in Northern and Central Italy at the end of 2013, obtained from the market research firm IRI. We complement these with hand-collected information on which supermarkets are in malls and on mall size, obtained from public online directories. We focus exclusively on Northern and Central Italy because the structure of grocery markets in the South differs markedly, with traditional stores and open-air markets still playing a very important role, and relatively few instances of large malls. Data on population and demographics are obtained from the 2011 official census, while data on (tax) income at the municipality level are available for 2013 from the Ministry of Economy and Finance.

Market Definition and Industry Players

Defining the relevant market in this industry requires specifying both which store formats are direct competitors, and the geographical units that constitute separate local grocery markets. The Italian Antitrust Authority distinguishes between smaller stores with floor space up to 1,500 m^2 (16,146 ft^2) and stores above this threshold, pointing out that stores in these two categories differ fundamentally in location, assortments, and regulation they are subject to (see AGCM 2013, Viviano et al. 2012). From the point of view of retail groups, larger stores are more efficient, and have seen the fastest growth in this industry in the last 15 years, suggesting that firms and consumers prefer these modern formats. Price surveys also indicate that larger stores offer, on average, lower prices. We consider stores with a floor space of at least 1,500 m^2 (16,146 ft^2)[26] as the relevant market for our study, since these stores seem to be most relevant to economic outcomes in a grocery market. These are also the stores that are most likely to be in direct competition with the grocery anchors present in large malls. We define large malls as shopping centers including at least 50 independent shops.

No existing administrative unit provides a natural way of defining local grocery markets in Italy, and in antitrust cases the extension of geographical markets is determined on an ad-hoc basis. Evidence collected by various European Antitrust Authorities indicates that most consumers travel little to do their grocery shopping,[27] and evidence from marketing research points to the fact that supermarkets make most of their revenues from customers living in a 2 km (1.24 mi.) radius, while large shopping malls attract shoppers who drive up to 30 minutes. Since commuting patterns capture consumers' daily movements better than administrative units, we start from the geographical commuting areas

[26]For comparison, median store size for US supermarkets was 46,500 ft^2 in 2013 according to Food Marketing Institute, an industry association.

[27]UK's Competition Commission considers all large stores in a radius of 10-15 minutes by car to belong to the same market.

as defined by ISTAT, the national statistical agency, and split commuting areas that are too large.[28] We exclude large cities with more than three hundred thousand inhabitants in a municipality, as the density of highly urban areas makes it very hard to separate distinct markets. We are left with 484 local grocery markets. We report summary statistics for these markets in Table 5, considering separately markets with large malls and markets with no large malls. The latter are systematically smaller, have a slightly lower per capita income, and have on average one supermarket.

The Italian supermarket industry is quite diverse, as it includes cooperative firms, independent local supermarket groups, and foreign firms. Networks of consumers' and retailers' cooperatives, mainly operating under the signs of Coop Italia and Conad, all affiliated with the national umbrella organization Legacoop, have the largest market share and play an important role. Despite their formal organizational form and historical origins, they are managed efficiently and their behavior can be assimilated to that of their profit maximizing competitors.[29] There are then several national groups, all based in the North of the country, which own and operate networks of relatively large stores in partially overlapping geographical areas. Based on IRI data, Esselunga, Bennet, PAM, Finiper and Selex are the groups that have more than 2.5% market share in 2013. Two large French retail multinationals, Auchan and Carrefour, have also entered the Italian market both via acquisitions of local supermarket groups, and by opening new stores. They operate mainly large format stores, but have not been very profitable in the Italian market.

A particular feature of the Italian supermarket industry is that it is subject to extensive regulation, and entry in local markets can be delayed significantly by zoning and other laws. Schivardi and Viviano (2010) exploit geographical variation in how the 1998 retail liberalization reform is implemented, to show that this regulation has an important impact on the industry. We consider a static setup, capturing the cross-section of the industry in 2013, and argue that while regulation matters, it is unlikely to block entry altogether in a local market. While regulation may increase entry costs, in our modeling setup all players that find profitable to enter a market, are eventually able to do so.[30]

[28]We split the commuting area along municipality borders if it contains more than two towns that have at least fifteen thousand inhabitants, and are in a radius of 20 minutes of driving distance.

[29]Bentivogli and Viviano (2012) find in a survey of Italian cooperatives that their strategic response to changes in the economic environment is substantially similar to that of their privately owned counterparts.

[30]See for instance Schaumans and Verboven (2008) for a strategic model of entry with binding entry restrictions.

Table 5: Demographics of Local Grocery Markets

Variable	Mean	Std. Dev.	Median	Max	Min
Large Mall in Market	0.130	0.337	0	1	0
421 Markets with no Large Malls:					
Population	44,629.22	40,341.88	31,730	297,510	3,276
Surface, in km^2	329.90	242.72	275.72	1,969.64	25.19
Tax Income Per Capita, in EUR	13,223.8	1,730.34	13,204.92	18,288.90	8,020.68
# of Supermarkets	1.46	1.95	1	16	0
# of Players in Market	0.85	0.93	1	3	0
63 Markets with Large Malls:					
Population	117,614.10	56,195.42	103,925	249,852	35,768
Surface, in km^2	447.84	377.92	359.95	2,243.54	95.33
Tax Income Per Capita, in EUR	14,411.47	1,650.48	14,475.88	18,627.36	10,333.89
# of Supermarkets	3.77	2.89	3	13	0
# of Players in Market	1.58	0.87	2	3	0

This table reports summary statistics on the geographical grocery markets we consider in our analysis. The variable Large Mall in market is a dummy that equals one if a shopping center with at least 50 stores and a grocery anchor is present in the geographical market at the end of 2013. The supermarket industry players we consider are (i) cooperative groups Coop Italia and Conad, (ii) Italian independent supermarket groups, (iii) French supermarket groups Auchan and Carrefour. We code the entry variable as equal to one if the supermarket industry player has at least one supermarket with a floor space of 1500 m^2 or greater in operation at the end of 2013 in a geographical market.

Descriptive Regressions

To gain insight on the impact of large malls on grocery markets, we estimate descriptive linear regressions and ordered probit models.[31] The dependent variable is either the number of supermarkets in a geographical market or the number of supermarket industry players operating in a market. The coefficient estimates we obtain, reported in Table 6, point to a small and negative covariation between market structure outcomes and the presence of large malls in a grocery market. These regressions suggest a weak negative effect of malls on supermarket industry outcomes, but do not shed light on the potential differences in the impact of large malls on the behavior of different industry groups. Moreover, the counterfactual market structure that would emerge if malls were to be removed from some geographical grocery markets also depends on the competitive effect that supermarket industry groups have on each other's entry decisions.

Probit regressions that model the probability entry of supermarket groups as a function of market demographics, presence of large malls, and the entry behavior of competitors are a first step to describe heterogeneity in supermarket groups' profits and competitive effects. We consider separately the three types of players in this industry, cooperatives, independent Italian groups and French groups, and estimate the model:

$$y_{i,m} = 1\left\{ x_m' \beta_i + \sum_{j \neq i} y_{j,m} \Delta_j + \varepsilon_{i,m} \geq 0 \right\}, \tag{7.1}$$

where the outcome $y_{i,m} \in \{0,1\}$ denotes presence in market m in 2013 for industry player i, x_m are market level covariates, and $\varepsilon_{i,m}$ is a firm-market level unobservable, assumed to be iid Normally distributed in the probit regression. In the specifications we estimate, vectors x_m include a constant, measures of market size, and dummies for the presence of large malls in a market. We report in Table 7 estimates for a model in which market size is captured by market population times log income per capita, but results are very similar when considering different measures of market size. We control either for a dummy that indicates market in the home regions of each player, or for a full set of region dummies.

Unlike the results of Table 6, coefficient estimates of these probit regressions cannot be interpreted causally. In fact, $\varepsilon_{i,m}$ and $y_{j,m}$ are likely to be correlated in our data. The positive signs of competitive effects parameters are due to endogeneity, as players enter markets that have unobservably high profitability. There is some indication in these estimates of a strong negative covariation between the presence of large malls in the market and the decision of cooperatives to enter or stay in a market, while for the other players the covariation between entry and the presence of malls is smaller. These results do not

[31] The ordered probit model is equivalent to the specification of Bresnahan and Reiss (1991b). It can be interpreted as a game theoretic model in which players have the same payoffs and complete information.

