Estimation of Discrete Games with Weak Assumptions on Information*

Lorenzo Magnolfi† and Camilla Roncoroni‡

First Draft: November, 2015; This Draft: July, 2017

Abstract

We propose a method to estimate static discrete games with weak assumptions on the information available to players. We do not fully specify the information structure of the game, but allow instead for all information structures consistent with players knowing their own payoffs and the distribution of opponents' payoffs. To make this approach tractable we adopt a weaker solution concept: Bayes Correlated Equilibrium (BCE), developed by Bergemann and Morris (2016). We characterize the sharp identified set under the assumption of BCE behavior and no assumptions on equilibrium selection, and find that in simple games with modest variation in observable covariates identified sets are narrow enough to be informative. In an application, we estimate a model of entry in the Italian supermarket industry and quantify the effect of large malls on local grocery stores. Parameter estimates and counterfactual predictions 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

JEL Codes: C57, L10

*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, Ken Hendricks, John Lazarev, Niccolò Lomys, Chiara Margaria, Francesca Molinari, Marcelo Sant'Anna, Che-Lin Su and Ed Vytlacil for helpful discussions. We also thank several seminar participants at John's Hopkins, Northwestern, NYU Stern, Penn State, UT-Austin, UW-Madison, Yale, Warwick and at the Eief-Igier-Bologna IO Workshop, ES NASM 2016, ES ESM 2016, EARIE 2016, XXXI Jornadas de Economía Industrial and AEA 2017 for comments. We gratefully acknowledge financial support from the Cowles Foundation. All errors and omissions are our own.

† *Corresponding Author*: Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison WI, USA. `magnolfi@wisc.edu`.

‡ Department of Economics, University of Warwick, Coventry CV4 7AL, UK. `c.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] Game-theoretic models' equilibrium predictions, and thus the map between the data and parameters of interest, depend crucially on the assumptions on the information that players have on each other's payoffs. However, the nature of firms' information about their competitors is often ambiguous in applications. Moreover, restrictive assumptions, when not satisfied in the application at hand, may 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 on 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, without assumptions on equilibrium selection. 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, adopted by the pioneering work in this area (Bjorn and Vuong, 1985; Jovanovic, 1989; Bresnahan and Reiss, 1991a; Berry, 1992); and private information (Seim, 2006; de Paula and Tang, 2012). Likewise, it nests the class of information structures considered by Grieco (2014). Moreover, our model is flexible in other dimensions: we allow the information structure of the game to vary across markets and to be asymmetric, i.e. agents may be informed about opponents' payoffs with varying levels of accuracy.

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, 2016), has the property of describing BNE predictions for a range 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. Exploiting the convexity of the set of equilibria, we also provide a tractable characterization of the sharp identified set of parameters without explicitly modeling equilibrium

[1] See for instance Bresnahan and Reiss (1991b), Berry (1992), 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.

selection. 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 mis-specification 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 a confidence set for parameters in the identified set using techniques developed in Chernozhukov, Hong and Tamer (2007).

We apply our method to the investigation of the effect of large malls on the grocery retail industry in Italy. The disagreement on the impact of the presence of these big outlets echoes the US debate on "Wal-Mart effects." Advocates of stricter regulation of large retailers claim that the superstores in malls drive out existing supermarkets and leave consumers without the option of shopping at local stores. Economic theory[2] and some of the existing evidence from other countries suggest however that local stores might benefit from the agglomeration economies created by the mall, or be differentiated enough not to suffer the competition of grocery-anchored shopping centers.

We estimate a static entry game using our robust method, and find mixed evidence on the effect of large malls on supermarkets. For all players in the industry the competition from a rival supermarket group seems to have a larger effect on profits than the competition from malls has. This is consistent with a substantial degree of differentiation between malls and local supermarkets, and thus a limited effect of malls on the availability of grocery stores. Our findings are in line with previous studies that have found a limited impact of supercenters on entry by small grocery retailers in the US (Ellickson and Grieco, 2013).

We compare these estimates with those obtained using a model of complete information. Results differ in important ways: in particular, we do not reject high values (in absolute value) of competitive effects, which are rejected under strong assumptions on information. This is because the assumption of complete information imposes that players fully anticipate competitors' decisions. As a consequence, the more restrictive complete information model may lead to underestimate how much players' profits are affected by the presence of competitors in a market.

In a counterfactual, we evaluate the effect on market structure of removing large malls from markets that currently have no other supermarket. Under weak assumptions on information, we find that the absence of the mall may or may not foster the emergence of a

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

market structure with at least two competing industry players. The model with complete information predicts instead that removing large malls results in a substantial increase in the average maximal probability[3] of observing at least two entrants. In this application, a model with restrictive assumptions on information leads us to strong conclusions, which are dispelled once more robust methods are adopted.

This article contributes to the literature on identification and estimation of static discrete games, recently surveyed by de Paula (2013). We follow Tamer (2003), Berry and Tamer (2006), who do not restrict equilibrium selection and allow for set identification of parameters. In particular, we rely on ideas in Beresteanu, Molchanov and Molinari (2011), who provide a useful characterization of the sharp identified set for models with convex predictions.[4]

Grieco (2014) is the first to estimate a game-theoretic model that relaxes the standard assumptions of either complete or perfectly private information. His model defines a parametric class of information structures where players receive both public and private signals; the relative precision of these signals is pinned down by the data. We adopt a complementary approach as we consider a model that is strictly more general, but we do not perform inference on the information structure.[5] Our emphasis on identification and estimation under weak assumptions on information is similar to the spirit of Dickstein and Morales (2016), who examine firms' export decisions, and develop a method to estimate payoff parameters without fully specifying firms' information on their expected revenues.

We build on the theoretical work of Bergemann and Morris (2013, 2016). They define the equilibrium concept used in this article and describe its property of offering robust predictions for games with incomplete information. Their characterization, developed in the context of theoretical work, inspires our use of a similarly robust framework in empirical applications. Aradillas-Lopez and Tamer (2008) study identification for a less restrictive solution concept, rationalizability.[6] Our approach is neither more general nor more restrictive than theirs, as they relax the assumption of equilibrium play, but work with restrictive assumptions on information.

Our study of the effect of the presence of large malls on local supermarkets is related to

[3]Since our model is partially identified and has multiple equilibria, it does not yield a unique counterfactual prediction. We follow Ciliberto and Tamer (2009) in reporting the average across markets and the maximum over equilibrium selections of the probability of observing a market structure outcome.

[4]Galichon and Henry (2011) provide an alternative characterization of the sharp identified set in game-theoretic models.

[5]The literature on discrete games faces a comparable trade-off between recovering a structural component and flexibility with respect to equilibrium selection. Whereas some studies parametrize and recover features of the equilibrium selection mechanism (Bajari, Hong and Ryan, 2010), other leave it unspecified (Tamer, 2003). Our approach with respect to the information structure is comparable to the latter studies.

[6]Yang (2009) examines estimation of discrete games of complete information under Nash behavior, using non-sharp restrictions imposed by Correlated Equilibrium in order to simplify computation. The assumption of Correlated Equilibrium under complete information is nested in our approach.

several papers that use structural models of market structure to examine the effect of entry of large store formats, especially Wal-Mart in the US, on other retailers, such as Jia (2008), Beresteanu, Ellickson and Misra (2010) and Arcidiacono et al. (2016). In a companion paper, Magnolfi and Roncoroni (2016), we study the role of political connections in shaping market structure in the Italian supermarket industry.

The structure of the article is as follows. In the following section, we define a general class of a discrete games. In Section 3 we discuss identification in this class of models, and motivate the use of BCE in empirical games. In Section 4 we compare our robust approach to models with more restrictive assumptions on information. In Section 5 we lay out sufficient conditions for identification in a more restrictive class of discrete games, and show evidence on the informativeness of our robust identified set. In Section 6 we develop the empirical application. Section 7 concludes.

2 Model

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

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

$$\left\{ P_{y|x} : x \in X \right\}$$

with $P_{y|x} \in \mathbb{P}_y$, the set of probability distributions over the finite set Y. The finite set of players N is also observable.

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.

2.1 Payoff Environment

Each player i is characterized by 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 vector

$\theta_\varepsilon \in \Theta_\varepsilon$.[7] Payoffs to player i, denoted by π_i, depend on action profiles and payoff types. Payoffs are also affected by observable covariates x, and finite dimensional payoff parameters $\theta_\pi \in \Theta_\pi$, so that for every player i and every pair (x, θ_π) there is a map:

$$\pi_i(\cdot; x, \theta_\pi) : Y \times \mathcal{E}_i \to \mathbb{R}.$$

We assume that ε is independent of the vector x. A realization of x and a vector of parameters $\theta = (\theta_\pi, \theta_\varepsilon) \in \Theta$ pins down a payoff structure. We want to identify, from data on behavior and market observable characteristics x, the vector of parameters θ. Although we present a model with ε independent of x and finite dimensional parameters θ, these restrictions are not necessary for the general discussion of robust identification in Section 3.[8]

We introduce here 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 $Y_i = \{0, 1\}$. The possible outcomes are either a duopoly when $(1, 1)$ is realized, or monopolies when either $(1, 0)$ or $(0, 1)$ are realized, or a market with no entrants with $(0, 0)$. Payoffs are:

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

so that the payoff parameter vector is $\theta_\pi = (\beta_i, \Delta_i)_{i=1,2}$. The parameter Δ_i, often called *competitive effect*, quantifies the effect of entry by firm i on firm $-i's$ payoffs. Payoff types ε_i are iid and uniformly distributed on $[-1, 1]$. Payoffs may be visualized in the following matrix:

	Player 2:	**Out**	**Enter**
Player 1:		0	1
Out	0	(0, 0)	$\left(0,\ x_2^T \beta_2 + \varepsilon_2\right)$
Enter	1	$\left(x_1^T \beta_1 + \varepsilon_1,\ 0\right)$	$\left(x_1^T \beta_1 + \Delta_2 + \varepsilon_1,\ x_2^T \beta_2 + \Delta_1 + \varepsilon_2 \right)$

[7] This specification allows for correlation among players' payoff types. See also Xu (2014) and Wan and Xu (2015) for models that allow for correlated payoff types.

[8] We do not pursue non-parametric identification of the payoff structure in this article, as it is not directly related to our main goal of achieving robustness with respect to assumptions on information. Similarly, we maintain for simplicity (but with some loss of generality) the assumptions of scalar, continuously distributed payoff types. We present this parametric setup to preserve the link with the previous literature and with applied work, although all results in Section 3 would go through without these assumptions. See Lewbel and Tang (2015) for an example of non-parametric 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.

2.2 Informational Environment

We assume that every player i knows the realization of her payoff type, as well as parameters and x.[9] 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 market x, the set of signals a player may receive and the probability of receiving them, given the realization of the vector of payoff types. Formally:

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

where T^x is subset of a separable metric space and represents the support of the vector of signals $\tau^x = (\tau_i^x)_{i\in N}$. The probability kernel $\left\{P^x_{\tau|\varepsilon} : \varepsilon \in \mathcal{E}\right\}$ 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, as we allow the informational environment to change with observable characteristics of the payoff environment.

Let $\mathcal{S}$ denote the collection of all possible information structures S. More formally, $\mathcal{S}$ is a general non-parametric class of information structures:

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

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 $\underline{S}$, in which payoff types are private information. 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: $P^x_{\tau_i|\varepsilon}([\tau_i = \varepsilon]) = 1$ for all $\varepsilon \in \mathcal{E},\ x \in X,\ i \in N$. Instead, in the minimal information structure $\underline{S}$ signals τ_i^x are uninformative: $P^x_{\tau_i|\varepsilon} = P^x_{\tau_i}$ for all $\varepsilon \in \mathcal{E},\ x \in X,\ i \in N$.

Our framework naturally accommodates the class of flexible information structures proposed in Grieco (2014).

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

$$\pi_i(y, \varepsilon; \theta_\pi) = y_i \left(\Delta y_{-i} + \varepsilon_i\right),$$

[9]Although this assumption is crucial for our point identification result in Section 5, it is not needed for our Proposition 1, and we could in principle allow for a different *baseline* level of information. We choose to assume that players know everything that is known to the econometrician, as this assumption is plausible in the economic environments we consider and provides identification power in applications. Aradillas-Lopez (2010) describes semi-parametric inference procedures for models in which the part of players' payoffs that is unobserved to the econometrician is private information, and players may be imperfectly informed about the part of opponents' payoffs that is observable to the econometrician.

and payoff types ε_i may be decomposed in two parts, $\varepsilon_i = \eta_i^1 + \eta_i^2$, and $\tau_i = (\eta_{-i}^1, \eta_i^1)$. The vector (η_{-i}^1, η_i^1) represents a publicly observed component of the payoff type that is correlated across players, whereas η_i^2 is a privately known component of the payoff type, independent across players. See Appendix E in Supplementary Materials online for more discussion of this example.

The model does not constrain the information structure to be the same across different markets, and encompasses other relevant cases, such as *privileged information* S^P, in which only some players know the type of the other players.[10] In this case, the signal spaces for all players are $T_i^x = \mathcal{E}$. For an informed player i, signals τ_i^x are distributed according to $P_{\tau_i|\varepsilon}^x([\tau_i = \varepsilon]) = 1$ for all $\varepsilon \in \mathcal{E}$, $x \in X$, whereas for an uninformed player j signals τ_j^x are distributed according to $P_{\tau_j|\varepsilon}^x = P_{\tau_j}^x$.

2.3 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 a 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. Let σ denote the equilibrium strategy profile, where:

$$\sigma = \times_{i \in N} \sigma_i, \ \ \sigma_i \in (\mathbb{P}_{Y_i})^{\mathcal{E}_i \times T_i^x}.$$

Let $BNE^x(\theta, S)$ denote the set of all BNE strategy profiles for the game $\Gamma^x(\theta, S)$; as $\Gamma^x(\theta, S)$ may have multiple equilibria, the set of equilibria $BNE^x(\theta, S)$ may not be a singleton.[11]

The informational environment of the game has important implications for equilibrium behavior. When players receive informative signals on their opponents' payoff types, their beliefs and hence their equilibrium behavior 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.[12]

Example. (1, Continued) Figure 1 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 distinct

[10]This information structure resembles the one in the proprietary information model of common value auctions (Engelbrecht-Wiggans, Milgrom and Weber, 1983).