Table 6: Regressions of Market Structure on Presence of Large Malls

Variable	# of Supermarkets		# of Players in Market	
	(1)	(2)	(3)	(4)
Large Mall in Market	-0.437	-0.222	-0.150	-0.242
	(0.278)	(0.165)	(0.145)	(0.175)
Market Size	3.764***	2.658***	1.213***	1.766***
	(0.236)	(0.158)	(0.109)	(0.143)
Constant	0.167		0.022	
	(0.378)		(0.230)	
N	484	484	484	484
R^2	0.677	0.255	0.434	0.225
Model	Linear Regression	Ordered probit	Linear Regression	Ordered probit

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

This table reports the coefficient estimates and standard errors from linear regressions (columns (1) and (3)) and ordered probit models (columns (2) and (4)) in which the dependent variable is the number of supermarkets with a floor space of 1500 m^2 or greater (columns (1) and (2)), or the number of supermarket industry players (columns (3) and (4)), present in a geographical market at the end of 2013. The supermarket industry players we consider are (i) cooperative groups Coop Italia and Conad, (ii) Italian independent supermarket groups, (iii) French supermarket groups Auchan and Carrefour. We code the entry variable as equal to one if the supermarket industry player has at least one supermarket with a floor space of 1500 m^2 or greater in operation at the end of 2013 in a geographical market. The variable Large Mall in market is a dummy that equals one if a shopping center with at least 50 stores and a grocery anchor is present in the geographical market at the end of 2013. The variable market size is the product of population times log of tax income per capita in the geographical market. All regressions include fixed effects for the 13 administrative regions included in our analysis. Values of R^2 refer to pseudo-R^2 for the ordered probit regressions.

Table 7: Probit Regressions of Entry

Variable	COOP Entry		IT Entry		FR Entry	
	(1)	(2)	(3)	(4)	(5)	(6)
Large Mall in Market	-0.686***	-0.722***	-0.154	-0.182	0.111	0.098
	(0.238)	(0.243)	(0.233)	(0.248)	(0.213)	(0.220)
Market Size	1.575***	1.640***	1.323***	1.644***	0.828***	0.882***
	(0.215)	(0.225)	(0.224)	(0.268)	(0.182)	(0.198)
Competitive Effects Δ_{ji}						
COOP Present in 2013			0.533***	0.452***	0.136	0.234
			(0.154)	(0.172)	(0.167)	(0.186)
IT Present in 2013	0.490***	0.569***			0.226	0.252
	(0.150)	(0.170)			(0.170)	(0.192)
FR Present in 2013	0.174	0.132	0.140	0.043		
	(0.187)	(0.196)	(0.186)	(0.206)		
Home Region Dummies						
COOP Home Region	1.174***					
	(0.157)					
IT Home Region			1.083***			
			(0.143)			
FR Home Region					0.906***	
					(0.157)	
Constant	-2.103***	-2.488***	-1.517***	-0.204	-2.026***	-1.722***
	(0.180)	(0.303)	(0.129)	(0.232)	(0.159)	(0.275)
N	484	481	484	447	484	447
Region FE	N	Y	N	Y	N	Y

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

This table reports coefficient estimates and standard errors from Probit regressions in which the dependent variable is entry of one supermarket industry player in a geographical grocery market. We consider three such players, (i) cooperative groups Coop Italia and Conad (COOP), (ii) Italian independent supermarket groups (IT), (iii) French supermarket groups Auchan and Carrefour (FR). We code the entry variable as equal to one if the supermarket industry player has at least one supermarket with a floor space of 1500 m^2 or greater in operation at the end of 2013 in a geographical market. The variable Large Mall in market is a dummy that equals one if a shopping center with at least 50 stores and a grocery anchor is present in the geographical market at the end of 2013. The variable market size is the product of population times log of tax income per capita in the geographical market. Competitive effects are coefficients on variables that describe entry of competing players. Home region dummies are equal to one for markets in regions where the players' headquarters are located. N is the number of observations. Columns (1), (3) and (5) do not include region fixed effects for the 13 administrative regions covered by our analysis, while columns (2), (4) and (6) do include these fixed effects.

change significantly when we control for players' presence in 2000 in unreported regressions. Our descriptive exploration of the data seems to indicate a weak effect of large malls on supermarket industry outcomes, with a considerable degree of heterogeneity in the way the presence of malls impacts the entry decisions of different supermarket industry players. To obtain a more reliable picture of the effects of large malls, and to perform counterfactual analysis taking into account the interdependent nature of entry decisions, we move to a game theoretic model. This requires a discussion of the assumptions we make on players' information.

7.2 Game Theoretic Model

We estimate now a static model of strategic interaction among players in the supermarket industry. Each player chooses whether to be present in each of the local geographical markets. This decision takes into account the exogenous characteristics of the market, the endogenous presence of the other players, and market-level characteristics unobserved to the econometrician. Payoffs from entry for player i in market m are:

$$\pi_i^{x,\theta_\pi} = x_m'\beta_i + \sum_{j \neq i} y_{j,m}\Delta_j + \varepsilon_{i,m},$$

while payoffs from staying out of the market are ized to zero. Market level covariates x_m in this specification include a measure of market size, a dummy for the presence of large malls in the market, and a home region dummy. The vector of unobservable payoff types $(\varepsilon_{i,m})_{i \in \mathcal{I}}$ is jointly distributed according to a distribution $F(\varepsilon;\rho)$. We assume that for every i, $\varepsilon_{i,m}$ has a Logistic distribution with zero mean and unit variance. The correlation of payoff types is modeled by a copula, with correlation between any couple of $\varepsilon_{i,m}$,$\varepsilon_{j,m}$ equal to ρ.[32] As discussed in Berry (1989), this specification of entry profits can either be interpreted as a "reduced form" assumption, justified on the grounds of parsimony and difficulties in modeling the nature of post-entry competition, or can be given a structural interpretation. A similar form of profits (up to monotone transformations), can in fact be derived under the assumptions of post-entry Cournot competition, constant elasticity demand and marginal costs identical across firms.[33] In order to reduce the complexity of the game theoretic model, we consider a game with three players, lumping together cooperatives, independent Italian groups and French groups.[34]

[32]While in principle we could allow for asymmetric correlations among players' payoff types, we do not do so in this application for data limitations.

[33]A Cournot model would also result in a complicated nonlinear expression for competitive effects. For parsimony, we assume instead that the entry of player j has a constant effect Δ_j on the payoffs of every other player i.

[34]For this model, for example, action $y_i = 1$ played by independent Italian groups represents the event of entry by at least one independent Italian supermarket chain.

We assume that the presence of large malls is exogenous to outcomes in the supermarket industry. This assumption is motivated by three arguments. First, since malls mostly host non-grocery shops, they have a "catchment area" that is substantially larger than the one of supermarkets, and attract consumers from a region that only partly coincides with local grocery markets. Second, malls require large areas for development, and this severely limits location choice in densely populated Italy, pushing developers to locate malls far from their ideal location. Third, obtaining permits for building large malls is a long process subject to shocks due to changing attitudes of local and national regulators, and can drive malls to locations that are only viable because consumer travel relatively far for non-grocery shopping. The exogenous entry assumption for the large player also follows Grieco (2014) and Ackerberg and Gowrisankaran (2006).

We estimate the model under the assumption that players observe their own shock, and can receive additional informative signals on the realizations of their competitors' payoffs. This approach nests not only for the information structures typical of the empirical games literature, but also allows for asymmetries in players' information structures that are relevant for this empirical setting but not compatible with existing models. As suggested by the probit estimates in Table 7, the behavior of French groups seems to be markedly different from that of the other players. Industry analysts point out that the rollout of French groups in the Italian market was largely unprofitable, and part of this can be attributed to a fundamentally different information structure, which might have left them with ex-post regret. This is hardly the case for independent Italian groups, which have been more profitable and might know local markets better. Hence, we adopt the assumption that the data are generated by BCE behavior, and estimate the confidence set for Θ_I^{BCE}. To compare our method to standard techniques, we also obtain a confidence set for parameters maintaining the assumption that data are generated by pure strategy Nash equilibrium behavior for the game of complete information.

Confidence Sets

Column (1) in Table 8 presents the estimated 95% confidence intervals of the identified set under the assumptions of BCE behavior. We report, for each parameter of the model, the lowest and highest value it takes in the confidence set.

Results for constant and coefficients on market size and home regions dummies are in line with the results of probit regressions, pointing to an intuitive positive correlation between market size and payoff from establishing presence in a local grocery market. The effects of operating in a home region is not significantly different from zero for any of the groups we examine. This is in contrast with the probit results in Table 7, which find evidence of positive effects of home region on the likelihood of presence of a player in a local market.

The evidence on the effect of the presence of large malls on the presence of supermarket groups is mixed. We do not find the effect of malls to be significantly different from zero for any of the players, although the confidence sets for the effect of large malls on the presence of French Supermarket groups (and, to a lesser extent, the confidence set for the same parameter in the payoff of Cooperatives) lie mostly on the negative real line. Probit estimates find largely positive competitive effects, as those models ignore the endogeneity of competitors' entry. The game theoretic model provides evidence that competitors' presence in a local market makes entry less profitable for supermarket groups. Projected confidence sets for the correlation parameter ρ are firmly positive, pointing to a robust correlation of payoff types among players in the same market.

In column (2) we report the projections of the 95% confidence intervals for the identified set under the assumptions of pure strategy Nash behavior and complete information. It is interesting to compare the estimates obtained under these more restrictive assumptions with the one obtained with our method. For the constant, market size parameters, and home region parameters the confidence sets corresponding to the two models are largely similar. The assumption of complete information makes a difference, however, for the estimates of the effect of large malls and of competitive effects. While the sign of the effect of malls is not identified under weak assumptions on information, with complete information this effect is estimated to be negative for two out of three players in the industry.

The importance of assumptions on information is most highlighted when we consider the estimates of the competitive effects that players have on each other. For all supermarket groups the interval estimated under weak assumptions on information is shifted to the left with respect to the one estimated under the assumption of perfect information. This means that the competitive effects estimated under the restrictive assumption of perfect information are mostly smaller, in absolute value, than those obtained with a model with weak assumptions on information. By assuming Nash behavior under perfect information, we impose that those players who decide to operate in a market with a competitor know that the competitor will be present. Instead, under BCE behavior, the equilibrium expectations allow for uncertainty about opponents' behavior. Hence, more negative values for the competitive effects parameters cannot be rejected, as they enter players' payoffs in expectation. These differences in estimated confidence sets are important, since they have an impact on the counterfactual predictions, and the respective policy implications, that the two models provide. The interval for the correlation parameter ρ is wider for the model with perfect information on payoff shocks: under the assumption of complete information we cannot reject very high values of correlation among payoff types. Weaker assumptions on information offer ways of rationalizing correlation in players' actions that are alternative to correlation in payoff shocks, thus allowing to reject very high correlation parameters.

It is not surprising that the set we estimate under the restrictive assumption of complete information is not nested in the estimated set under the weaker BCE assumption. Indeed, our robust identification result predicts that the complete information estimates are expected to be a subset of the BCE estimates only when the more restrictive assumption is not falsified by the data. If instead the more restrictive assumption is not supported by the data, there is no reason to expect estimates obtained under that assumption to lie inside the robust estimated set.[35]

Counterfactuals

We consider the counterfactual scenario in which regulation prevents the construction of large shopping malls in small markets. This counterfactual is a way to quantify how market structure is affected by the presence of large malls. We examine in particular the eight small geographical grocery markets that have a large shopping center but no supermarkets in the current market configuration, and compute predicted market structures under our parameter estimates. We then compute predicted outcomes of the entry game between supermarkets once the large shopping center is removed.

The multiplicity of parameter vectors that are included in the confidence region, as well as the multiplicity of equilibria given a certain parameter vector, imply that the model does not yield point predictions on which market structure will emerge. There are then several ways to evaluate how tighter regulation that prevents entry of large malls in these markets affects market outcomes.

We follow Ciliberto and Tamer (2009), and focus on the changes in average upper bounds on the probability of market outcomes, such as observing entry of a certain player, or observing entry by at least two players. More formally, consider an outcome y in the sigma algebra of the set $Y = \{0,1\}^3$. For each market with covariates x, and a fixed parameter value $\theta \in \hat{\Theta}_I$ we can find the upper bound on the probability of outcome y as:

$$\overline{p}_{x,\theta}(y) = \max_{\nu \in E^{BCE}(x,\theta)} \int \nu(y,\varepsilon)\, d\varepsilon,$$

so that, when averaging across markets $x \in \tilde{X}$, we have $\overline{p}_{\theta}(y) = \frac{1}{|\tilde{X}|}\sum_x \overline{p}_{x,\theta}(y)$. The same procedure yields, for all markets with counterfactual covariates x', upper bounds $\overline{p}_{x',\theta}(y)$ and average upper bounds $\overline{p}_{\theta}^{CF}(y)$, so that for every parameter value we have the difference in average upper bounds:

$$\Delta_{\overline{p}}^{\theta,y} = \left(\overline{p}_{\theta}^{CF}(y) - \overline{p}_{\theta}(y)\right).$$

We report in Table 9 the values of $\min_{\theta \in \hat{\Theta}_I} \Delta_{\overline{p}}^{\theta,y}$ and $\max_{\theta \in \hat{\Theta}_I} \Delta_{\overline{p}}^{\theta,y}$ for several market out-

[35] A similar result is observed in Haile and Tamer (2003) and in Dickstein and Morales (2015).

Table 8: Confidence Sets

Parameter	Weak Assumptions on Info - BCE	Complete Information - Nash
	(1)	(2)
Constant	[-2.35, -0.23]	[-3.26, -1.16]
Market Size	[3.66, 7.02]	[2.65, 6.79]
Home Region:		
Cooperatives	[-0.52, 1.43]	[-0.66, 1.47]
Indep. Italian Supermarket Groups	[-0.41, 1.96]	[-0.59, 1.88]
French Supermarket Groups	[-1.89, 1.82]	[-1.19, 1.55]
Presence of Large Malls:		
Cooperatives	[-2.79, 0.90]	[-3.04, -0.14]
Indep. Italian Supermarket Groups	[-2.33, 2.41]	[-3.08, 0.26]
French Supermarket Groups	[-4.18, 0.29]	[-4.65, -0.35]
Competitive Effects:		
Cooperatives	[-6.04, -1.99]	[-3.11, -0.01]
Indep. Italian Supermarket Groups	[-5.78, -1.35]	[-3.23, -0.19]
French Supermarket Groups	[-8.29, -2.63]	[-3.52, -0.02]
ρ - **Correlation Of Unobservable Profitability**	[0.43, 0.90]	[0.38, 1.00]

This table reports estimates of the game theoretic model of entry played among the three main players in the Italian supermarket industry. Payoffs are described by equation (7.1), and payoff types ε_m are iid across markets, distributed according to the distribution a Logistic distribution with zero mean and correlation ρ. Estimates are reported for each parameter value as projections of $\hat{\Theta}_{I,.95}$, the .95 confidence set for the identified set. The three numbers we report for each coefficient k correspond to $\left[\inf_{\theta\in\hat{\Theta}_{I,.95}}\theta^k, \sup_{\theta\in\hat{\Theta}_{I,.95}}\theta^k\right]$. The numbers we report in column (1) are obtained by estimating the model with weak assumptions on information, whereas numbers in column (2) are obtained from the model with complete information. See Appendix A for computational details.

comes y.

The two models are consistent in predicting a decrease in the upper bound of the average probability that small markets remain without supermarkets for most of the parameters in the identified set. Predictions on the change in probability of entry for different groups are different, in particular the BCE model allows for a sharp drop (more than -50%) of the upper bound of the probability that Cooperatives operate in a market after large malls are removed. Instead, the complete information model allows for a sharp increase (up to 70%) of the upper bound of the probability that French groups operate in a market.

Predictions on the change in probability of two key outcomes, entry by at least one or two players, are also affected by the assumptions maintained on information. In particular, the model that assumes perfect information predicts positive changes in the upper bound of the probability of observing at least one or at least two players in a market. This supports the view that preventing entry by large malls in small geographical grocery markets increases the likelihood of obtaining outcomes that are more conducive to consumer welfare. However, removing strong assumptions on information and considering predictions from the BCE model yields a fairly different picture. Under the assumption of BCE behavior, the change in the upper bound of the probability of having at least one or at least two supermarket industry players in a market does not have an unambiguously positive sign. Thus, the conclusion that removing large malls would increase the average probability that underserved markets end up with at least one or two supermarkets seems to rest on very restrictive assumptions on information, and does not stand once these assumptions are removed.