[11]The set $BNE^x(\theta, S)$ may be empty; we assume that for the true information structure $BNE^x(\theta, S) \neq \emptyset$.

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

informational environments result in radically different equilibrium behavior. In Panel (A) we represent equilibrium outcomes for a game of complete information such as the one analyzed by Bresnahan and Reiss (1991a) and Tamer (2003). For every realization of ε, common knowledge for players, there is one or more equilibrium outcome. In Panel (B), equilibrium behavior takes the form of threshold strategies: each player does $y_i = 1$ iff $\varepsilon_i \geq 1/5$. In Panel (C) the privileged information structures results in equilibria where player 1 knows ε and can condition her behavior on the realizations of both ε_1 and ε_2. Player 2 only knows ε_2 and follows a threshold strategy. There is a continuum of such equilibria with thresholds $\varepsilon_2^* \in [1/8, 1/4]$.

[Figure 1 about here.]

For each equilibrium strategy $\sigma \in BNE^x(\theta, S)$ we can formulate the following prediction on behavior:

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

$$q_\sigma(y) = \int_{\mathcal{E}} \int_T \left(\prod_{i \in N} \sigma_i(\varepsilon_i, \tau_i)(y_i) \right) \mathrm{d}P_{\tau|\varepsilon} \mathrm{d}F,$$

for all $y \in Y$.

The set $BNE^x(\theta, S)$ of equilibria might not be a singleton, and we do not make any specific assumption on equilibrium selection. For any pair (θ, S), we define a prediction correspondence $Q^{BNE}_{\theta,S} : X \rightrightarrows \mathbb{P}_Y$

$$Q^{BNE}_{\theta,S}(x) = \mathrm{co}\left[\{ q \in \mathbb{P}_Y : \exists\, \sigma \in BNE^x(\theta, S) \text{ such that } q = q_\sigma \} \right],$$

where co[·] takes the convex hull of a set.[13] The prediction correspondence describes the set of distributions over actions y that may be obtained in a game $\Gamma^x(\theta, S)$ under the assumption of BNE play. The convex hull operator takes care of the multiplicity of equilibria: since we do not make assumptions on the equilibrium selection mechanism, we allow for all possible distributions over equilibria. In the next section, we consider identification in this model.

3 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 2. We are interested

[13]The set of BNEs of games in this class will not typically be convex. Convexification of the set of predictions captures the idea that distributions over outcomes may arise from arbitrary equilibrium selection mechanisms (taking the form of probability distributions over equilibria).

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 obtain for the parameters of interest if we allow for any $S \in \mathcal{S}$, and for any Bayes Nash Equilibrium constructed given any S.

Building on 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. (Data generating process) *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, for any restriction on information $\mathcal{S}' \subseteq \mathcal{S}$ let:

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

be the identified set of parameters consistent with BNE behavior. The sharp identified set of parameters without further assumptions on information is thus defined as:

$$\Theta_I^{BNE}(\mathcal{S}) = \left\{\theta \in \Theta | \exists\ S \in \mathcal{S} \text{ such that } P_{y|x} \in Q^{BNE}_{\theta,S}(x),\ P_x - a.s.\right\}. \tag{3.1}$$

This is the set of parameters whose implications, without restrictions on equilibrium selection or information structure, are compatible with the observables. All parameters $\theta \in \Theta_I^{BNE}(\mathcal{S})$ are *observationally equivalent*, as 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 $\Theta_I^{BNE}(\mathcal{S})$ is our object of interest. It captures all the restrictions on parameters that may be obtained under weak assumptions on the information structure.

Nevertheless, definition (3.1) seems hardly useful in practice, as computing correspondences $Q^{BNE}_{\theta,S}$ for all S in the large class $\mathcal{S}$ is an analytical challenge. In fact, a brute-force approach would require specifying all possible information structures, and finding the corresponding sets of equilibria. In the following subsections we propose a method to identify the set $\Theta_I^{BNE}(\mathcal{S})$ that sidesteps the difficulties inherent in a direct approach by relying on the connection between equilibrium behavior and robustness to assumptions on information.

3.1 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 $\Theta_I^{BNE}(\mathcal{S})$. We start with the definition of BCE, which follows Bergemann and Morris (2016).[14]

[14] Although not identical, this definition is equivalent to the one in Bergemann and Morris (2016), and is proposed in a previous working paper by the same authors. We adopt it as it's well suited for our purposes of identification and estimation.

Definition 2. (BCE) A Bayes Correlated Equilibrium $\nu \in \mathbb{P}_{Y,\mathcal{E}}$ for the game $\Gamma^x(\theta, \underline{S})$ is a probability measure ν over actions profiles and payoff types that is *consistent with the prior*: for all $\varepsilon \in \mathcal{E}$,

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

and *incentive compatible*: for all i, ε_i, y_i such that $\nu(y_i|\varepsilon_i) > 0$,

$$\sum_{y_{-i} \in Y_{-i}} \int_{\mathcal{E}_{-i}} \pi_i(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi) \nu(y_{-i}, \mathrm{d}\varepsilon_{-i}|y_i, \varepsilon_i) \geq$$

$$\geq \sum_{y_{-i} \in Y_{-i}} \int_{\mathcal{E}_{-i}} \pi_i(y_i', y_{-i}, \varepsilon_i; x, \theta_\pi) \nu(y_{-i}, \mathrm{d}\varepsilon_{-i}|y_i, \varepsilon_i), \forall y_i' \in Y_i.$$

The BCE concept is a natural generalization of Correlated Equilibrium to an incomplete information environment, under the assumptions that players have a common prior on the distribution of payoff types and on the signal structure. 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, hence its marginal over payoff types reflects the common knowledge of the underlying distribution of ε. Moreover, players need to be best-responding to equilibrium beliefs, as summarized by the BCE distribution. When best-responding, different actions played by the same type may be justified by different equilibrium beliefs, as in the standard notion of Correlated Equilibrium.

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. Although in principle BCE may be defined for any incomplete information game, we use Definition 2 in what follows, and use $BCE^x(\theta)$ to denote the set of BCE for the game $\Gamma^x(\theta, \underline{S})$, reflecting our choice of baseline information structure. The set $BCE^x(\theta)$ is convex, because the equalities and inequalities that define it 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 q_ν defined as:

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

The observable implications of BCE behavior in a structure characterized by $(\theta, \underline{S})$ are described by the prediction correspondence $Q_\theta^{BCE} : X \rightrightarrows \mathbb{P}_Y$, defined as:

$$Q_\theta^{BCE}(x) = \{q \in \mathbb{P}_Y : \exists\, \nu \in BCE^x(\theta) \text{ such that } q = q_\nu\}.$$

Because the set $BCE^x(\theta)$ is convex, any convex combination of BCE predictions is also a BCE prediction. Therefore, $Q_\theta^{BCE}(x)$ captures equilibrium predictions with no restrictions on equilibrium selection.

Figure 2 shows the set of BCE outcomes for Example 1, for the case with no covariates x. 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 about here.]

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:

$$\Theta_I^{BCE} = \left\{\theta \in \Theta \text{ such that } P_{y|x} \in Q_\theta^{BCE}(x)\ P_x - a.s.\right\}. \tag{3.2}$$

3.2 BCE Identification

Bergemann and Morris (2013) establish the robust prediction property of BCE. In our setup, this property translates into the equivalence, for any given θ, between the BCE predictions Q_θ^{BCE} and the union of BNE equilibrium predictions $Q_{\theta,S}^{BNE}$ taken over $S \in \mathcal{S}$. Figure 2 illustrates this result by representing the polytope Q_θ^{BCE} as well as the sets of BNE predictions $Q_{\theta,S}^{BNE}$ for the three information structures described in Example 2, and for the payoff structure described in Example 1.

Leveraging on the robust prediction property of BCE, we establish the following proposition:

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. $\Theta_I^{BCE} = \Theta_I^{BNE}(\mathcal{S})$.

Proof. See Appendix B. □

Proposition 1 is the foundation for the use of the BCE behavioral assumption for identification. The adoption of BCE allows to characterize the set of parameters consistent with equilibrium behavior and a common prior, with minimal assumptions on information. In light of Proposition 1, our approach is not aimed at changing the behavioral assumption that we impose on players, but rather at relaxing assumptions on information. The object $\Theta_I^{BNE}(\mathcal{S})$, impossible to characterize when relying on BNE behavior, is easily defined by relying on BCE behavior.

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 $\mathcal{S}' \subset \mathcal{S}$ selects a nonempty subset $\Theta_I^{BNE}(\mathcal{S}') \subset \Theta_I^{BCE}$. In fact, we show in Section 4 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.

3.3 Support Function Characterization of the Identified Set

We argued with Proposition 1 that Θ_I^{BCE} is the set of all parameters compatible with the observables and with the non-parametric class of information structures $\mathcal{S}$. To estimate and compute Θ_I^{BCE}, we need however a more practical characterization, as it's not immediately obvious how to compute the set as defined in equation (3.2).

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 as in Beresteanu, Molchanov and Molinari (2011).[15] Let B denote the unit ball in $\mathbb{R}^{|Y|}$ and let $h\left(\cdot; Q_\theta^{BCE}(x)\right) : B \to \mathbb{R}$ denote the *support function* of the set $Q_\theta^{BCE}(x)$:

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

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

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

We have then:

$$\begin{aligned} \Theta_I^{BCE} &= \left\{\theta \in \Theta |\, b^T P_{y|x} \leq h\left(b; Q_\theta^{BCE}(x)\right) \ \forall\, b \in B,\ P_x - a.s.\right\} \\ &= \left\{\theta \in \Theta |\ \max_{b \in B} \min_{q \in Q_\theta^{BCE}(x)} \left[b^T P_{y|x} - b^T q\right] = 0,\ P_x - a.s.\right\}. \end{aligned} \tag{3.3}$$

The computation of this object can be further simplified leveraging on the characterization of $BCE^x(\theta)$. Because the inner program is a linear constrained minimization problem, we

[15] Because BCE yields a set of predictions that is already convex, we do not need to use Aumann expectations as in Beresteanu, Molchanov and Molinari (2011). Appendix F in Supplementary Materials describes how our characterization of the identified set maps into their framework.

can consider its dual maximization program: this way it is possible to verify whether θ belongs to the identified set Θ_I^{BCE} by solving a single constrained maximization problem.[16] Appendix A provides more computational details.

3.4 Inference

Suppose that we observe an iid sample of players' choices and covariates $\{y_j, x_j\}_{j=1}^n$. To apply existing inferential methods, we assume that the set of covariates X is discrete.[17] We adopt an extremum estimation approach to perform inference. We redefine the identified set characterized in (3.3) as the set of minimizers of a non-negative criterion function G,[18] or

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

for

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

The sample analogue of the population criterion function is:

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

where $\hat{P}_{y|x_j}$ is the empirical frequency of strategy profile y in observations with covariates $x = \bar{x}$. 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. (Consistent estimator) *Assume that:*

---

[16] The computation of the the criterion function that we use for inference, described in the next subsection, is similarly simplified.

[17] Although 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 intensive for the estimation of our model. Other recent methods, such as Andrews and Soares (2010), Bugni (2010), Armstrong and Chan (2016), Kaido, Molinari and Stoye (2016) are instead designed for models that generate a finite number of (conditional) moment inequalities, and hence do not apply to our setup. For a recent overview of methods in this area, see Canay and Shaikh (2017).

[18] Since the set of predictions $Q_\theta^{BCE}(x)$ is a subset of the $(|Y|-1)$-dimensional simplex, in our application it is sufficient to adopt the equivalent criterion function:

$$G'(\theta) = \int_X \sup_{b'} \left[b'^T P'_{y|x} - h\left(b'; Q'^{BCE}_\theta(x)\right)\right] \mathrm{d}P_x,$$

where $b' \in \mathbb{R}^{|Y|-1}$, $P'_{y|x}$ is defined as the first $|Y|-1$ elements of $P_{y|x}$ and $Q'^{BCE}_\theta(x)$ is the set of the first $|Y|-1$ elements of BCE predictions.

1. *The map* $\theta_\pi \to \pi_i(y, \varepsilon_i; x, \theta_\pi)$ *is continuous for all* i, x, y *and* ε_i, *the quantity*

$$|\pi_i(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi) - \pi_i(y'_i, y_{-i}, \varepsilon_i; x, \theta_\pi)|$$

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

2. *The parameter space* Θ *is compact;*

3. *The following uniform convergence condition holds:* $\sup_{\theta\in\Theta} \sqrt{n}|G_n(\theta) - G(\theta)| = O_p(1)$*;*

4. *The sample criterion function* G_n *is stochastically bounded over* Θ_I *at rate* $1/n$*.*

Then, the set $\hat{\Theta}_I = \{\theta \in \Theta | \, nG_n(\theta) \leq a_n\}$ *is a consistent estimator of* Θ_I^{BCE} *for* $a_n \to \infty$ *and* $\frac{a_n}{n} \to \infty$*.*

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. As in Ciliberto and Tamer (2009) we perform inference by constructing confidence regions C_n for the identified parameters $\theta \in \Theta_I^{BCE}$. The regions C_n have the coverage property:

$$\liminf_{n\to\infty} P\{\theta \in C_n\} \geq 1 - \alpha, \, \forall\theta \in \Theta_I^{BCE}.$$

Appendix C in Supplementary Materials describes the details of how we obtain C_n.