8 Conclusion

In this paper we present a method to estimate empirical discrete games, focusing on entry examples, under weak assumptions on the structure of the information available to players about each other's payoffs. Assumptions on information matter, since the different equilibrium predictions implied by different information structures translate in parameter estimates that might be biased if the information structure is misspecified. We are able to avoid strong parametric assumptions on information by adopting a broad equilibrium concept, Bayes Correlated Equilibrium (BCE), defined by Bergemann and Morris (2013, 2015). We argue that BCE is weak enough to make our method robust to assumptions on information, but informative enough to yield useful confidence sets for parameters. In an application, in which we study the effect of large malls on competition among supermarket groups in local grocery markets, we show that strong assumptions on information can bias counterfactual policy evaluations, while our method allows the analyst to avoid

Table 9: Counterfactual Change in Probability of Outcomes

Outcome	Weak Assumptions on Info - BCE	Complete Information - Nash
	(1)	(2)
No Entry	[-0.44, 0.06]	[-0.42, 0.06]
Entry by Cooperatives	[-0.54, 0.40]	[-0.03, 0.53]
Entry by Italian Groups	[-0.21, 0.11]	[-0.14, 0.45]
Entry by French Groups	[-0.08, 0.17]	[-0.32, 0.69]
Entry by at least 1 Player	[-0.03, 0.32]	[0.05, 0.43]
Entry by at least 2 Players	[-0.50, 0.33]	[0.16, 0.44]

This table reports the counterfactual change in average upper bounds in probabilities $\Delta_{\overline{p}}^{\theta,y} = \left(\overline{p}_{\theta}^{CF}(y) - \overline{p}_{\theta}(y)\right)$, where $\overline{p}_{\theta}^{CF}(y)$ is average upper bound of probability of outcome y in the counterfactual scenario in which large malls are removed, and $\overline{p}_{\theta}(y)$ is average upper bound of probability of outcome y in simulated with actual values of market level covariates. The average is taken across eight small geographical markets that currently have a large mall and no supermarkets. In brackets we report $\left[\min_{\theta\in\hat{\Theta}_I} \Delta_{\overline{p}}^{\theta,y}, \max_{\theta\in\hat{\Theta}_I} \Delta_{\overline{p}}^{\theta,y}\right]$. The numbers we report in column (1) are obtained with the estimates from the model with weak assumptions on information, whereas numbers in column (2) are obtained from the model with complete information.

restrictive assumptions.

We use throughout this paper an entry application, but also recognize that entry is a dynamic phenomenon. With its weaker equilibrium assumption, our method can be seen as a good candidate to capture with a static model behavior that is generated by a more complicated underlying dynamic model. Moreover, the use of (pure) Nash Equilibrium in games of complete information in entry applications is often justified on the grounds that this information structure and equilibrium notion, which impose no regret on players, is apt to capture a rest point of dynamic interaction. In work in progress, we argue that the theory of learning in games provides a rigorous rationale for the use of BCE to capture the steady state of repeated play when agents adopting myopic adaptive behavior rules have to learn to play equilibrium strategies.

There are other avenues for future research left open by this paper. Our method for the estimation of games under weak assumptions on information could be applied beyond discrete games, starting with models of auctions. Using BCE to allow bidders to have information on each other's valuation, as in Bergemann, Brooks and Morris (2015), seems relevant for several applied contexts, but will require careful thought on identification, feasible computation, and inference. We also do not pursue in this paper identification of information structures. While trying to recover an information structure from data on binary outcomes might be too optimistic, richer data like those generated by play in games with continuous actions could allow to identify the information structure of the game that generates the observable outcomes.

References

AGCM, I. A. A. (2013): "Indagine Conoscitiva sul Settore della GDO," *General Fact Finding Investigation*.

Aguirregabiria, V. and A. Magesan (2015): "Identification and Estimation of Dynamic Games when Players' Beliefs are Not in Equilibrium," *Working Paper*.

Aguirregabiria, V. and J. Suzuki (2015): "Empirical Games of Market Entry and Spatial Competition in Retail Industries," .

Aliprantis, C. D. and K. C. Border (1994): *Infinite Dimensional Analysis*, Springer.

Andrews, D. W. K. and X. Shi (2013): "Inference Based on Conditional Moment Inequalities," *Econometrica*, 81, 609–666.

Aradillas-Lopez, A. (2010): "Semiparametric Estimation of a Simultaneous Game with Incomplete Information," *Journal of Econometrics*, 157, 409–431.

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

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010a): "Estimating Static Models of Strategic Interactions," *Journal of Business & Economic Statistics*, 28, 469–482.

Bajari, P., H. Hong, and D. Nekipelov (2010b): "Game Theory and Econometrics: A Survey of Some Recent Research," in *Advances in Economics and Econometrics: Tenth World Congress*, vol. 3, 3–52.

Bajari, P., H. Hong, and S. P. Ryan (2010c): "Identification and Estimation of a Discrete Game of Complete Information," *Econometrica*, 78, 1529–1568.

Bentivogli, C. and E. Viviano (2012): "Le Trasformazioni del Sistema Produttivo Italiano: Le Cooperative," *Occasional Papers, Bank of Italy*.

Beresteanu, A., I. Molchanov, and F. Molinari (2011): "Sharp Identification Regions in Models with Convex Moment Predictions," *Econometrica*, 79, 1785–1821.

Bergemann, D., B. Brooks, and S. Morris (2015): "First Price Auctions with General Information Structures: Implications for Bidding and Revenue," *Cowles Foundation Discussion Paper, Yale University*.

Bergemann, D. and S. Morris (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2015): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, forthcoming.

Berry, S. T. (1989): "Entry in the Airline Industry," *Unpublished Ph.D. Dissertation, Univ. of Wisconsin-Madison*.

——— (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 6, 889–917.

Berry, S. T. and E. Tamer (2006): "Identification in Models of Oligopoly Entry," *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, 2, 46–85.

Border, K. C. (1989): *Fixed point theorems with applications to economics and game theory*, Cambridge university press.

Borkovsky, R., P. Ellickson, B. Gordon, V. Aguirregabiria, P. Gardete, P. Grieco, T. Gureckis, T.-H. Ho, L. Mathevet, and A. Sweeting (2015): "Multiplicity of Equilibria and Information Structures in Empirical Games: Challenges and Prospects," *Marketing Letters*, 26, 115–125.

Bresnahan, T. F. and P. C. Reiss (1991a): "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57–81.

——— (1991b): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 977–1009.

Brock, W. A. and S. N. Durlauf (2001): “Discrete Choice with Social Interactions,” *The Review of Economic Studies*, 68, 235–260.

Chang, J. T. and D. Pollard (1997): “Conditioning as disintegration,” *Statistica Neerlandica*, 51, 287–317.

Chen, X., E. Tamer, and A. Torgovitsky (2011): “Sensitivity Analysis in Semiparametric Likelihood Models,” *Working Paper*.

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

Ciliberto, F. and E. Tamer (2009): “Market Structure and Multiple Equilibria in Airline Markets,” *Econometrica*, 77, 1791–1828.

Datta, S. and K. Sudhir (2013): “Does Reducing Spatial Differentiation Increase Product Differentiation? Effects of Zoning on Retail Entry and Format Variety,” *Quantitative Marketing and Economics*, 11, 83–116.

de Paula, Á. (2013): “Econometric Analysis of Games with Multiple Equilibria,” *Annual Review of Economics*, 5, 107–131.

de Paula, Á. and X. Tang (2012): “Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information,” *Econometrica*, 80, 143–172.

Ellickson, P. B. and P. L. Grieco (2013): “Wal-Mart and the geography of grocery retailing,” *Journal of Urban Economics*, 75, 1–14.

Ellickson, P. B., S. Houghton, and C. Timmins (2013): “Estimating Network Economies in Retail Chains: A Revealed Preference Approach,” *The RAND Journal of Economics*, 44, 169–193.

Galichon, A. and M. Henry (2011): “Set Identification in Models with Multiple Equilibria,” *The Review of Economic Studies*, 78, 1264–1298.

Grieco, P. L. E. (2014): “Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets,” *The RAND Journal of Economics*, 45, 303–340.

Haile, P. A. and E. Tamer (2003): “Inference with an Incomplete Model of English Auctions,” *Journal of Political Economy*, 111, 1–51.

Igami, M. (2011): “Does Big Drive Out Small?” *Review of Industrial Organization*, 38, 1–21.

Jia, P. (2008): “What Happens when Wal-mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry,” *Econometrica*, 76, 1263–1316.

Jovanovic, B. (1989): “Observable Implications of Models with Multiple Equilibria,” *Econometrica*, 1431–1437.

Kashaev, N. (2015): “Testing for Nash behavior in entry games with complete information,” *Working Paper*.

Kline, B. (2014): “The Empirical Content Of Games With Bounded Regressors,” *Quantitative Economics*, forthcoming.

——— (2015): “Identification of complete information games,” *Journal of Econometrics*, 189, 117–131.

Lewbel, A. and X. Tang (2015): “Identification and Estimation of Games with Incomplete Information Using Excluded Regressors,” *Journal of Econometrics*, 189, 229–244.

Magnolfi, L. and C. Roncoroni (2015): “Political Connections as Barriers to Entry,” *Working Paper*.

Manski, C. F. (1988): “Identification of Binary Response Models,” *Journal of the American statistical Association*, 83, 729–738.

——— (2003): *Partial Identification of Probability Distributions*, Springer.

——— (2009): *Identification for Prediction and Decision*, Harvard University Press.

Mazzeo, M. J. (2002): “Product Choice and Oligopoly Market Structure,” *The RAND Journal of Economics*, 33, 221–242.

Mazzeo, M. J., K. Seim, and M. Varela (2014): “The Welfare Consequences of Mergers with Endogenous Product Choice,” *Working Paper*.

Morales, E. and M. J. Dickstein (2015): “What do Exporters Know?” *Working Paper*.

Navarro, S. and Y. Takahashi (2012): “A Semiparametric Test of Agents’ Information Sets for Games of Incomplete Information,” *Working Paper*.

Nishida, M. (2014): “Estimating a Model of Strategic Network Choice: The Convenience-Store Industry in Okinawa,” *Marketing Science*, 34, 20–38.

Paarsch, H. J. (1992): “Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions,” *Journal of Econometrics*, 51, 191–215.

Schaumans, C. and F. Verboven (2008): “Entry and regulation: evidence from health care professions,” *The Rand journal of economics*, 39, 949–972.

Schivardi, F. and E. Viviano (2011): “Entry Barriers in Retail Trade,” *The Economic Journal*, 121, 145–170.

Seim, K. (2006): “An Empirical Model of Firm Entry with Endogenous Product-Type Choices,” *The RAND Journal of Economics*, 37, 619–640.

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

——— (2010): "Partial Identification in Econometrics," *Annual Review of Economics*, 2, 167–195.

Tang, X. (2010): "Estimating Simultaneous Games with Incomplete Information under Median Restrictions," *Economics Letters*, 108, 273–276.

Toivanen, O. and M. Waterson (2005): "Market Structure and Entry: Where's the Beef?" *RAND Journal of Economics*, 36, 680–699.

Viviano, E. (2008): "Entry Regulations and Labour Market Outcomes: Evidence from the Italian Retail Trade Sector," *Labour Economics*, 15, 1200–1222.

Viviano, E., L. Aimone Gigio, E. Ciapanna, D. Coin, F. Colonna, F. Lagna, and R. Santioni (2012): "La Grande Distribuzione Organizzata e l'Industria Alimentare in Italia," *Occasional Papers, Bank of Italy*.

Xu, H. (2014): "Estimation of Discrete Games with Correlated Types," *The Econometrics Journal*, 17, 241–270.

Yang, Z. (2009): "Correlated Equilibrium and the Estimation of Discrete Games of Complete Information," *Working Paper*.

Zheng, F. (2016): "Spatial Competition and Preemptive Entry in the Discount Retail Industry," *Working Paper*.

Appendix A - Computational Details

Computation of G_n

In order to make the inferential procedure outlined above feasible, we need to compute G_n,[36] which requires computing $\sup_{b\in B}\left[P'_{y|x}b-h\left(Q_\theta^{BCE}(x);b\right)\right]$. We can rewrite $m(\theta,x)$ as the objective of the following program:

$$
\begin{array}{lcl}
\max_{b}\min_{(p,\nu)} & b'\left(P_{y|x}-p\right) & (P0) \\
s.t. & \sum_{j=1}^{|Y|} b_j^2 \leq 1, & \\
 & 0 \leq \nu(y,\varepsilon), & \forall y\in Y, \varepsilon\in\mathcal{E} \\
 & p(y)=\sum_{\varepsilon\in\mathcal{E}}\nu(y,\varepsilon), & \forall y\in Y \\
 & \sum_{y\in Y}\nu(y,\varepsilon)=f(\varepsilon;\theta_\varepsilon) & \forall\varepsilon\in\mathcal{E} \\
 & \sum_{y\in Y,\varepsilon\in\mathcal{E}}\nu(y,\varepsilon)=1, & \\
 & \sum_{y_{-i}\in Y_{-i}},\sum_{\varepsilon_{-i}\in\mathcal{E}_{-i}}\nu(y,(\varepsilon_i,\varepsilon_{-i}))[\pi_i^{x_j,\theta}(y,(\varepsilon_i,\varepsilon_{-i}))- & \forall i,j,y_i,y_i',\varepsilon_i \\
 & \pi_i^{x_j,\theta}\left(\left(y_i',y_{-i}\right),(\varepsilon_i,\varepsilon_{-i})\right)\geq 0 &
\end{array}
$$

We compute the value of the program $(P0)$ by approximating the infinite dimensional object ν. We do this by discretizing the set $\mathcal{E}$, which we substitute with the discrete set $\mathcal{E}^r$ containing r elements, so that the object ν has dimension $|Y|\times r^{|N|}\equiv d_\nu$. $(P0)$ is in the class of bilinear optimization problems with convex constraints; we use duality to find an equivalent program that can be handled by standard solvers. Let first $\tilde{p}=P_{y|x}-p$, and $(\tilde{p},\mathrm{vec}(\nu))=(z_1,z_2)$; all vectors are row vectors. The program now reads:

$$
\begin{array}{lcl}
\max_{u}\min_{z_1,z_2\geq 0_{d_\nu}} & \left(u,0_{d_\nu}\right)(z_1,z_2)^T & (P1) \\
s.t. & \|u\|\leq 1, & \\
 & A_{eq}z^T=a^T, & \\
 & A_{ineq}z^T\leq 0'_{d_{ineq}}, &
\end{array}
$$

in which A_{eq}, A_{ineq} and a are matrices that stack, respectively, linear equality constraints, linear inequalities and constants, and d_{ineq} is the number of rows of A_{ineq}. By duality, we

[36] The computation of the population criterion function G, necessary for the tables presented in Sections 5 and 6, is analogous.

obtain the equivalent program:

$$\begin{aligned}\max_{u,\lambda_{eq},\lambda_{ineq}\geq 0_{d_\nu}} \quad & a\left(\lambda_{eq},\lambda_{ineq}\right)^T \qquad (P2)\\ s.t. \quad & \|u\| \leq 1,\\ & \left(A^T\right)^r_{1:|Y|}\left(\lambda_{eq},\lambda_{ineq}\right)^T = -u^T,\\ & \left(A^T\right)^r_{|Y|+1:d_A}\left(\lambda_{eq},\lambda_{ineq}\right)^T \geq 0_{d_z},\end{aligned}$$

where $A = \begin{bmatrix} A_{eq} \\ A_{ineq} \end{bmatrix}$, the row vectors λ_{eq} and λ_{ineq} are the Lagrange multipliers associated to the constraints of $(P1)$, and $\left(A^T\right)^r_{1:|Y|}$ and $\left(A^T\right)^r_{|Y|+1:d_A}$ denote the first $|Y|$ and the last rows of the matrix A^T. By strong duality, as well as existence of BCE, $(P2)$ has the same value than $(P1)$, and we compute it using the solver KNITRO.

Computation of $\mathcal{L}_n$

We compute the confidence set for the identified set $\mathcal{L}_n(\hat{c})$ following the procedure outlined in Ciliberto and Tamer (2009).

1. We construct deterministic parameter grids using Halton sets around the parameter values of Probit regressions, and compute G_n for these points. We select the ten values of θ that have the lowest $G_n(\theta)$, and use them as starting points of a Simulated Annealing routine, which runs for ten thousand iterations.

2. We pool all the parameters visited by Simulated Annealing, and consider the corresponding set $\tilde{\Theta}$ as an approximation of Θ. We define as $g_n = \min_{\theta'\in\tilde{\Theta}} G_n(\theta')$, and can then obtain for all $\theta \in \tilde{\Theta}$:

$$\tilde{G}_n(\theta) = G_n(\theta) - g_n.$$

3. We extract $T = 100$ subsamples of size $n_t = n/3$. For each subsample i, we the criterion function using the subsampled observations, so that:

$$G_n^i(\theta) = \frac{1}{n_t}\sum_{j=1}^{n_t}\sup_{b\in B}\left[\hat{P}^{i\prime}_{y|x_j} b - h\left(Q_\theta^{BCE}\left(x_j\right); b\right)\right],$$

and then we find $g_n^i = \min_{\theta\in\Theta} G_n^i(\theta)$.

4. We choose the cutoff value $\hat{c}_0 = g_n/4$, and define the set:

$$\hat{\Theta}_I(\hat{c}_0) = \left\{\theta \in \Theta | \, \tilde{G}_n(\theta) \leq \hat{c}_0\right\}.$$

5. We obtain then $\tilde{G}_n^i(\theta) = G_n^i(\theta) - g_n^i$ for all $\theta \in \tilde{\Theta}$, and compute $\hat{c}_1$ as the 95th percentile of the distribution across subsamples of the statistic:

$$\tilde{L}_n^i(\hat{c}_0) = \sup_{\theta \in \hat{\Theta}_I(\hat{c}_0)} \left[G_n^i(\theta) - g_n^i\right].$$

6. Iterating steps 4,5 we obtain a sequence $\hat{c}_\ell, \hat{\Theta}_I(\hat{c}_\ell)$, which converges when

$$\sup_{\theta \in \hat{\Theta}_I(\hat{c}_0)} \left[G_n(\theta) - g_n\right] \leq \hat{c}_\ell.$$

We denote such $\hat{c}_\ell$ as $\hat{c}$, and report as $\mathcal{L}_n(\hat{c})$ the set $\hat{\Theta}_I(\hat{c})$.