We have thus far motivated the use of BCE to perform identification of payoff parameters under weak assumptions on information, and developed a tractable characterization of the identified set that easily extends to an inferential procedure. Several issues remain open. How does our approach compare with existing methods? What happens when restrictions on information imposed in estimation are not valid in the data generating process? We turn to these questions in the next section.

4 Assumptions on Information and Identification

The prevalent approach in the literature on estimation of games is to restrict the class of admissible information structures. In practice, this is done by choosing a class of information structures $\mathcal{S}' \subset \mathcal{S}$ such that the set of equilibrium predictions $Q_{\theta,S}^{BNE}$ is analytically tractable for $S \in \mathcal{S}'$, and by focusing the analysis on the set $\Theta_I^{BNE}(\mathcal{S}')$. For instance, seminal papers 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 Sweeting (2009), Bajari et al. (2010), and de Paula and Tang (2012) restrict $\mathcal{S}'$ to only contain the minimal information structure $\underline{S}$, whereby signals τ^x are uninformative.

Ideally, the restriction imposed on the information structure $\mathcal{S}'$ is *true*, that is $\mathcal{S}' = \{S_0\}$, or at least *well-specified* 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. In typical applications there is, however, little evidence on the nature of S_0. If instead $S_0 \notin \mathcal{S}'$, the model is *mis-specified*, and one of the following three scenarios will occur. Either (i) the mis-specification 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 mis-specification results in a nonempty identified set, selecting arbitrarily a region of $\Theta_I^{BNE}(\mathcal{S})$ that does not contain θ_0, or (iii) the model is *falsified* by the data, that is $\Theta_I^{BNE}(\mathcal{S}') = \emptyset$.

In the latter case, no parameter θ can rationalize the observables given the restriction on information $\mathcal{S}$'. 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 matching the data. If no such values exist for $\mathcal{S}' \subset \mathcal{S}$, then all the information structures $S \in \mathcal{S}'$ are falsified. Although we do not usually observe directly data on information, assumptions on information could be falsified because distinct (sets of) assumptions may have markedly different predictions. When estimation is performed under a mis-specified assumption, estimates may be inconsistent.[19] We show in the following subsection how assumptions on information may affect identification in the context of a simple example.

4.1 Impact of Strong Assumptions on Identification

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

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

with ε_i iid according to a uniform distribution on the interval [-1,1]. The parameter Δ belongs to the interval $\Theta = [-1, 0]$.

Restrictive assumptions on information have substantial impact on identification in this

[19] As pointed out by Ponomareva and Tamer (2011), estimating a misspecified model may result in tight bounds, which however may 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.

game. To see this more clearly, consider the non-sharp identified set:

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

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 only allow for pure-strategy equilibria we can immediately recover the parameter $\Delta \in \tilde{\Theta}_I^{BNE}$ by solving the equation:

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

If instead we adopt the restriction of fully private payoff types, that is $\mathcal{S}' = \{\underline{S}\}$, the symmetric BNE is characterized by the equilibrium probability of entry $\sigma_i([y_i = 1]) = \int_{\mathcal{E}_i} \sigma_i(\varepsilon_i)(1)\mathrm{d}F_i$ that solves the equation:

$$\sigma_i([y_i = 1]) = 1 - F_i(\overline{\varepsilon}),$$

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

$$\overline{\varepsilon} + \Delta\sigma_i([y_i = 1]) = 0,$$

so that $\sigma_i([y_i = 1]) = \frac{1}{2-\Delta}$. The corresponding implication 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 enters. There are in this case multiple equilibria: player 2 has a threshold strategy characterized by the value $\overline{\varepsilon}_2$, and player 1 always enters when $\varepsilon_1 > -\Delta$ and enters only if $\varepsilon_2 < \overline{\varepsilon}_2$ when $0 < \varepsilon_1 < -\Delta$. Optimality in the choice of $\overline{\varepsilon}_2$ implies that $\overline{\varepsilon}_2 \in \left[-\frac{\Delta(1+\Delta)}{2}, -\frac{\Delta}{2}\right]$, and the corresponding implication for the probability of duopoly is:

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

Suppose now that the true value of the parameter in the data generating process is $\Delta_0 = -0.5$. For a certain value of $P_y(1,1)$ observed in the data, restrictions on the information structure yield different identified sets. Table 1 summarizes the identified set $\tilde{\Theta}_I^{BNE}(\mathcal{S}')$ under several combinations of $\mathcal{S}'$ and S_0.

[Table 1 about here.]

From Table 1, it appears that overstating the amount of information available to players leads to an identified parameter that is lower, in absolute value, than the true parameter value.[20] 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 also enters. In particular, in the model with complete information players are certain that their competitor also enters when the equilibrium outcome is $(1, 1)$. Hence this model predicts, for a given parameter value, the lowest $P_y(1, 1)$ across all information structures. On the other hand, a model with some level of incomplete information generates a higher frequency of duopolies, as players are more likely to enter given a belief that does not assign probability one to the presence of a competitor. Attenuation bias is induced from mis-specified complete information models even if we use for identification moments other than $P_y(1, 1)$.[21]

In this example, we use just one moment from the distribution of the observables to get an intuition of the direction of the bias: the full set of moments would always falsify the alternative mis-specified models. Moreover, for models with richer action spaces and parametrization, it is harder to predict the direction of the bias resulting from mis-specification of the information structure and to link it to moments of the data in an intuitive fashion. Nevertheless, the example conveys the idea that mis-specification of the information structure may result in significant bias in the identified parameters: estimation of games under the assumption of BCE avoids this bias.

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. When relaxing identifying restrictions there is, in principle, a trade-off between robustness and informativeness of the identified sets. Nevertheless, when variation in covariates allows the econometrician to observe games in which strategic considerations are negligible, the assumption of BCE behavior is sufficient for point identification of several features of the model.

A Two-Parameter Example We consider first a two-player entry game with no covariates, and focus on how assumptions on equilibrium behavior and information affect the

[20]This type of attenuation bias has already been recognized in the literature by Bergemann and Morris (2013), and in the context of dynamic games by Aguirregabiria and Magesan (2016).

[21]In fact, complete information models maximize, for a given value of the parameter and across information structures, the probability of observing a monopoly. If data are generated by a model with some incomplete information, hence featuring a relatively lower probability of monopoly for given Δ_0, the identified value of Δ is attenuated also when using $P_y(0, 1)$ or $P_y(1, 0)$ for identification.

identification of competition effects. In Figure 3, we represent identified sets under different behavioral assumptions for data generated by Nash equilibrium play under complete information. Our method is a compromise between the goals of robustness and informativeness. In fact, the identified set Θ_I^{BCE} in red is much larger than the set obtained under the (correct) behavioral assumption of Nash Equilibrium with complete information (in yellow). The latter model imposes more restrictive assumptions on information, so that the corresponding identified set is small.

Figure 3 also shows the identified sets obtained under weaker assumptions on behavior. In blue, we show the identified set under level-1 rationality. This approach is strictly more general than ours, but has little identifying power.[22] We also represent in the figure the identified set under rationalizability (corresponding to level-2 rationality for this model). This behavioral assumption is defined for a complete information environment, so the identified set is not a superset of Θ_I^{BCE}, but relaxes the assumption of equilibrium play. Rationalizability yields no lower bounds for competition effects, hence the identified obtained under BCE is more informative.

[Figure 3 about here.]

Point Identification The previous example shows that the assumption of Bayes Correlated Equilibrium results in tighter identification than non-equilibrium behavioral restrictions do. At the same time, the figure also shows that Θ_I^{BCE} may be much larger than $\Theta_I^{BNE}\left(\overline{S}\right)$ when S_0, the information structure in the data generating process, coincides with $\overline{S}$ and thus there may be concerns on the informativeness of Θ_I^{BCE}. We argue that introducing a key source of identifying power, variation in exogenous covariates x, shrinks the identified set Θ_I^{BCE}. In particular, full-support variation of covariates that only enter one player's payoff yields point identification under the assumption of BCE, as it does for models relying on more restrictive informational assumptions. 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.[23] To formalize this intuition in a simple setting, we focus on the class of two-player binary games with linear payoffs and present the identification result in Proposition 3:

[22]For this model, identified sets for competition effects are unbounded under level-1 rationality even if we allow for the presence of observable covariates x in payoffs.

[23]Several other works in the literature establish point identification of players' utility functions under "at infinity" variation in game-theoretic models, for different sets of assumptions on information, equilibrium selection and parametric restrictions on primitives. See for instance Bajari, Hong and Ryan (2010), Grieco (2014) and Kline (2015). Notice that the identification strategy proposed in Kline (2016) , which does not rely on large support assumptions but requires the existence of "unique potential outcomes" for some realizations of the unobservables, in general does not apply to our model.

Assumption 3. (Two-player entry game with linear payoffs) *Let* $|N| = 2$ *and* $Y = \{0,1\}^2$; *let payoffs be:*

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

Assume moreover:

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

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

Proof. See Appendix B. □

The proposition relies on the occurrence of values of covariates for which one player has a dominant strategy for almost all payoff types: if this is the case, identification of payoffs proceeds as in single-agent binary choice models. The assumption of BCE behavior guarantees that players do not select dominated strategies with positive probability, and have equilibrium beliefs, and this is sufficient for point identification.[24] We also show that in our model it is possible to obtain bounds for the parameters that characterize the joint distribution of payoff types.

Because we are imposing weak restrictions on information and we are allowing for players to receive correlated signals, we may be concerned that the model has no identification power with respect to the correlation between payoff types. However, players know their payoff type, and cannot be induced to enter by signals that systematically mislead them about the probability of entry of competitors, as they have a common prior over types. This restriction implied by our equilibrium assumption helps to generate useful bounds on the correlation of payoff types.

[24] Kline (2015) establishes sufficiency of level-2 rationality for point identification of payoffs in complete information games. Though we argue that BCE is sufficient for point identification of payoffs in incomplete information games, weaker behavioral notions may suffice.

Identification with Finite Support Although we do not expect the large support assumptions of the proposition to always hold in applications, Proposition 3 indicates a source of variation that helps identification also in the case of covariates with finite support. To illustrate the the identifying power of BCE in the latter case, we compute identified sets for a simple two-player binary game with payoffs linear in covariates. We present in Table 2 projections of Θ_I^{BCE}. The set is computed for different data generating processes, characterized by information structures ($S_0 = \bar{S}$, $S_0 = \underline{S}$, and $S_0 = S^P$),[25] and for two examples of uniformly distributed covariates with finite support, X' and X''. Covariates $X' = X_1' \times X_2' \times X_C'$ are characterized by $X_i' = X_C' = \{-1, 0, 1\}$; covariates $X'' = X_1'' \times X_2'' \times X_C''$ are instead characterized by player-specific $X_i'' = \{-3, 0, 3\}$ for $i = 1, 2$ and $X_C'' = X_C'$. Results indicate that discrete sets of covariates have some identifying power in this model; the size of the identified set, as measured by projections for each parameter, shrinks considerably as we increase variation in covariates.

[Table 2 about here.]

6 Application: the Impact of Large Malls on Local Supermarkets

The emergence of large grocery-anchored malls in Italy, a relatively recent phenomenon, has sparked a debate on their impact on local retailers. If malls' "anchor" grocery stores represent a strong competitor to local supermarkets, as their critics argue,[26] the presence of shopping centers might generate a market structure with either few local supermarkets or monopolies. This may hurt consumers, who benefit from the availability of local stores. Others contend that format differentiation results in little competition between local supermarkets and anchors. Additionally, the economic activity linked to large malls may generate spillovers that strengthen local demand. Consequently, restrictive regulation on entry by malls would ultimately be harmful to consumers.

In this section we quantify the effects of the presence of malls on local supermarkets. To this aim, we estimate a game-theoretic model in which industry players decide strategically whether to operate stores in local grocery markets and the presence of large malls may affect supermarkets' expected profits.

[25] We also select an equilibrium for those DGPs characterized by games with multiple equilibria. In particular, for $S_0 = \overline{S}$ we select with equal probability one of the two pure-strategy equilibria, and for $S_0 = S^P$ we select the equilibrium that maximizes the probability of entry by player 2. Different equilibria in the DGP result in distinct identified sets, but don't change qualitatively the informativeness of our identified sets.

[26] 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 previous five years. See `http://www.confesercenti.it/blog/imprese-dei-centri-storici-sondaggio-confesercenti-swg-fisco-ha-inciso-negativamente-per-8-su-10/`.

The empirical methods developed in the previous sections of this paper are well suited to estimate our model of market structure in the Italian supermarket industry. The institutional features of the industry offer limited guidance on the information available to players, and firms condition their entry decisions on both private and public information. In particular, local authorities may impose costs on entrants that vary across stores and are mostly private information to firms.[27] Moreover, industry players are likely to be heterogeneous in their ability to collect and process private information.[28]

We model the cross-section of equilibrium market-structure outcomes as a simultaneous game, following a large literature (Bresnahan and Reiss, 1991b; Berry, 1992; Mazzeo, 2002; Seim, 2006; Ciliberto and Tamer, 2009). We do so for three reasons: first, we model an industry that went through a sudden expansion following a regulatory change, making the outcome of this expansion suitable for static equilibrium modeling. Second, even if it's possible to collect data on supermarkets opening date, it's much harder to obtain information on when exactly the decision of entering a market is taken and becomes common knowledge. In fact, industry sources mention heterogeneous and possibly long lags between the final decision to open a store and the store's opening date. Finally, although dynamic methods are appealing for applications where inter-temporal incentives are of first-order importance, most empirical models of dynamic games require strong assumptions on the nature of information and of unobserved heterogeneity that we want to avoid.

Previous studies of market structure in retail industries have explored aspects that are absent from our analysis, which provides instead greater flexibility with respect to the information structure. For instance, economies of density (Holmes, 2011) and chain-effects (Jia, 2008) have been found to be important in the US discount retail industry, but are unlikely to be as important in the Italian supermarket industry, which operates over a much smaller geographical area where no pair of geographical markets is more than a few hundred miles apart. [29]

We also estimate the game under the assumption of complete information, and discuss the consequences of imposing more restrictive assumptions.[30] The no-regret property of pure Nash equilibria in games of complete information is often viewed as a plausible feature of

[27]For example, firms may be required to build roads or parking lots when developing a new grocery store. These requirements are typically the result of private negotiations with local authorities.

[28]In Magnolfi and Roncoroni (2016) we explore in more depth one of the possible sources of this heterogeneity: firms' political connections.

[29]For other studies of entry in retail industries that model these aspects, see Ellickson, Houghton and Timmins (2013), Nishida (2014) and Zheng (2016) for multi-store firms and economies of density, Mazzeo (2002), Seim (2006) and Datta and Sudhir (2013) for endogenous product and location choice, and Aradillas-Lopez and Rosen (2016) for multi-store firms. See Aguirregabiria and Suzuki (2016) for a recent survey of structural models of competition in retail.

[30]Minimal information $\underline{S}$ would also represent a natural benchmark. However, sharp inference in models of incomplete information games with (i) no restrictions on equilibrium multiplicity and selection, and (ii) arbitrary correlation among payoff types has never been implemented in empirical applications to the best of our knowledge.

the long-run industry equilibrium captured by a static model (Ciliberto and Tamer 2009). This argument is not particularly strong for our setting. Our cross-section captures the industry at the end of a 15-years period of growth that followed an overhaul of regulation in 1998. Both accounting data and trade press sources indicate, however, that many stores are operating at a loss in 2013, so that regret for not having anticipated the level of competition cannot be ruled out.

Results from the application of the robust method are consistent with a substantial degree of differentiation between the grocery stores in malls and local supermarkets. In particular, we do not reject high values (in absolute value) of competitive effects, whereas low values (in absolute value) for the effect of malls on supermarkets are not rejected.

Adopting weak assumptions on information is key for this finding: the model with complete information generates confidence sets for parameters that are not nested into those produced by the more general model, with lower bounds for competitive effects that are closer to zero. This result echoes the intuition developed in Section 4 that overstating the amount of information available to players attenuates competitive effect parameters. As a consequence, in our counterfactual analysis we find that a market structure with at least two competing industry players may not be more likely in the absence of the mall. In contrast, the model with complete information predicts an increase of the probability of observing two or more local stores upon removing the mall from small geographical markets.

6.1 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 sourced from the market research firm IRI. We complement these with hand-collected information on malls and mall size, obtained from public online directories. We focus 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 an important role and relatively few instances of large malls. We obtain data on population and demographics from the 2011 official census, and data on (tax) income at the municipality level for 2013 from the Ministry of Economy and Finance.

Market Definition and Industry Players

To define the relevant markets for our study we need to specify both which store formats are direct competitors and the geographical extent of grocery markets. The Italian antitrust authority distinguishes between stores with floor space up to 1,500 m^2 (16,146 ft^2) and stores above this threshold, pointing out that these two categories differ fundamentally in location, product-line, and applicable regulation (see AGCM - Italian antitrust authority, 2013; Viviano et al., 2012). Larger stores have seen the fastest growth in this industry in the last 15 years, suggesting that firms and consumers prefer these modern formats. Since larger stores seem the most relevant to welfare outcomes and the most likely to compete with the grocery anchors in malls, we consider stores with a floor space of at least 1,500 m^2 (16,146 ft^2)[31] as the relevant market for our study.

No existing administrative unit provides a natural way of defining local grocery markets in Italy. Because commuting patterns capture consumers' daily movements better than administrative units do, we delimit markets starting from the geographical commuting areas defined by ISTAT, the national statistical agency, and split commuting areas that are too large.[32][33] The geographic extension of these markets is consistent with industry sources and previous studies.[34] We also drop from our sample large cities with more than three hundred thousand inhabitants in a municipality, as the density of highly urbanized areas makes it hard to separate distinct markets. This leaves us with 484 local grocery markets. We report summary statistics for these markets in Panel (A) of Table 3, 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.

[Table 3 about here.]

Firms operating in the Italian supermarket industry are heterogeneous. Coop Italia and Conad, networks of consumers' and retailers' cooperatives affiliated with the national umbrella organization Legacoop, have the largest market share. Despite their organizational form, they are managed efficiently and we assume that, in their entry behavior, they follow the same logic as their profit maximizing competitors. Several independent firms, all based in the North of the country, own and operate networks of large stores. Based on IRI data, five such firms (Esselunga, Bennet, PAM, Finiper and Selex) have a market share greater than 2.5% in 2013. Two large French retail multinationals, Auchan and Carrefour, have also entered the Italian market mostly in the early 2000s. Given the similarities among supermarket groups with comparable organizational structures, we conduct our analysis referring to the three types of market players mentioned above: cooperative groups, independent Italian supermarket groups, and French multinationals.

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

[32]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.

[33]Ellickson, Grieco and Khvastunov (2017) propose an alternative, data driven, approach to market definition in spatially differentiated retail industries that however requires store-level revenue data.

[34]Evidence collected by various European Antitrust Authorities indicates that most consumers travel little to do their grocery shopping. For example, UK's Competition Commission considers all large stores in a radius of 10-15 minutes by car to belong to the same market. 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. Pavan, Pozzi and Rovigatti (2017) use the same Italian commuting areas we use as a basis for market definition in their study of gasoline markets.

We define large malls as shopping centers including at least 50 independent shops, including a grocery anchor. Although these anchor supermarkets are not regarded by industry experts as very successful in their own right, they receive rent subsidies from mall operators, as they are believed to attract consumers that shop at other stores in the mall. Malls' "catchment area" is substantially larger than that of supermarkets, attracting shoppers who drive up to 30 minutes from a region that only partially coincides with the local grocery market. Most large malls are developed by local investors or specialized national firms.

The Italian supermarket industry is subject to extensive regulation, and entry in local markets may be delayed significantly by zoning and other laws.[35] We assume that all players that found profitable to enter a market were able, by year 2013, to do so. Regulation for large malls and zoning laws vary across regions; the large areas required for the development of malls are hard to find in densely populated areas, and lengthy negotiations with local authorities are often necessary.

To gain insight on the impact of large malls on grocery markets, we estimate descriptive linear regressions and ordered probit models.[36] 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 Panel (B) of Table 3, point to a small and negative covariation between market structure outcomes and the presence of large malls in a grocery market. These regressions however do not shed light on the potential differences in the impact of large malls on the behavior of different industry groups. In addition, the counterfactual market structure that would emerge if malls were not present in some geographical markets also depends on the competitive effect that supermarket industry groups have on each other's entry decisions.

6.2 Game-theoretic Model

We estimate 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 other players, and firm-market specific characteristics unobserved to the econometrician. Payoffs from entry for player i in market m are:

$$\pi_i(\cdot; x, \theta_\pi) = x_{im}^T \beta_i + \sum_{j \neq i} y_{j,m} \Delta_j + \varepsilon_{i,m},$$

[35]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.

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

whereas payoffs from staying out of the market are normalized to zero.[37]

Market level covariates x_{im} include a measure of market size, a dummy for the presence of large malls in the market, and a home-region dummy. The measure of market size is the product of population and logarithm of income in a market. We assume that the coefficient measuring the effect of market size on profits is the same across players. The effect of malls on supermarket players, the focus of our analysis, is heterogeneous across players. The home-region dummy has a player-specific coefficient and is excluded from the payoff of competitors. 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 Normal copula, with correlation between any pair $(\varepsilon_{i,m}, \varepsilon_{j,m})$ equal to ρ. We allow for player-specific competition effects: every player j, if on the market, reduces competitor's payoff by Δ_i. Data limitations impose a constraint on the number of parameters that we can precisely estimate; at the same time our model is flexible enough to preserve the dimensions of heterogeneity (in competitive effects and effects of malls) that are key to our application.

Although in principle supermarket groups may choose to enter a geographical market with several stores, or to vary store format, we assume that player's actions y_i are binary to reduce the complexity of the model. Moreover, we consider a game with three players, lumping together cooperatives, independent Italian groups and French groups.[38] In other words, player i (for example, independent Italian groups) can take action $y_{im} \in \{0, 1\}$ in market m (where, for example, $y_{im} = 1$ corresponds to entry by at least one Italian group with at least one supermarket in market m). These substantial simplifications respond to the need to limit the complexity of the model while maintaining the flexibility necessary to consider the counterfactuals that address our research question.

We assume that the presence of large malls is exogenous to outcomes in the supermarket industry. This is equivalent to maintain both econometric exogeneity and exogeneity from the point of view of the model. In fact, the presence of malls is also the result of an entry decision: we assume that this decision does not depend on the presence of local supermarkets,[39] while we allow it to depend on observed market characteristics x and on unobservable factors $\tilde{\varepsilon}$. The exogeneity assumption requires independence between ε and $\tilde{\varepsilon}$, conditional on x. This is a strong assumption, but there are reasons to believe that it

[37] This specification of entry profits may be interpreted as a "reduced form", justified on the grounds of parsimony and difficulties in modeling post-entry competition. A structural interpretation of this linear profit function is discussed in Berry (1989).

[38] As in Ciliberto and Tamer (2009), this assumption is appropriate as long as these players behave similarly in the markets in our sample. In the industry we examine the similarities among cooperatives, French and Italian groups in terms of size, ownership and organizational structure support the assumption of similar strategic behavior.

[39] Similar assumptions of exogenous entry by for the large player are maintained in Grieco (2014) and Ackerberg and Gowrisankaran (2006).

may hold in our data. First, malls have a larger catchment area than supermarkets, as they can attract consumers from a region that only partly overlaps with the local grocery market. Second, restrictive regulation and the limited availability of suitably large areas for development may push developers to locate malls far from their ideal location, in regions that are only viable because consumers travel relatively far for non-grocery shopping.[40]

We estimate the model under weak assumptions on the information structure, hence assuming that data are generated by BCE behavior. This approach not only nests all the information structures adopted thus far in the empirical games literature, but also allows for asymmetries in players' information that are relevant for this empirical setting and not compatible with existing models. To compare our method with standard techniques, we also obtain a confidence set for parameters under the assumption that data are generated by pure-strategy Nash equilibrium behavior in the game of complete information as in Ciliberto and Tamer (2009).

Proposition 3 guides our intuition on what variation in the observables identifies the parameters. Although our model includes a covariate that is firm specific, the home-region dummy, this variable does not have full-support, so our parameters are set identified. Bounds on the β parameters are identified by covariation of observable characteristics and entry patterns. Identification of Δ_j comes from the difference between the probability of entry for firms $-j$ in markets where x_j makes firm j unlikely to enter, and the corresponding probability in markets where x_j makes firm j very likely to entry. The model offers some identification power with respect to the parameter ρ, which captures correlation between unobservable payoff types. In particular, correlation between entry decisions across firms in markets that have different profitability across firms (based on data and other parameters) helps establish an upper bound on ρ. Similarly, correlation between entry decisions across firms in markets that have uniform profitability across firms helps establish a lower bound on ρ.

6.3 Results

The first column in Table 4 presents projections of the estimated 95% confidence set for parameters in 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.

[Table 4 about here.]

Results for the coefficient on market size indicate that the dimension of a local grocery market affects positively the profitability of entry. The effect of operating in a home-region

[40]In addition to the arguments that support the independence of unobservable payoff determinants, most large malls are already built or planned before the expansion of the supermarket industry considered in our model.

is not significantly different from zero for any of the groups we examine.

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 lie mostly on the negative real line. The game-theoretic model provides evidence that competitors' presence in a local market makes entry less profitable: the confidence set includes parameter vectors with large negative competitive effects. Projected confidence sets for the correlation parameter ρ are firmly positive, pointing to a substantial correlation among unobserved determinants of supermarkets' profits.[41]

In the second column of Table 4 we report the projections of the 95% confidence intervals for parameters in 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. Although 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. The competitive effects estimated under the assumption of complete information are mostly smaller, in absolute value, than those obtained with a model with weak assumptions on information. This finding is in line with our discussion in Section 4.1: by assuming complete information, we impose that those players who decide to operate in a market have correct expectations on competitors' presence. Instead, under BCE behavior, the equilibrium expectations allow for uncertainty about competitors' behavior. Hence, more negative values for the competitive effects parameters cannot be rejected. The interval for the correlation parameter ρ is smaller for the model with complete information on payoff shocks, and includes only very high values. To clarify this finding, consider that weaker assumptions on information offer ways of rationalizing correlation in players' actions that are alternative to correlation in payoff shocks, thus leading not to reject lower values of ρ. Moreover, high values of ρ help rationalize the high frequency of outcomes in the data where players play similar actions in a restrictive model that does not allow for incomplete information.

[41] The dimensions of the confidence set depends on several factors such as the model's specification and on the sample size, and it's not immediate to compare it with other findings in the literature. Grieco (2014) finds smaller intervals for most payoff parameters, which is not surprising as his model is more restrictive and his sample size is larger; the different parametrization and scale of coefficients, however, makes this comparison hard.

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, the identified set is empty and there is no reason to expect estimates obtained under that assumption to lie inside the robust estimated set.[42] [43]

Counterfactuals

We consider the counterfactual scenario in which regulation prevents the construction of large shopping malls in small markets. This counterfactual allows to quantify how market structure is affected by the presence of large malls. We examine in particular the eight small geographical grocery markets[44] in our dataset that have a large shopping center but no supermarkets in the current market configuration, and compute predicted outcomes of the entry game between supermarkets once the large shopping center is removed.