Appendix B - Proofs

To prove Proposition 1, we first restate formally Bergemann and Morris (2015) robust prediction property in our context as Lemma 1.

Lemma (**1**). *For all $\theta \in \Theta$ and $x \in X$,*

1. *if $p \in Q_\theta^{BCE}(x)$, then $p \in Q_{\theta,S}^{BNE}(x)$ for some $S \in \mathcal{S}_0$.*
2. *Conversely, for all $S \in \mathcal{S}_0$, $Q_{\theta,S}^{BNE}(x) \subseteq Q_\theta^{BCE}(x)$.*

Proof. Fix $\theta \in \Theta$ and $x \in X$ throughout.

1. Consider $p \in Q_\theta^{BCE}(x)$. By definition of $Q_\theta^{BCE}(x)$, there exists then $\nu \in E_{\theta,x}^{BCE}$ such that $p = p_\nu$. We construct an information structure $S = \left\{\left(T^x, \left\{\mathbb{P}_{\tau|\varepsilon,x} : \varepsilon \in \mathcal{E}\right\}\right)_{x \in X}\right\}$ by defining for all $x \in X$:

$$T^x = Y,$$

and a probability kernel $\left\{\mathbb{P}_{\tau|\varepsilon,x} : \varepsilon \in \mathcal{E},\ x \in X\right\}$[37] such that:

$$\int_E \mathbb{P}_{\tau|\varepsilon,x}\{y\}\, d\mathbb{P}_\varepsilon^{\theta_\varepsilon}\{\varepsilon\} = \nu_x\{y, E\},\ \forall E \in \mathcal{B}(\mathcal{E}) : \mathbb{P}_\varepsilon^{\theta_\varepsilon}\{E\} > 0,\ y \in Y,$$

where $\mathbb{P}_\varepsilon^{\theta_\varepsilon}$ is the probability measure corresponding to the distribution of payoff types $F(\cdot;\theta_\varepsilon)$. Then, the incentive compatibility condition of the BCE distribution implies that the strategy profile s with:

$$\begin{aligned} s_i(y_i|\varepsilon_i, \tau_i = y_i) &= 1, \\ s_i(y_i|\varepsilon_i, \tau_i \neq y_i) &= 0,\ \forall y_i \in Y_i, \end{aligned}$$

[37] For the existence of such a kernel, see Chang and Pollard (1997).

is a BNE of the game $\Gamma^x(\theta, S)$, and $p_s = p_\nu$. Then, $p \in Q^{BNE}_{\theta,S}(x)$.

2. Suppose instead that $p = \sum_{k=1}^{K} \alpha_k p_{s_k} \in Q^{BNE}_{\theta}(x)$ for $K < \infty$, $\alpha_k \in (0,1)$ and $s_k \in E^{BNE}_{\theta,S,x}$ for all $k = 1,...,K$. Then, for each s_k there exists $\nu^k \in E^{BCE}_{\theta,x}$ such that:

$$\nu^k(y,E) = \int_{\varepsilon \in E} \left[\int_{\tau \in T} \prod_{i \in N} \{s_i(y_i|\varepsilon_i, \tau_i)\} \, d\mathbb{P}_{\tau|\varepsilon,x}\{\tau\} \right] d\mathbb{P}^{\theta_\varepsilon}_{\varepsilon}\{\varepsilon\},$$

for all $y \in Y$ and $E \in \mathcal{B}(\mathcal{E})$. Hence, $\sum_k \alpha_k \nu^k = \nu \in E^{BCE}_{\theta,x}$, and the corresponding $p_\nu = p \in Q^{BCE}_{\theta}(x)$. □

Proposition (1). *Let Assumptions 1 and 2 hold. Then,*

1. *the identified set under BCE behavior contains the true parameter value, $\theta_0 \in \Theta^{BCE}_I$, and*
2. *$\Theta^{BCE}_I = \Theta^{BNE}_I(\mathcal{S}_0)$.*

Proof. By Assumption 2, $P_{y|x} \in Q^{BNE}_{\theta_0,S_0}(x)$, while by Lemma 1, we have $Q^{BNE}_{\theta_0,S_0}(x) \subseteq Q^{BCE}_{\theta_0}(x)$. Since this happens $P_x - a.s.$, it follows by the definition of Θ^{BCE}_I that $\theta_0 \in \Theta^{BCE}_I$. Moreover, if $\theta \in \Theta^{BNE}_I(\mathcal{S})$ for any $\mathcal{S} \subseteq \mathcal{S}_0$, it follows that $\exists\ S \in \mathcal{S}$ such that $P_{y|x} \in Q^{BNE}_{\theta,S}(x)\, P_x-$ *a.s.* By Lemma 1 again, we have $Q^{BNE}_{\theta,S}(x) \subseteq Q^{BCE}_{\theta}(x)$, so that $\theta \in \Theta^{BCE}_I$. Hence, $\Theta^{BNE}_I(\mathcal{S}) \subseteq \Theta^{BCE}_I$. Consider instead $\theta \in \Theta^{BCE}_I$; by definition of Θ^{BCE}_I, there must be a collection of $(\nu^x)_{x \in X: P_x(X)=1}$ such that $p_{\nu^x} \in Q^{BCE}_{\theta}(x)$. It follows that, by Lemma 1, $p_{\nu^x} \in Q^{BNE}_{\theta,S}(x)\, P_x - a.e.$ for some $S \in \mathcal{S}_0$. Hence, $\Theta^{BCE}_I \subseteq \Theta^{BNE}_I(\mathcal{S}_0)$. □

Proposition (2). *Assume that:*

(i) the map $\theta_\pi \to \pi^{x,\theta_\pi}_i(y, \varepsilon_i)$ is continuous for all i, x, y and ε_i, the quantity

$$|\pi^{x,\theta_\pi}_i(y_i, y_{-i}, \varepsilon_i) - \pi^{x,\theta_\pi}_i\left(y'_i, y_{-i}, \varepsilon_i\right)|$$

is bounded above, and the map $\theta_\varepsilon \to F(\cdot;\theta_\varepsilon)$ is continuous for all ε;

(ii) the following uniform convergence condition holds: $\sup_\Theta \sqrt{n}|G_n(\theta) - G(\theta)| = O_p(1)$;

(iii) the sample criterion function G_n is stochastically bounded over Θ_I at rate $1/n$.

Then, the set $\hat{\Theta}_I = \{\theta \in \Theta |\, nG_n(\theta) \leq \log n\}$ is a consistent estimator of Θ^{BCE}_I.

Proof. We start by proving that the correspondence $Q : \theta \rightrightarrows Q^{BCE}_{\theta}(x)$ is upper hemi-continuous for all $x \in X$. This correspondence is a compound correspondence between the BCE equilibrium correspondence $\theta \rightrightarrows E^{BCE}_{\theta}(x)$ and the marginal operator $\nu \to \int_{\mathcal{E}} \nu(y,\varepsilon)\, d\varepsilon$. The compound correspondence of two upper hemi-continuous correspondences is hemi-continuous (Proposition 11.23 in Border 1989). Since $\nu \to \int_{\mathcal{E}} \nu(y,\varepsilon)\, d\varepsilon$ is a continuous function mapping into a compact set, establishing upper hemi-continuity of Q^{BCE}_{θ} can be accomplished by establishing the same property for the equilibrium correspondence $\theta \rightrightarrows E^{BCE}_{\theta}(x)$.

This requires establishing that the graph $\vartheta = \left\{(\theta, \nu) : \theta \in \Theta,\ \nu \in E_{\theta}^{BCE}(x)\right\}$ is closed. Consider a sequence $\theta^k \to \overline{\theta} \in \Theta$, for $\left\{\theta^k\right\}_{k=1}^{\infty} \in \Theta$, and a corresponding sequence $\left\{\nu_k^{x,\theta_\pi^k}\right\}_{k=1}^{\infty}$ such that $\nu_k^{x,\theta_\pi^k} \in E_{\theta^k}^{BCE}(x)$ for all k, and ν_k^{x,θ_π^k} converges setwise to the distribution $\overline{\nu}$. This happens if $\overline{\nu}$ is a BCE distribution in $E_{\overline{\theta},x}^{BCE}$. Consistency of $\overline{\nu}$ with respect to the prior follows for the continuity of the map $\theta_\varepsilon \to F(\cdot;\theta_\varepsilon)$. Incentive compatibility of $\overline{\nu}$ results from continuity of $\theta_\pi \to \pi_i^{x,\theta_\pi}$, and dominated convergence. Hence, $\overline{\nu}_x \in E_{\overline{\theta}}^{BCE}(x)$, so that the graph ϑ is closed. This establishes that the correspondence Q_θ^{BCE} is closed.