There are several ways to summarize the model's counterfactual predictions. In general, game-theoretic models do not yield deterministic predictions on counterfactual market structures, but rather predicted probability distributions over outcomes. Moreover, the multiplicity of parameter vectors in the confidence region, as well as the multiplicity of equilibria, implies that the model predicts multiple probability distributions on market structures. We follow Ciliberto and Tamer (2009), and focus on the changes in average upper bounds of the probability of market outcomes of interest, such as the entry of a specific player or entry of at least one or two players.[45]

More formally, consider an outcome of interest as a subset $\hat{Y}$ of admissible market structures, $\hat{Y} \subseteq Y$. For each market with covariates x, and a fixed parameter value in the confidence set $\theta \in C_n$ we can find the upper bound on the probability of outcome $\hat{Y}$ as:

$$\begin{aligned} q_{\hat{Y}}(\theta, x) &= \max_{\nu \in BCE^x(\theta)} \sum_{y \in \hat{Y}} \int \nu(y, \mathrm{d}\varepsilon) \\ &= h\left(b(\hat{Y}); Q_\theta^{BCE}(x)\right), \end{aligned}$$

[42] A similar result is observed in Haile and Tamer (2003) and in Dickstein and Morales (2016).

[43] This discussion suggests a possible procedure for rejecting assumptions on information, although the implementation is not straightforward in our inferential setup, and we do not pursue formal testing in this paper. For testing procedures in game-theoretic models, see also Takahashi and Navarro (2012), who develop testing procedures to distinguish between information structures, and Kashaev (2015) who proposes a test for Nash behavior in complete information games.

[44] For further details on these markets, see Appendix D in Supplementary Materials.

[45] An alternative exercise would use in the counterfactual the equilibrium distributions that best fit the data, as in Grieco (2014). Such a counterfactual yields sharper predictions but does not allow for the possibility that the counterfactual policy may affect also equilibrium selection.

where the second equality holds as $q_{\hat{Y}}(x;\theta)$ is equal to the support function of the set $Q_{\theta}^{BCE}(x)$ evaluated at an appropriate value $b(\hat{Y})$. We average across markets $x \in \hat{X}$ and obtain $q_{\hat{Y}}(\theta) = \frac{1}{|\hat{X}|}\sum_x q_{\hat{Y}}(\theta, x)$. An identical procedure yields, for the same markets but with counterfactual covariates x', upper bounds $q_{\hat{Y}}^{CF}(\theta, x')$ and average upper bounds $q_{\hat{Y}}^{CF}(\theta)$. For every parameter value we obtain the difference in average upper bounds:[46]

$$D_{\hat{Y}}(\theta) = \left(q_{\hat{Y}}^{CF}(\theta) - q_{\hat{Y}}(\theta)\right).$$

We report in Table 5 the values of $\min_{\theta \in C_n} D_{\hat{Y}}(\theta)$ and $\max_{\theta \in C_n} D_{\hat{Y}}(\theta)$ for several market outcomes $\hat{Y}$. We also present the same counterfactual object for the complete information model.[47]

[Table 5 about here.]

Since the confidence sets for our model do not determine the sign of the effect of malls, it's unsurprising that counterfactual predictions on the effect of removing malls are inconclusive. In contrast, the model with complete information predicts a decrease of the probability of no entry for most parameter values. Predictions on the change in probability of entry for distinct supermarket groups are also different: the BCE model allows for a smaller increase of the upper bound of the probability that each individual player operates in a market.

Predictions on the change in probability of entry by at least one or two players are prominently affected by the assumptions maintained on information. In particular, the model that assumes complete information predicts positive changes in the probability of observing 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 desirable for consumer welfare. However, removing strong assumptions on information and considering predictions from the BCE model yields a 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 players in a market does not have an unambiguously positive sign. Thus, the conclusion that removing large malls would favor entry in under-served markets seems to rest on restrictive assumptions on information, and does not stand once these assumptions are removed.

[46] The upper bound is also a natural object of interest for those outcomes that the counterfactual policy seeks to foster. In fact, the policy may include actions by the regulator that help firms to select equilibria that maximize the probability of such outcomes. Additional results for lower bounds on probabilities of outcomes, and for disaggregated markets, are available as robustness checks in Appendix D in Supplementary Materials.

[47] The corresponding object for the complete information model is obtained using an analogous procedure in which, however, upper bounds on probabilities are generated by Nash equilibrium behavior.

7 Conclusion

In this article, 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, because the equilibrium predictions implied by different information structures translate in parameter estimates that may be biased if the information structure is mis-specified. We are able to avoid strong assumptions on information by adopting a broad equilibrium concept, Bayes Correlated Equilibrium (BCE), defined by Bergemann and Morris (2013, 2016). 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 restrictive assumptions on information may drive counterfactual policy evaluations, whereas our method allows to avoid restrictive assumptions.

There are several avenues for future research left open by this article. 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 (2017), is potentially relevant for many applied contexts, but will require an appropriate approach to identification, feasible computation, and inference, given the different characteristics of auction models. We also do not pursue in this article inference on information structures. Although trying to recover an information structure from data on binary outcomes may be too optimistic, richer data like those generated by play in games with continuous actions may allow to identify the information structure of the game that generates the observable outcomes.

References

Ackerberg, D. A. and G. Gowrisankaran (2006): "Quantifying Equilibrium Network Externalities in the ACH Banking Industry," *The RAND Journal of Economics*, 37, 738–761.

AGCM (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 (2016): "Empirical Games of Market Entry and Spatial Competition in Retail Industries," *Handbook on the Economics of Retail and Distribution.*

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

Andrews, D. W. K. and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78, 119–157.

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

Aradillas-Lopez, A. and A. M. Rosen (2016): "Inference in Ordered Response Games with Complete Information," *Working Paper.*

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

Arcidiacono, P., P. Bayer, J. R. Blevins, and P. B. Ellickson (2016): "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition," *The Review of Economic Studies*, 83, 889–931.

Armstrong, T. B. and H. P. Chan (2016): "Multiscale Adaptive Inference on Conditional Moment Inequalities," *Journal of Econometrics*, 194, 24–43.

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

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

Beresteanu, A., P. Ellickson, and S. Misra (2010): "The dynamics of retail oligopoly," *Working Paper.*

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 (2017): "First-Price Auctions With General Information Structures: Implications for Bidding and Revenue," *Econometrica*, 85, 107–143.

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

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

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: 9th World Congress of the Econometric Society*, 2, 46–85.

Bjorn, P. A. and Q. H. Vuong (1984): "Simultaneous Equations Models for Dummy Endogenous Variables: a Game Theoretic Formulation with an Application to Labor Force Participation," *Working Paper.*

Boyd, S. and L. Vandenberghe (2004): *Convex Optimization*, Cambridge University Press.

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.

Bugni, F. A. (2010): "Bootstrap Inference in Partially Identified Models Defined by Moment Inequalities: Coverage of the Identified set," *Econometrica*, 78, 735–753.

Canay, I. A. and A. M. Shaikh (2017): "Practical and Theoretical Advances for Inference in Partially Identified Models," *Advances in Economics and Econometrics: 11th World Congress of the Econometric Society.*

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

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.

Dickstein, M. J. and E. Morales (2016): "What do Exporters Know?" *Working Paper.*

Ellickson, P. B., P. L. Grieco, and O. Khvastunov (2017): "Measuring Competition in Spatial Retail," .

Ellickson, P. B. and P. L. E. 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.

Engelbrecht-Wiggans, R., P. R. Milgrom, and R. J. Weber (1983): "Competitive Bidding and Proprietary Information," *Journal of Mathematical Economics*, 11, 161–169.

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.

Holmes, T. J. (2011): "The Diffusion of Wal-Mart and Economies of Density," *Econometrica*, 79, 253–302.

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.

Kaido, H., F. Molinari, and J. Stoye (2016): "Confidence Intervals for Projections of Partially Identified Parameters," *Working Paper.*

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

Kline, B. (2015): "Identification of Complete Information Games," *Journal of Econometrics*, 189, 117–131.

——— (2016): "The Empirical Content of Games with Bounded Regressors," *Quantitative Economics*, 7, 37–81.

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 (2016): "Political Connections and Market Structure," *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.

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.

Pavan, G., A. Pozzi, and G. Rovigatti (2017): "Strategic Entry and Potential Competition: Evidence from Compressed Gas Fuel Retail," *Working Paper*.

Ponomareva, M. and E. Tamer (2011): "Misspecification in Moment Inequality Models: back to Moment Equalities?" *The Econometrics Journal*, 14, 186–203.

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.

Sweeting, A. (2009): "The Strategic Timing Incentives of Commercial Radio Stations: An Empirical Analysis using Multiple Equilibria," *The RAND Journal of Economics*, 40, 710–742.

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

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

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*.

Wan, Y. and H. Xu (2014): "Semiparametric Identification of Binary Decision Games of Incomplete Information with Correlated Private Signals," *Journal of Econometrics*, 182, 235–246.

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*.

Zhu, T., V. Singh, and A. Dukes (2011): "Local Competition, Entry, and Agglomeration," *Quantitative Marketing and Economics*, 9, 129–154.

Appendix A - Computation of G and G_n

To find the identified set and conduct inference, we need to compute G and G_n. First, let us approximate the infinite dimensional object ν by discretizing the set $\mathcal{E}$. Let $\mathcal{E}^r \subset \mathcal{E}$ be the discretized set, with $|\mathcal{E}^r| = r$. For every market x, we need to solve the program:

$$\begin{array}{rcr}
\max\limits_{b} \min\limits_{q,\nu\geq 0} & b^T \left(P_{y|x} - q\right) & (P0) \\
s.t. & b^T b - 1 & \leq 0 \\
\forall y \in Y & q\left(y\right) - \sum_{\varepsilon} \nu\left(y,\varepsilon\right) & = 0 \\
\forall \varepsilon \in \mathcal{E} & \sum_{y} \nu\left(y,\varepsilon\right) - f\left(\varepsilon;\theta_{\varepsilon}\right) & = 0 \\
 & \sum_{y,\varepsilon} \nu\left(y,\varepsilon\right) - 1 & = 0 \\
\forall i, y_i, y_i', \varepsilon_i & \sum_{y_{-i}} \sum_{\varepsilon_{-i}} \nu\left(y,\varepsilon_i,\varepsilon_{-i}\right)\left(\pi_i\left(y_i', y_{-i}, \varepsilon_i, \varepsilon_{-i}; x, \theta\right) - \pi_i\left(y, \varepsilon_i, \varepsilon_{-i}; x, \theta\right)\right) & \leq 0
\end{array}$$

Given the discretization, ν has now dimension $|Y| \times r^{|N|} = d_\nu$. Also, $f\left(\cdot;\theta_\varepsilon\right)$ denotes the corresponding approximation of the prior distribution. We then transform $(P0)$ by defining new variables $\tilde{p} = P_{y|x} - q$, and $(\tilde{p}, \text{vec}\left(\nu\right)) = (z_1, z_2)$. . As the set of predictions is a subset of the $(|Y|-1)$-dimensional simplex, we also consider the object: $(\tilde{b}, 0)^T \left(P_{y|x} - q\right)$, where $\tilde{b}$ is equal to the first $|Y|-1$ elements of b. Al vectors are column vectors. The transformed program is:

$$\begin{array}{rcl}
\max\limits_{\tilde{b}} \min\limits_{z_1, z_2 \geq 0_{d_\nu}} & \left(\tilde{b}, 0, \mathbf{0}_{d_\nu}\right)^T \left(z_1, z_2\right) & (P1) \\
s.t. & \tilde{b}^T \tilde{b} & \leq 1 \\
 & A_{eq} z & = a \\
 & A_{ineq} z & \leq \mathbf{0}_{d_{ineq}},
\end{array}$$

where A_{eq}, A_{ineq} and a are matrices that stack, respectively, linear equality constraints, linear inequalities and constants, d_{ineq} is the number of rows of A_{ineq} and we use $\mathbf{0}_{d_\nu}$ to denote the $d_\nu-$vector of zeros. (P1) can be simplified taking the dual problem of the minimization problem. We obtain:

$$\begin{array}{rcl}
\max\limits_{\tilde{b}, \lambda_{eq}, \lambda_{ineq} \geq 0_{d_\nu}} & \left(a, \mathbf{0}_{d_{ineq}}\right)^T \left(\lambda_{eq}, \lambda_{ineq}\right) & (P2) \\
s.t. & \tilde{b}^T \tilde{b} & \leq 1 \\
 & \left(A^T\right)_{1:|Y|} \left(\lambda_{eq}, \lambda_{ineq}\right) & = (\tilde{b}, 0) \\
 & \left(A^T\right)_{|Y|+1:d_A} \left(\lambda_{eq}, \lambda_{ineq}\right) & \geq \mathbf{0}_{d_z},
\end{array}$$

where $A=\begin{bmatrix} A_{eq} \\ A_{ineq} \end{bmatrix}$, the row vectors λ_{eq} and λ_{ineq} are the dual variables associated to the constraints of $(P1)$, and $\left(A^T\right)_{1:|Y|}$ and $\left(A^T\right)_{|Y|+1:d_A}$ 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 $(P1)$ (Boyd and Vandenberghe, 2004) and we compute it using the solver KNITRO in the modeling environment AMPL. Computation of $G(\theta)$ for the two-player game of Table 2 with $r=50$ takes less than 30 seconds of CPU time on a 3.4Ghz quad-core Intel processor. Computation times for the function $G_n(\theta)$ in our application, with $r=10$, are similar. Parallel computation of $G(\theta)$ for different values of θ is not supported by AMPL, but can be implemented using the script Parampl.[48] Further details on how to compute Θ_I and $\mathcal{C}_n$ are in Appendix C in the Supplementary Materials online.

Appendix B - Proofs

Lemma 1 is a preliminary result needed to prove Proposition 1. In the lemma we restate and adapt to our context the robust prediction property of BCE, established as Theorem 1 in Bergemann and Morris (2016).

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

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

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

1. Consider $q \in Q_\theta^{BCE}(x)$. Then there exists $\nu \in BCE^x(\theta)$ such that $q = q_\nu$. We need to show that there exists an information structure S and a strategy profile σ such that $q_\sigma = q_\nu$ and $q_\sigma \in Q_{\theta,S}^{BNE}(x)$. To this aim, let $T^x = Y$ and define a probability kernel $\left\{P_{\tau|\varepsilon}^x : \varepsilon \in \mathcal{E}\right\}$[49] such that:

$$\int_E \left(P_{\tau|\varepsilon}[\tau = y]\right) \mathrm{d}F = \nu(y, E), \; \forall E \in \mathcal{B}(\mathcal{E}) : \int_E \mathrm{d}F > 0, \; y \in Y.$$

Also, for all ε_i, τ_i, let $\sigma_i(\varepsilon, \tau_i)(y_i) = 1$ if $y_i = \tau_i$, and $\sigma_i(\varepsilon_i, \tau_i)(y_i) = 0$ if $y_i \neq \tau_i$. Hence, the incentive compatibility conditions of BCE guarantee that σ is a BNE of the game $\Gamma^x(\theta, S)$.

2. Suppose that $q = \sum_{k=1}^K \alpha_k q_k \in Q_\theta^{BNE}(x)$ for $K < \infty$, $\sum_{k=1}^K \alpha_k = 1$ and $\sigma_k \in BNE^x(\theta, S)$ for all $k = 1, ..., K$. Then, for each σ_k we can obtain $\nu_k \in BCE^x(\theta)$ as:

$$\nu_k(y, E) = \int_E \int_T \left(\prod_{i \in N} \sigma_i(\varepsilon_i, \tau_i)(y_i)\right) \mathrm{d}P_{\tau|\varepsilon} \mathrm{d}F,$$

[48] Available at www.parampl.com. We thank Arthur Olszak for kind and patient support with Parampl.

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

for all $y \in Y$ and $E \in \mathcal{B}(\mathcal{E})$. Hence, $\sum_k \alpha_k \nu_k = \nu \in BCE^x(\theta)$, and the corresponding $q_\nu = q \in Q_\theta^{BCE}(x)$. $\square$

Proposition 1. *Let Assumptions 1 and 2 hold. Then:*

1. *The identified set under BCE behavior contains the true parameter value,* $\theta_0 \in \Theta_I^{BCE}$, *and*

2. $\Theta_I^{BCE} = \Theta_I^{BNE}(\mathcal{S})$.

Proof. 1. By Assumption 2, almost surely with respect to P_x, $P_{y|x} \in Q_{\theta_0,S_0}^{BNE}(x)$. Also, by Lemma 1, $Q_{\theta_0,S_0}^{BNE}(x) \subseteq Q_{\theta_0}^{BCE}(x)$. It follows, by the definition of Θ_I^{BCE}, that $\theta_0 \in \Theta_I^{BCE}$.

2. Let $\theta \in \Theta_I^{BNE}(\mathcal{S}')$ for some $\mathcal{S}' \subseteq \mathcal{S}$. Then $\exists\, S \in \mathcal{S}'$ such that $P_{y|x} \in Q_{\theta,S}^{BNE}(x)\, P_x - a.s.$ Since, by Lemma 1 again, we have $Q_{\theta,S}^{BNE}(x) \subseteq Q_\theta^{BCE}(x)$, $\theta \in \Theta_I^{BCE}$ and $\Theta_I^{BNE}(\mathcal{S}') \subseteq \Theta_I^{BCE}$. Consider instead $\theta \in \Theta_I^{BCE}$; by definition of Θ_I^{BCE}, there must be a collection of $(\nu^x)_{x\in X:}$ such that $p_{\nu^x} \in Q_\theta^{BCE}(x)$. It follows that, by Lemma 1, $p_{\nu^x} \in Q_{\theta,S}^{BNE}(x)\ P_x - a.s.$ for some $S \in \mathcal{S}$. Hence, $\Theta_I^{BCE} \subseteq \Theta_I^{BNE}(\mathcal{S})$. $\square$

Proposition 2. *Assume that:*

1. *The map* $\theta_\pi \to \pi_i(y, \varepsilon_i; x, \theta_\pi)$ *is continuous for all* i, x, y *and* ε_i, *the quantity*

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

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

2. *The parameter space* Θ *is compact;*

3. *The following uniform convergence condition holds:* $\sup_{\theta\in\Theta} \sqrt{n}|G_n(\theta) - G(\theta)| = O_p(1)$;

4. *The sample criterion function* G_n *is stochastically bounded over* Θ_I *at rate* $1/n$.

Then, the set $\hat{\Theta}_I = \{\theta \in \Theta |\, nG_n(\theta) \leq a_n\}$ *is a consistent estimator of* Θ_I^{BCE} *for* $a_n \to \infty$ *and* $\frac{a_n}{n} \to \infty$.

Proof. We want to show that our setup satisfies the condition C.1 in Chernozhukov, Hong and Tamer (2007); the consistency of $\hat{\Theta}_I$ follows by their Theorem 3.1. To this aim, we need to establish that the function $G(\theta)$ is lower semicontinuous.

We start by showing that $\theta \rightrightarrows Q_\theta^{BCE}(x)$ is upper hemi-continuous for all $x \in X$. This correspondence is a compound correspondence between the BCE equilibrium correspondence $\theta \rightrightarrows BCE^x(\theta)$ and the marginal operator $\nu \to \int_{\mathcal{E}} \nu(y, \mathrm{d}\varepsilon)$. The latter is a continuous function mapping into a compact set. For the the equilibrium correspondence: consider a sequence $\theta^k \to \overline{\theta} \in \Theta$, for $\left\{\theta^k\right\}_{k=1}^{\infty} \in \Theta$, and a corresponding sequence $\{\nu_k\}_{k=1}^{\infty}$ such that $\nu_k \in BCE^x(\theta^k)$ for all k, and ν_k converges to $\overline{\nu}$. To see that $\overline{\nu} \in BCE^x(\bar{\theta})$, notice that (i) consistency of $\overline{\nu}$ follows for the continuity of the map $\theta_\varepsilon \to F(\cdot;\theta_\varepsilon)$ and absolute continuity of $\nu^m(y;\cdot)$, and (ii) incentive compatibility of $\overline{\nu}$ results from the continuity of $\theta_\pi \to \pi_i(\cdot;x,\theta_\pi)$ (this is shown by contradiction, as in Milgrom and Weber, 1985). Therefore the correspondence Q_θ^{BCE} is upper hemi-continuous.

Then, the map

$$\tilde{h}:\theta \to h\left(b;Q_\theta^{BCE}(x)\right) = \sup_{q\in Q_\theta^{BCE}(x)} b^T q$$

is upper semicontinuous (Lemma 17.30 in Aliprantis and Border, 1994), for all values of x, b. It follows that the map $\theta \to -h\left(b;Q_\theta^{BCE}(x)\right)$ is lower semicontinuous, and so is $\theta \to \sup_{b\in B}\left(b^T P_{y|x} - h\left(b;Q_\theta^{BCE}(x)\right)\right)$, point-wise supremum of a family of lower semicontinuous functions (Proposition 2.41 in Aliprantis and Border 1994). Hence, the function $G(\theta)$ is lower semicontinuous: for a sequence $\theta_n \to \theta$ in Θ:

$$\begin{aligned}\liminf_{n\to\infty} G(\theta_n) &= \liminf_{n\to\infty}\int \sup_{b\in B}\left[b^T P_{y|x} - h\left(b;Q_{\theta_n}^{BCE}(x)\right)\right]\mathrm{d}P_x\\ &\geq \int \liminf_{n\to\infty}\sup_{b\in B}\left[b^T P_{y|x} - h\left(b;Q_{\theta_n}^{BCE}(x)\right)\right]\mathrm{d}P_x\\ &\geq \int \sup_{b\in B}\left[b^T P_{y|x} - h\left(b;Q_{\theta_n}^{BCE}(x)\right)\right]\mathrm{d}P_x = G(\theta)\end{aligned}$$

where the first inequality holds by Fatou's Lemma, and the second inequality holds for the lower semi continuity of $\theta \to \sup_{b\in B}\left(b^T P_{y|x} - h\left(b;Q_\theta^{BCE}(x)\right)\right)$. □

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:

$$P_{y_2=1|x} = \int_{\left\{\varepsilon_2:\varepsilon_2\geq -x_c^T\beta^C - x_2^T\beta_2^E\right\}} \mathrm{d}F_2(\cdot;\theta_\varepsilon), \tag{7.1}$$

where $F_i(\cdot;\theta_\varepsilon)$ is the marginal over ε_i of $F(\cdot;\theta_\varepsilon)$. The model implies the following link between the observables and the structure, for all $x \in X$ and $\nu^x \in BCE^x(\theta)$:

$$\begin{aligned} P_{y_2=1|x} &= \nu^x\left([y_1=1,y_2=1]\right)+\nu^x\left(\left[y_1=0,y_2=1,\varepsilon:\varepsilon_2<-x_c^T\beta^C-x_2^T\beta_2^E\right]\right)+ \\ &+ \nu^x\left(\left[y_1=0,y_2=1,\varepsilon:\varepsilon_2\geq-x_c^T\beta^C-x_2^T\beta_2^E\right]\right) \end{aligned}$$

Assume $\beta_{1k}^E > 0$ without loss of generality, and let $x_{1k} \to -\infty$. Conditional on such values of x, $\pi_1(1, y_2, \varepsilon_1; x, \theta_\pi) < 0$ for all values of y_2 $\varepsilon_1 - a.s.$ By the equilibrium optimality condition, $\nu^x(y_1 = 1|y_2, \varepsilon_1) = 0$ whenever $\pi_1(1, y_2, \varepsilon_1; x, \theta_\pi) < 0$. It follows that:

$$\lim_{x_{1k}\to-\infty}\nu^x\left([y_1=1,y_2=1]\right) \leq \lim_{x_{1k}\to-\infty}\int_{\mathcal{E}_1}\nu^x\left([y_1=1]\,|\varepsilon_1\right)\mathrm{d}F_1\left(\cdot;\theta_\varepsilon\right)=0.$$

Moreover, $\lim_{x_{1k}\to-\infty}\nu^x\left(\left[y_1=0,y_2=1,\varepsilon:\varepsilon_2<-x_c^T\beta^C-x_2^T\beta_2^E\right]\right) = 0$, as in the limit $\varepsilon_2 < -x_c^T\beta^C - x_2^T\beta_2^E$ implies $y_2 = 0$. For a similar application of the (IC) property of BCE,

$$\nu^x\left(\left[y_1=0,y_2=1,\varepsilon:\varepsilon_2\geq-x_c^T\beta^C-x_2^T\beta_2^E\right]\right)=\int_{\left\{\varepsilon_2:\varepsilon_2\geq-x_c^T\beta^C-x_2^T\beta_2^E\right\}}\mathrm{d}F_2\left(\cdot;\theta_\varepsilon\right).$$

The result in equation 7.1 follows; this equation describes a single-agent threshold crossing model: under Assumption 3, $\left(\beta^C,\beta_2^E\right)$ and F_i are point-identified (Manski, 1988).
Player 1's parameter β_1 is identified by asymmetric argument. To prove identification of Δ parameters, consider instead $x_{1k} \to \infty$; the same steps lead to:

$$\lim_{x_{1k}\to\infty}P_{y_2=1|x}=\int_{\left\{\varepsilon_2:\varepsilon_2\geq-x_c^T\beta^C-x_2^T\beta_2^E-\Delta_1\right\}}\mathrm{d}F_2\left(\cdot;\theta_\varepsilon\right).$$

2. Let β, Δ be identified. We can derive (non-sharp) bounds on the joint distribution of payoff types $F(\varepsilon;\theta_\varepsilon)$. Let

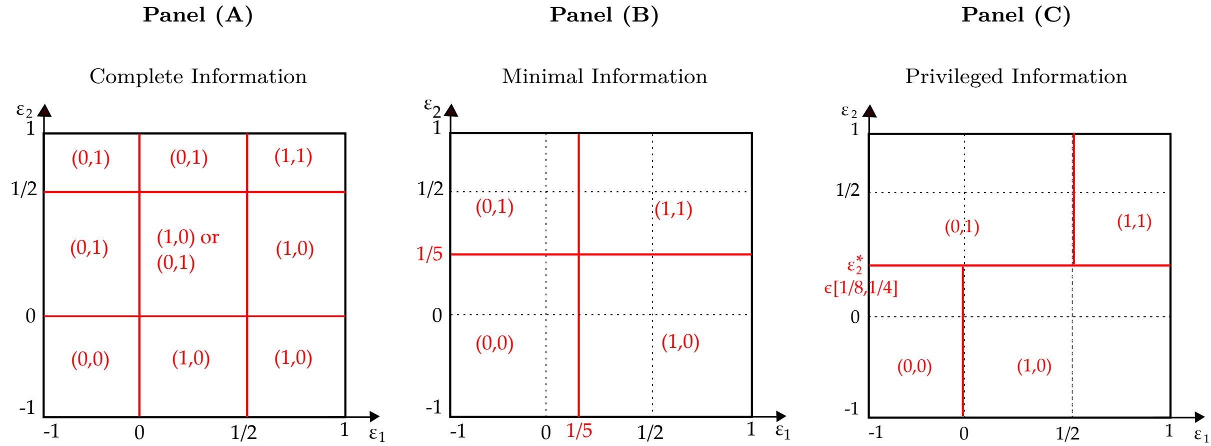
$$\underline{\mathcal{E}}^{(1,1)}\left(x,\theta\right)=\left\{\varepsilon_1\geq-x_c^T\beta^C-x_1^T\beta_1^E-\Delta_2,\varepsilon_2\geq-x_c^T\beta^C-x_2^T\beta_2^E-\Delta_1\right\}.$$