Then, the map

$$h : \theta \to h\left(Q_\theta^{BCE}(x), b\right) = \sup_{p \in Q_\theta^{BCE}(x)} p'b$$

is upper semicontinuous (Proposition 14.30 in Aliprantis and Border 1994), for all values of x, b. By definition of upper semi continuity, the set $\left\{\theta \in \Theta : h\left(Q_\theta^{BCE}(x), b\right) \geq c\right\}$ is closed for all $c \in \mathbb{R}$, which implies that $\left\{\theta \in \Theta : -h\left(Q_\theta^{BCE}(x), b\right) \leq -c\right\}$ is also closed for all $-c \in \mathbb{R}$, so that the map $\theta \to -h\left(Q_\theta^{BCE}(x), b\right)$ is lower semicontinuous. Then the map $\theta \to \sup_{b \in B}\left(P'_{y|x}b - h\left(Q_{\theta_n}^{BCE}(x); b\right)\right)$, pointwise supremum of a family of lower semicontinuous functions, is lower semicontinuous as well (Proposition 2.41 in Aliprantis and Border 1994). Hence, the function $G(\theta)$ is lower semicontinuous, since for a sequence $\theta_n \to \theta$ in Θ, we have that:

$$\begin{aligned}
\liminf_{n\to\infty} G(\theta_n) &= \liminf_{n\to\infty} \int_X \sup_{b \in B}\left[P'_{y|x}b - h\left(Q_{\theta_n}^{BCE}(x); b\right)\right] dP_x \\
&\geq \int \liminf_{n\to\infty} \sup_{b \in B}\left[P'_{y|x}b - h\left(Q_{\theta_n}^{BCE}(x); b\right)\right] dP_x \\
&= \int \sup_{b \in B}\left[P'_{y|x}b - h\left(Q_{\theta}^{BCE}(x); b\right)\right] dP_x \\
&= G(\theta),
\end{aligned}$$

where the first inequality holds by Fatou's Lemma.

Then, our setup satisfies the condition C.1 in Chernozhukov, Hong and Tamer (2007), and the consistency of $\hat{\Theta}_I$ follows by their Theorem 3.1. □

Proposition (**3**). *Suppose the econometrician observes the distribution of the data* $\left\{P_{y|x} : x \in X\right\}$, *generated by BCE play of a game. Then, under Assumption 3,*

1. *payoff parameters* β^C, β^E *and* Δ *are point identified as in single-agent threshold crossing models; and*

2. *the structure implies bounds on the payoff type parameter* θ_ε.

Proof. 1. Consider first the identification of β^C, β_2^E. We want to show that, for appropriate values of x, we have:

$$P_{y_2=1|x} = \int_{\left\{\varepsilon_2:\varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E\right\}} dF_2(\varepsilon_2;\theta_\varepsilon)$$

for $F_i(\cdot;\theta_\varepsilon)$ marginal distribution over ε_i defined from $F(\cdot;\theta_\varepsilon)$. The model implies the following link between the observables and the structure, for all $x \in X$ and some BCE measure $\nu^x \in E_\theta^{BCE}(x)$:

$$P_{y_2=1|x} = \nu^x\{y_1 = 1, y_2 = 1\} + \nu^x\{y_1 = 0, y_2 = 1\}.$$

By definition of ν^x, whenever $\nu^x\{y_1 = 1|\varepsilon_1\} > 0$ we have:

$$\sum_{y_2 \in \{0,1\}} \int_{\mathcal{E}_2} \left[\pi_1^{x,\theta_\pi}((1,y_2),\varepsilon_1)\right] d\nu^x\{y_2,\varepsilon_2|1,\varepsilon_1\} \geq 0. \tag{8.1}$$

Let $\beta_{1k}^E > 0$ without loss of generality, and consider $x_{1k} \to -\infty$. Conditional on such values of x, $\pi_1^{x,\theta_\pi}(1,y_2,\varepsilon_1) < 0$ for all values of y_2 a.e.-ε_1. Hence, by (8.1),

$$\lim_{x_{1k}\to-\infty} \nu^x\{y_1 = 1|\varepsilon_1\} = 0,$$

a.e.-ε_1. This implies that:

$$\begin{aligned}
\lim_{x_{1k}\to-\infty} \nu^x\{y_1 = 1, y_2 = 1\} &= \lim_{x_{1k}\to-\infty} \int_{\mathcal{E}_1} \nu^x\{y_1 = 1, y_2 = 1|\varepsilon_1\} dF_1(\varepsilon_1;\theta_\varepsilon) \\
&\leq \lim_{x_{1k}\to-\infty} \int_{\mathcal{E}_1} \nu^x\{y_1 = 1|\varepsilon_1\} dF_1(\varepsilon_1;\theta_\varepsilon) \\
&= 0.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\lim_{x_{1k}\to-\infty} P_{y_2=1|x} &= \lim_{x_{1k}\to-\infty} \nu^x\left\{y_1 = 0, y_2 = 1, \left\{\varepsilon : \varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E\right\}\right\} + \\
&\quad \nu^x\left\{y_1 = 0, y_2 = 1, \left\{\varepsilon : \varepsilon_2 < -z'\beta^C - x_2'\beta_2^E\right\}\right\}
\end{aligned}$$

and for $\varepsilon_2 < -z'\beta^C - x_2'\beta_2^E$, at the limit:

$$\begin{aligned}
\lim_{x_{1k}\to-\infty} \quad & \int_{\mathcal{E}_1} \left(z'\beta^C + x_2'\beta_2^E + \Delta_1 + \varepsilon_2\right) d\nu^x \{y_1 = 1, \varepsilon_1 | y_2 = 1, \varepsilon_2\} + \\
& \int_{\mathcal{E}_1} \left(z'\beta^C + x_2'\beta_2^E + \varepsilon_2\right) d\nu^x \{y_1 = 0, \varepsilon_1 | y_2 = 1, \varepsilon_2\} \\
= \lim_{x_{1k}\to-\infty} \quad & \int_{\mathcal{E}_1} \left(z'\beta^C + x_2'\beta_2^E + \varepsilon_2\right) d\nu^x \{y_1 = 0, \varepsilon_1 | y_2 = 1, \varepsilon_2\} < 0,
\end{aligned}$$

so that (8.1) implies that $\nu^x\{y_2 = 1|\varepsilon_2\} = 0$ for such ε_2. Similarly, (8.1) implies that $\nu^x\{y_2 = 0|\varepsilon_2\} = 0$ for $\varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E$. Hence,

$$\begin{aligned}
\lim_{x_{1k}\to-\infty} P_{y_2=1|x} &= \lim_{x_{1k}\to-\infty} \nu^x \left\{y_1 = 0, y_2 = 1; \left\{\varepsilon : \varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E\right\}\right\} \\
&= \lim_{x_{1k}\to-\infty} \int_{\left\{\varepsilon:\varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E\right\}} \nu^x \{y_1 = 0, y_2 = 1|\varepsilon\} dF(\varepsilon;\theta_\varepsilon) \\
&= \int_{\left\{\varepsilon_2:\varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E\right\}} dF_2(\varepsilon_2;\theta_\varepsilon).
\end{aligned}$$

The last equation describes a single-agent threshold crossing model, and under Assumption 3 we have that $\left(\beta^C, \beta_2^E\right)$ and F_i are point-identified (Manski, 1988).
Player 1's parameter β_1 are identified by an identical argument. To prove identification of Δ parameters, consider instead $x_{1k} \to \infty$; repeating the same steps leads to:

$$\lim_{x_{1k}\to-\infty} P_{y_2=1|x} = \int_{\left\{\varepsilon_2:\varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E - \Delta_1\right\}} dF_2(\varepsilon_2;\theta_\varepsilon).$$

2. Once β, Δ are identified by variation in covariates, we can derive (non-sharp) bounds on the joint distribution of payoff types $F(\varepsilon;\theta_\varepsilon)$, giving intuition on how some model restrictions identify the parameters θ_ε. Let:

$$\begin{aligned}
\overline{\mathcal{E}}(x,\theta) :&= \left\{\varepsilon_1 \geq -z'\beta^C - x_1'\beta_1^E - \Delta_2, \varepsilon_2 \geq -z'\beta^C - x_2'\beta_2^E - \Delta_1\right\}, \\
\underline{\mathcal{E}}(x,\theta) :&= \left\{\varepsilon_1 < -z'\beta^C - x_1'\beta_1^E, \varepsilon_2 < -z'\beta^C - x_2'\beta_2^E\right\}, \\
\widetilde{\mathcal{E}}(x,\theta) :&= \mathcal{E}(x,\theta)/(\overline{\mathcal{E}}(x,\theta) \cup \underline{\mathcal{E}}(x,\theta)).
\end{aligned}$$