For any $x \in X$, (IC) implies that for $\varepsilon \in \underline{\mathcal{E}}^{(1,1)}$ we have $\nu^x([y = (1,1)]\,|\varepsilon) = 1$ for every $\nu^x \in BCE^x(\theta)$. Similarly define a region $\underline{\mathcal{E}}^y(x,\theta)$ for any action profile y. Consider the partition of $\mathcal{E}$ formed by $\{\underline{\mathcal{E}}^y(x,\theta)\}_{y\in Y}$ and $\widetilde{\mathcal{E}}(x,\theta) = \mathcal{E}/\left(\cup_{y\in Y}\underline{\mathcal{E}}^y(x,\theta)\right)$. We can then construct the bounds:

$$\int_{\underline{\mathcal{E}}^y}\mathrm{d}F\left(\cdot;\theta_\varepsilon\right)\leq P_{y|x}\leq\int_{\underline{\mathcal{E}}^y\cup\tilde{\mathcal{E}}}\mathrm{d}F\left(\cdot;\theta_\varepsilon\right);$$

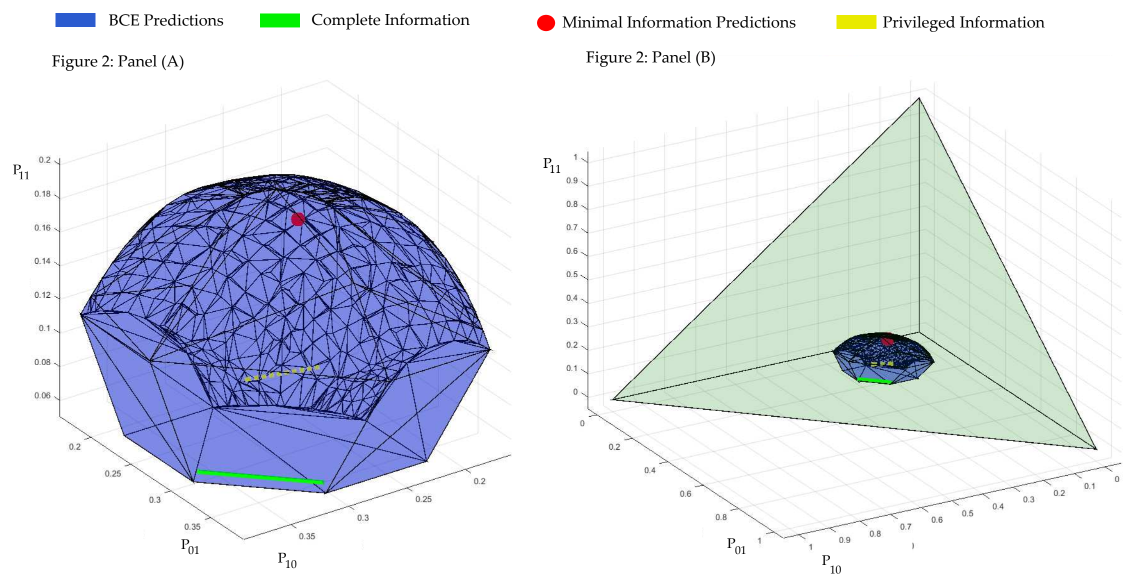
Variation in x shifts the regions $\overline{\mathcal{E}}$ and $\overline{\mathcal{E}} \cup \tilde{\mathcal{E}}$, and provides useful restrictions on θ_ε. $\square$

Figure 1: Information and Equilibrium Predictions



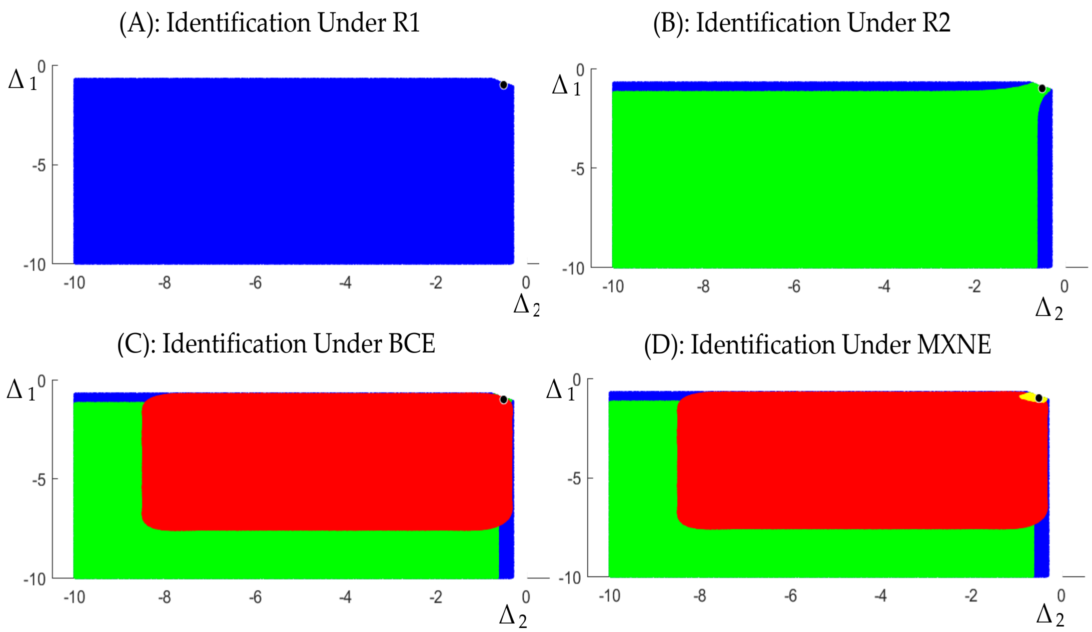
Notes: We represent BNE outcomes in the space $(\varepsilon_1, \varepsilon_2)$ for the two-player entry game in Example 1, with payoffs $\pi_i(y, \varepsilon) = y_i \left(-\frac{1}{2} y_j + \varepsilon_i\right)$ for $i = 1, 2$ and $\varepsilon_i \overset{iid}{\sim} U[-1, 1]$. (A) represents complete information pure-strategy Nash Equilibrium outcomes, (B) represents minimal information outcomes, (C) represents privileged information outcomes.

Figure 2: BCE Predictions



Note: - 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 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 $\varepsilon_i \overset{iid}{\sim} U[-1,1]$. The axes represent probabilities of outcomes P_y. (A) shows the set Q_θ^{BCE} containing the BNE predictions under different restrictions on information. (B) shows the set of BCE predictions inside the unit simplex.

Figure 3: Behavioral Assumptions and Identification



Note: - We represent the identified sets for Δ_1, Δ_2 under different restrictions on behavior in a two-player game with payoffs $\pi_i = y_i\,(y_j \Delta_i + \varepsilon_i)$, $\varepsilon_i \sim N\,(0, 1)$. Data are generated by Nash Equilibrium play with complete information. In the region with multiple equilibria we select with equal probability each of the three pure and mixed equilibria. The black dot represents $\Delta_1 = -1/2$ and $\Delta_2 = -1$, true parameters in the DGP. (A) represents, in blue, the identified set under the assumption of Level-1 rationality. In (B) we add, in green, the identified set under Level-2 rationality and complete information. The sets in (A) and (B) are not bounded from below. (C) includes in red Θ_I^{BCE}; in (D) we add, in yellow, the set $\Theta_I^{BNE}(\bar{S})$.

Table 1: Information and Identification

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

Note: - We report the identified sets for the two players entry model with payoffs $\pi_i(y, \varepsilon_i; \Delta) = y_i(\Delta y_{-i} + \varepsilon_i)$ for $i = 1, 2$ and $\varepsilon_i \sim U[-1, 1]$. The non-sharp identified sets $\tilde{\Theta}_I^{BNE}(\mathcal{S}')$ are obtained under restrictive assumptions on information $\mathcal{S}'$ (corresponding to rows) and true information structures S_0 (corresponding to columns). The true value of the parameter in the data generating process is $\Delta_0 = -0.5$. For $S_0 = S^P$, we generate the data with the equilibrium corresponding to the threshold $\overline{\varepsilon}_2 = 3/16$.

Table 2: Identification with Finite Support

Panel (A): $X^{'}$	β^C	β_i	Δ_1	Δ_2	ρ
θ_0	1	1	-1	-1	-
$S_0 = \bar{S}$	[.89,1.04]	[.89,1.04]	[-2.19,-.82]	[-2.19,-.82]	-
$S_0 = \underline{S}$	[.83,1.13]	[.89,1.21]	[-1.65,-.79]	[-1.65,-.80]	-
$S_0 = S^P$	[.79,1.04]	[.89,1.13]	[-2.04,-.72]	[-2.04,-.85]	-
Panel (B): $X^{''}$					
θ_0	1	1	-1	-1	-
$S_0 = \bar{S}$	[.95,1.04]	[.95,1.07]	[-1.06,-.88]	[-1.06,-.88]	-
$S_0 = \underline{S}$	[.95,1.04]	[.95,1.08]	[-1.06,-.88]	[-1.06,-.88]	-
$S_0 = S^P$	[.95,1.04]	[.95,1.05]	[-1.06,-.88]	[-1.06,-.88]	-
Panel (C): $X^{'}$ and correlated payoff types					
θ_0	1	1	-1	-1	0.8
$S_0 = \bar{S}$	[0.75 ,1.2]	[0.82 ,1.26]	[-1.83 ,-0.7]	[-1.83 ,-0.7]	[0.12 ,0.8]

Note: - We report projections of the identified sets for the two-player game with payoffs $\pi_i\left(y, \varepsilon_i; x, \theta_\pi\right) = y_i\left(x_c^T \beta^C + x_i^T \beta_i^E + \Delta_{-i} y_{-i} + \varepsilon_i\right)$ for $i = 1, 2$. Payoff types $\varepsilon_i \sim N(0,1)$ in (A) (B), and $\varepsilon \sim N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 0 & \rho \\ \rho & 0 \end{pmatrix}$ in (C). The first row in each panel reports the true parameters θ_0; subsequent rows report projections of Θ_I^{BCE} for different assumptions on S_0, the information structure of the game that generates the data. (A) and (C) report sets for data generated with $x \in X^{'}$, (B) reports sets with $x \in X^{''}$. Computational details are in Appendixes A and C.

Table 3: Descriptive Statistics and Regressions

Panel (A): 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

Panel (B): Regressions of Market Structure on Presence of Large Malls

Model	Linear Regression	Ordered probit	Linear Regression	Ordered probit
Variable (S.E. in parentheses)	# of Supermarkets		# of Players in Market	
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

Note: - Panel (B) reports results from linear regressions and ordered probit models. The dependent variable is the number of supermarkets of at least 1500 m^2, or the number of supermarket players. Market size is the product of population and log of tax income per capita. All regressions include fixed effects for 13 regions. Values of R^2 refer to pseudo-R^2 for the ordered probit regressions.

Table 4: Confidence Sets

Parameter	Weak Assumptions on Info - BCE	Complete Information - Nash
Constant	[-2.15 , -0.21]	[-3.26, -1.51]
Market Size	[3.00, 7.64]	[3.67, 6.23]
Home-region:		
Cooperatives	[-0.91, 1.95]	[-0.21, 1.16]
Indep. Italian Supermarket Groups	[-0.39, 2.62]	[-0.14, 1.66]
French Supermarket Groups	[-1.46, 1.96]	[-0.50, 1.15]
Presence of Large Malls:		
Cooperatives	[-3.26, 1.79]	[-2.37, 0.45]
Indep. Italian Supermarket Groups	[-3.77, 1.49]	[-2.63, -0.53]
French Supermarket Groups	[-2.94, 1.02]	[-4.39, -0.19]
Competitive Effects:		
Cooperatives	[-5.30, -1.11]	[-2.40, -0.73]
Indep. Italian Supermarket Groups	[-6.11, -1.69]	[-2.45, -1.34]
French Supermarket Groups	[-7.12, -1.55]	[-3.46, -0.39]
ρ - Correlation Of Unobservable Profitability	[0.36, 0.96]	[0.90, 0.99]

Note: - We report estimates for the game-theoretic model. For each parameter value, we report projections of C_n, the .95 confidence set for identified parameters. See Appendices A and C for computational details.

Table 5: Counterfactual Change in Probability of Outcomes

Outcome	Weak Assumptions on Info - BCE	Complete Information - Nash
No Entry	[-0.30, 0.28]	[-0.61, 0.04]
Entry by Cooperatives	[-0.20, 0.45]	[0.03, 0.78]
Entry by Italian Groups	[-0.16, 0.60]	[-0.19, 0.78]
Entry by French Groups	[-0.13, 0.53]	[-0.21, 0.64]
Entry by at least 1 Player	[-0.09, 0.35]	[-0.04, 0.61]
Entry by at least 2 Players	[-0.26, 0.42]	[0.05, 0.47]

Note: - We report counterfactual change in probability of market structure outcomes, or $\left[\min_{\theta\in C_n} D_{\hat{Y}}(\theta), \max_{\theta\in C_n} D_{\hat{Y}}(\theta)\right]$ for both the model with weak assumptions on information and for the complete information model. Additional counterfactual results are available in Appendix D in Supplementary Materials.

Supplementary Materials - For Online Publication

Appendix C - Further Computational Details

Computation of Identified Sets Θ_I^{BCE}

We describe in this appendix how we deal with numerical error when computing Θ_I^{BCE} to construct Figure 3 and Table 2 in the main text. The identified set is defined in Section 3.4 as:

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

where $G(\theta) = \int_X \sup_{b \in B} \left[b^T P_{y|x} - h\left(b; Q_\theta^{BCE}(x)\right)\right] \mathrm{d}P_x$. Appendix A outlines how to compute $G(\cdot)$, and the choice of discretization for $\mathcal{E}$; we denote with $\check{G}(\cdot)$ the computed $G(\cdot)$.

As a high-dimensional search over the whole set Θ is infeasible, we conduct a search over a subset $\check{\Theta}$. Moreover, since by construction $\check{G}(\cdot) > 0$, we specify a threshold and report the computed analog of the identified set:

$$\check{\Theta}_I^{BCE} = \left\{\theta \in \check{\Theta} | \check{G}(\theta) \leq c_I\right\}.$$

There is no general rule to construct an upper bound for this discretization error that is valid for every game and data generating process. However, for the two-player binary game with independent payoff types considered in Table 2, we find that r^{-1} (where r is the dimension of the discrete grid of ε_i that we use to compute $\check{G}(\cdot)$) is an upper bound of the discretization error if we restrict $Q_\theta^{BCE}(x)$ to $Q_\theta^{PSNE}(x)$. Since r^{-1} is representative of the order of magnitude of the discretization error, we use $c_I = r^{-1}$. Our findings on the informativeness of identified sets are very similar if we use higher values for c_I.