Since these three sets form a partition of $\mathcal{E}$, we have:

$$\begin{aligned}
P_{y_1=1,y_2=1|x} &= \nu^x \{y_1 = 1, y_2 = 1, \mathcal{E}\}, \\
&= \nu^x \left\{y_1 = 1, y_2 = 1, \overline{\mathcal{E}}(x,\theta)\right\} + \nu_x \left\{y_1 = 1, y_2 = 1, \tilde{\mathcal{E}}(x,\theta)\right\} + \\
&\quad \nu^x \{y_1 = 1, y_2 = 1, \underline{\mathcal{E}}(x,\theta)\};
\end{aligned}$$

For any $x \in X$, (IC) implies that for $\overline{\varepsilon}_i \in \overline{\mathcal{E}}(x,\theta)$ we have $\nu^x\{y_i = 1|\overline{\varepsilon}_i\} = 1$, and for $\underline{\varepsilon_i} \in \underline{\mathcal{E}}$ we have $\nu^x\left\{y_i = 0|\underline{\varepsilon_i}\right\} = 1$. Hence,

$$\begin{aligned} P_{y_1=1,y_2=1|x} &= \int_{\overline{\mathcal{E}}} dF + \int_{\varepsilon\in\tilde{\mathcal{E}}} \nu^x\{y_1 = 1, y_2 = 1|\varepsilon\}\, dF(\varepsilon;\theta_\varepsilon), \\ &\leq \int_{\left\{\overline{\mathcal{E}}\cup\tilde{\mathcal{E}}\right\}} dF(\varepsilon;\theta_\varepsilon); \end{aligned}$$

and

$$P_{y_1=1,y_2=1|x} \geq \int_{\overline{\mathcal{E}}} dF(\varepsilon;\theta_\varepsilon).$$

The other moments of the data imply analogous bounds. Variation in x shifts the regions $\overline{\mathcal{E}}$ and $\overline{\mathcal{E}} \cup \tilde{\mathcal{E}}$, and provides useful restrictions on θ_ε. □

Appendix C - Relation with Grieco (2014)

We show in this appendix that the model presented in Grieco (2014) fits within the class of models described in Section 3. Consider the following simplified version of Grieco's model for a game of two players $i = 1, 2$ with actions $y_i \in \{0, 1\}$. Payoffs are:

$$\pi_i = y_i\left[\theta y_{-i} + \eta_i^1 + \eta_i^2\right],$$

and payoff types η are distributed according to:

$$\begin{pmatrix} \eta_1^1 \\ \eta_2^1 \\ \eta_1^2 \\ \eta_2^2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2\rho & 0 & 0 \\ \sigma^2\rho & \sigma^2 & 0 & 0 \\ 0 & 0 & 1-\sigma^2 & 0 \\ 0 & 0 & 0 & 1-\sigma^2 \end{pmatrix}\right). \tag{8.2}$$

The realizations of $\left(\eta_1^1, \eta_2^1\right)$ are publicly observable, so that player i observes $\left(\eta_1^1, \eta_2^1, \eta_i^2\right)$. Define now:

$$\varepsilon_i = \eta_i^1 + \eta_i^2,$$

and notice that player i's beliefs on ε_{-i} conditional on the observables be summarized by the conditional density:

$$\varepsilon_{-i}|\left(\eta_i^1, \eta_{-i}^1, \eta_i^2\right) \sim N\left(\eta_{-i}^1, 1-\sigma^2\right). \tag{8.3}$$

We want to recast this model so that it fits the framework of Section 3, in which player i observes its own scalar payoff type ε_i as well as a signal t_i on the opponents' payoff type. We interpret η_{-i}^1 as the signal that player i gets on ε_{-i}, and η_i^1 as what player i knows that $-i$ knows about her payoff, so that $\left(\tau_i^1, \tau_i^2\right) = \left(\eta_i^1, \eta_{-i}^1\right)$. It follows that $\left(\tau_i^1, \tau_i^2\right) = \left(\tau_{-i}^2, \tau_{-i}^1\right)$, so signals are public. The joint distribution of signals and redefined payoff shocks, derived from (8.2) is thus:

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \tau_1^1 \\ \tau_1^2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \\ \sigma^2\rho & 1 & \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 & \sigma^2 & \sigma^2\rho \\ \sigma^2 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \end{pmatrix} \right), \tag{8.4}$$

so that the distribution of $\tau_1|\varepsilon$ is:

$$\mathbb{P}_{\tau_1|\varepsilon} = N\left(\begin{pmatrix} \varepsilon_2 \\ \varepsilon_1 \end{pmatrix}, \begin{pmatrix} 1-\sigma^2 & 0 \\ 0 & 1-\sigma^2 \end{pmatrix} \right),$$

and the distribution of ε is:

$$\mathbb{P}_{\varepsilon} = N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma^2\rho \\ \sigma^2\rho & 1 \end{pmatrix} \right).$$

Notice that (8.4) implies that the belief of player i about ε_{-i} conditional on her information set is:

$$\varepsilon_{-i}|(\tau_i, \varepsilon_i) \sim N\left(\tau_i^1, 1-\sigma^2\right),$$

which is identical to the belief (8.3). This shows that the model in Grieco (2014) is in the class of models described in Section 3.

Appendix D - BMM Representation of the Identified Set

Beresteanu, Molinari and Molchanov (2011), henceforth BMM, provide a computable characterization of the identified set of partially identified models making use of random set theory. In this appendix, we show how our characterization of the identified set maps into their framework.

Let $z = (x, y)$ and ε be respectively the vector of observable outcomes and covariates, and the vector of payoff types. The random vectors are defined on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be the sigma algebra generated by the random vector x. We also adopt the assumptions 3.1(i),(iii) and 3.2 in BMM, and substitute 3.1(ii) with the assumption of BCE behavior. We restate these assumptions below for ease of reference:

Assumption 4. *(i) The discrete set of strategy profiles of the game, Y, is finite.*

(ii) Payoffs $\pi_i(y, \varepsilon_i; x, \theta_\pi)$ have a known parametric form, and are continuous in x and ε_j.

(iii) The observed outcome y of the game is the result of BCE behavior in the game of minimal information $\underline{S}$.

(iv) The conditional distribution of outcomes $P_{y|x}$ is identified by the data, and ε has a continuous distribution function.

For a given realization of x, there is a set $E_\theta^{BCE}(x)$ of BCE equilibrium distributions ν; considering $x(\omega)$ as a random vector, $E_\theta^{BCE}(x(\omega)) = E_\theta^{BCE}(\omega)$ is a random set. Let $\mathrm{Sel}(E_\theta^{BCE})$ denote the set of all $\nu(\omega)$ measurable selections of $E_\theta^{BCE}(\omega)$. In order to characterize the identified set, we need to map these equilibria into observable outcomes of the game for each $\omega \in \Omega$. A realization of ω implies both a realization of $(x(\omega), \varepsilon(\omega))$, and also a BCE distribution $\nu(\omega)$, which in turn determine the following probability distribution over outcomes:

$$q[\nu(\omega)] = \nu\{\cdot|\varepsilon(\omega)\} \in \Delta^{Y-1},$$

where $\nu\{\cdot|\varepsilon(\omega)\}$ is the conditional distribution implied by the joint distribution $\nu(\omega)$ over $(Y \times \mathcal{E})$, and the realization $\varepsilon(\omega)$. We denote as $\tilde{Q}_\theta$ as the set of all such equilibrium predictions:

$$\tilde{Q}_\theta = \left\{q(\nu):\ \nu \in \mathrm{Sel}\left(E_\theta^{BCE}\right)\right\}.$$

Then the conditional Aumann expectation of this random set is:

$$\mathbb{E}\left(\tilde{Q}_\theta|x\right) = \left\{E(q(\nu)|x):\ \nu \in \mathrm{Sel}\left(E_\theta^{BCE}\right)\right\}.$$

Notice however that:

$$\begin{aligned} E(q(\nu)|x) &= E[\nu\{\cdot|\varepsilon(\omega)\}|x] \\ &= \int_{\mathcal{E}} \nu\{y|\varepsilon\}\, dF(\varepsilon) \\ &= \int_{\mathcal{E}} \nu\{y, \varepsilon\}\, d\varepsilon, \end{aligned}$$

so that $\mathbb{E}\left(\tilde{Q}_\theta|x\right) = Q_\theta^{BCE}(x)$. Hence, our characterization of the identified set is equivalent to the one proposed in BMM.