To construct $\check{\Theta}$, we proceed sequentially. We first specify $\check{\Theta}_1$ as a large Halton set of points around θ_0, then find:

$$Bds = \left[\left(\min_{\theta^k}\left\{\theta \in \check{\Theta}_1 : \check{G}(\theta) \leq c_I\right\}\right)_{k=1,\ldots,d_\Theta}, \left(\max_{\theta^k}\left\{\theta \in \check{\Theta}_1 : \check{G}(\theta) \leq c_I\right\}\right)_{k=1,\ldots,d_\Theta}\right]$$

and construct $\check{\Theta}_2$ as another Halton set within $Bds \times 1.2$. This procedure is aimed at constructing more precise boundaries for the identified set. Increasing the umber of points in $\check{\Theta}_1$ and $\check{\Theta}_2$ increases the precision in the computation of the identified set, at the cost of computing time. For Table 2, we use $|\check{\Theta}_1| = 20,000$ and $|\check{\Theta}_2| = 5,000$.

Computation of Identified Sets $\Theta_I^{BNE}\left(\overline{S}\right)$

In Figure 3 in the main text we compute the sharp identified set under the assumption of complete information and Nash equilibrium behavior, allowing for mixed strategies. The sharp identified set for this case can be obtained by first defining the criterion function:

$$G^{MXNE}(\theta) = \sup_{b \in Dir}\left[b^T P_{y|x} - \sup_{p \in Q_\theta^{MXNE}(x)} b^T p\right]_+ \tag{.1}$$

where Dir denotes the core-determining class (Galichon and Henry, 2011) and $Q_\theta^{MXNE}(x_j)$ contains the Nash equilibrium predictions for a game with covariates x and parameters θ. Since Dir is a discrete set, the computation of G^{MXNE} is fast for games with a small number of players and actions. Then, we have:

$$\Theta_I^{BNE}\left(\overline{S}\right) = \left\{\theta \in \check{\Theta} \middle| \, G^{MXNE}(\theta) = 0\right\}.$$

Figure 3 also shows the the identified sets under different behavioral assumptions, R1 and R2. The computation of the corresponding identified sets is analogous to our description of the construction of $\Theta_I^{BNE}\left(\overline{S}\right)$. Under the assumptions of R1 and R2, respectively, we obtain the functions G^{R1} and G^{R2} by substituting $Q_\theta^{R1}(x)$ and $Q_\theta^{R2}(x)$ for Q_θ^{MXNE} into the function G^{MXNE}. Notice that, as the set of predictions is relatively simple, the computation of Q_θ^{MXNE} (as well as of $Q_\theta^{R1}(x)$ and of $Q_\theta^{R2}(x)$) does not involve numerical simulation of the values of ε.

Computation of Confidence Sets C_n for Θ_I^{BCE}

To construct a confidence set s C_n for parameters in the identified sets Θ_I^{BCE} we follow the procedure outlined in Ciliberto and Tamer (2009). The procedure is based on the values of the empirical criterion G_n, whose computation is described in Appendix A. We get to our confidence set via the following steps:

1. We construct deterministic parameter grids using Halton sets around the parameter values of Probit regressions, and select among these 40 starting points for a Simulated Annealing routine, which runs for 10,000 iterations.

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

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

3. We extract $T = 100$ subsamples of size $n_t = n/4$. Subsample size can be an important tuning parameter in this class of models, as argued by Bugni (2014). We follow Ciliberto and Tamer (2009) in the choice of this parameter. For each subsample s, we compute the criterion function using the subsampled observations, so that:

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

and then we find $g_n^s = \min_{\theta \in \Theta} G_n^s(\theta)$ running a Nelder-Mead algorithm.

4. We choose the cutoff value $\hat{c}_0 = n g_n \cdot 1.25$, and define the set:

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

5. For all $\theta \in \hat{\Theta}_I(\hat{c}_0)$, we obtain then $\tilde{G}_n^s(\theta) = G_n^s(\theta) - g_n^s$ and the threshold $\hat{c}_1(\theta)$ as 95th percentile of the distribution across subsamples of the statistic:

$$\tilde{L}_n^s(\theta) = n_t\left(G_n^s(\theta) - g_n^s\right).$$

We compute then

$$\hat{c}_1 = \sup_{\theta \in \hat{\Theta}_I(\hat{c}_0)} \hat{c}_1(\theta),$$

and

$$\hat{\Theta}_I(\hat{c}_1) = \left\{\theta \in \check{\Theta} |\, n\tilde{G}_n(\theta) \leq \min\left(\hat{c}_1, \hat{c}_1(\theta)\right)\right\}.$$

6. Iterating steps 4,5 we obtain $\hat{c}_2$ and report the confidence set:

$$C_n = \left\{\theta \in \check{\Theta} |\, n\tilde{G}_n(\theta) \leq \min\left(\hat{c}_2, \hat{c}_2(\theta)\right)\right\}.$$

Further iterations of this procedure do not alter significantly our results.

We report results for confidence sets for parameters in the identified sets. For both Θ_I^{BCE} and $\Theta_I^{BNE}\left(\overline{S}\right)$, constructing confidence sets for the identified set, as opposed to constructing confidence sets for all points in the identified set, yields similar results (as in Ciliberto and Tamer,2009).

Computation of Confidence Sets for $\Theta_I^{BNE}\left(\overline{S}\right)$

The construction of the confidence set for parameters in $\Theta_I^{BNE}\left(\overline{S}\right)$ is analogous to the procedure followed to compute the confidence set under the assumption of BCE behavior, except that it is based on the empirical criterion function:

$$G_n^{PSNE}(\theta) = \frac{1}{n}\sum_{j=1}^{n} \sup_{b\in Dir}\left[b^T \hat{P}_{y|x_j} - \sup_{q\in Q_\theta^{PSNE}(x_j)} b^T q\right]_+ ,$$

where Dir contains vectors corresponding to core-determining class (Galichon and Henry, 2011) and $Q_\theta^{PSNE}(x_j)$ contains the pure-strategy Nash equilibrium predictions for a game with covariates x_j and parameters θ. We limit Nash equilibria to pure-strategy to maintain the parallel with Ciliberto and Tamer (2009), but the extension to mixed strategy is immediate and can be done by considering the empirical analogue of (.1). The confidence set for parameters identified under the assumption of pure-strategy Nash equilibrium and complete information is obtained going through the same steps 1.-6. described for the computation of C_n, where G_n is substituted with G_n^{PSNE}.

Appendix D - Additional Counterfactual Results

Tables 1 and 2 report additional counterfactual results and complement Table 5 in the main text. Table 1 reports counterfactual results for the model with weak assumptions on information. Panel (A) reports bounds on the difference in upper bound probabilities of market structure. Formally, let

$$\bar{D}_{Y'}(\theta, x) = \left(q_{Y'}^{CF}(\theta, x) - q_{Y'}(\theta, x)\right),$$

where x indexes the eight geographical markets that we consider in this counterfactual. We report in this table the bounds:

$$\left[\min_{\theta\in C_n} \bar{D}_{Y'}(\theta, x),\ \ \max_{\theta\in C_n} \bar{D}_{Y'}(\theta, x)\right]$$

for each market, as well as the median value across markets. Panel (B) reports instead bounds on the difference in lower bound probabilities. Formally, let

$$\underline{q}_{Y'}(\theta, x) = \min_{\nu\in BCE^x(\theta)} \sum_{y\in Y'} \int \nu(y, \mathrm{d}\varepsilon),$$

and define for each market x the difference in lower bound probabilities:

$$\underline{D}_{Y'}(\theta, x) = \left(\underline{q}_{Y'}^{CF}(\theta, x) - \underline{q}_{Y'}(\theta, x)\right).$$

We report in this table the bounds:

$$\left[\min_{\theta \in C_n} \underline{D}_{Y'}(\theta, x), \;\; \max_{\theta \in C_n} \underline{D}_{Y'}(\theta, x)\right]$$

for each market, as well as the median value across markets. Table 2 reports the same results for the model estimated under the assumption of complete information. Comparisons across the two tables yield the same insights that emerge from the discussion of the results in Table 5 in the main text.

[Table 1 about here.]

[Table 2 about here.]

Table 3 contains a full description of the markets included in our counterfactual exercise.

[Table 3 about here.]

Appendix E - 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 2. 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:

$$\pi_i(y, \eta) = y_i \left(\Delta y_{-i} + \eta_i^1 + \eta_i^2\right),$$

and payoff types η are distributed according to:

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

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:

$$\varepsilon_{-i} | \left(\eta_i^1, \eta_{-i}^1, \eta_i^2\right) \sim N\left(\eta_{-i}^1, 1-\sigma^2\right). \tag{.3}$$

We want to recast this model so that it fits the framework of Section 2, 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^1_{-i} as the signal that player i gets on ε_{-i}, and η^1_i as what player i knows that $-i$ knows about her payoff, so that $(\tau^1_i, \tau^2_i) = (\eta^1_i, \eta^1_{-i})$. It follows that $(\tau^1_i, \tau^2_i) = (\tau^2_{-i}, \tau^1_{-i})$, so signals are public. The distribution of ε is:

$$P_\varepsilon = N\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left(\begin{array}{cc} 1 & \sigma^2\rho \\ \sigma^2\rho & 1 \end{array}\right)\right).$$

The joint distribution of signals and redefined payoff shocks, derived from (.2) is thus:

$$\left(\begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \tau^1_1 \\ \tau^2_1 \end{array}\right) \sim N\left(\left(\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array}\right), \left(\begin{array}{cccc} 1 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \\ \sigma^2\rho & 1 & \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 & \sigma^2 & \sigma^2\rho \\ \sigma^2 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \end{array}\right)\right). \tag{.4}$$

Notice that (.4) implies that the belief of player i about ε_{-i} conditional on her information set is:

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

which is identical to the belief (.3).

Appendix F - BMM Representation of the Identified Set

Beresteanu, Molchanov and Molinari (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. *Assume that:*

1. *The discrete set of strategy profiles of the game, Y, is finite.*

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

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

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

Let us adapt our notation and denote the set of BCE equilibrium distributions ν with $BCE_\theta(x)$, for any given realization of x.. Considering $x(\omega)$ as a random vector, $BCE_\theta(x(\omega)) = BCE_\theta(\omega)$ is a random set. Let $\mathrm{Sel}(BCE_\theta)$ denote the set of all $\nu(\omega)$, measurable selections of $BCE_\theta(\omega)$. 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:

$$q(\nu(\omega)) = \nu(\cdot|\varepsilon(\omega)) \in \mathbb{P}_Y,$$

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

$$\tilde{Q}_\theta = \{q(\nu) : \ \nu \in \mathrm{Sel}(BCE_\theta)\}.$$

Then the conditional Aumann expectation of this random set is:

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

Notice however that:

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

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.

References

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

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

Bugni, F. A. (2016): "Comparison of inferential methods in partially identified models in terms of error in coverage probability," *Econometric Theory*, 32, 187–242.

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.

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.

Milgrom, P. R. and R. J. Weber (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operations Research*, 10, 619–632.

Table 1: Additional Counterfactual Results: BCE

Panel (A):	Bounds on Difference in Upper Bound Probabilities					
Outcome:	No Entry	Entry by Coop.	Entry by IT	Entry by FR	Entry by >=1	Entry by >=2
Mkt. 1	[-0.34, 0.27]	[-0.09, 0.32]	[-0.11, 0.59]	[-0.07, 0.46]	[-0.03, 0.16]	[-0.21, 0.45]
Mkt. 2	[-0.31, 0.26]	[-0.18, 0.47]	[-0.14, 0.64]	[-0.12, 0.57]	[-0.07, 0.34]	[-0.27, 0.5]
Mkt. 3	[-0.29, 0.3]	[-0.24, 0.5]	[-0.21, 0.65]	[-0.16, 0.56]	[-0.14, 0.48]	[-0.3, 0.36]
Mkt. 4	[-0.32, 0.24]	[-0.12, 0.3]	[-0.07, 0.39]	[-0.08, 0.5]	[-0.02, 0.19]	[-0.2, 0.44]
Mkt. 5	[-0.31, 0.32]	[-0.17, 0.42]	[-0.18, 0.69]	[-0.12, 0.55]	[-0.08, 0.38]	[-0.28, 0.49]
Mkt. 6	[-0.37, 0.35]	[-0.24, 0.53]	[-0.12, 0.57]	[-0.15, 0.56]	[-0.07, 0.4]	[-0.25, 0.45]
Mkt. 7	[-0.31, 0.25]	[-0.18, 0.48]	[-0.14, 0.66]	[-0.14, 0.59]	[-0.07, 0.32]	[-0.26, 0.48]
Mkt. 8	[-0.2, 0.22]	[-0.35, 0.57]	[-0.3, 0.59]	[-0.23, 0.45]	[-0.25, 0.52]	[-0.26, 0.21]
Average	[-0.3, 0.28]	[-0.2, 0.45]	[-0.16, 0.6]	[-0.13, 0.53]	[-0.09, 0.35]	[-0.26, 0.42]
Median	[-0.31, 0.27]	[-0.18, 0.47]	[-0.14, 0.62]	[-0.13, 0.56]	[-0.07, 0.36]	[-0.26, 0.45]

Panel (B):	Bounds on Difference in Lower Bound Probabilities					
Outcome:	No Entry	Entry by Coop.	Entry by IT	Entry by FR	Entry by >=1	Entry by >=2
Mkt. 1	[-0.16, 0.03]	[-0.09, 0.16]	[-0.05, 0.32]	[-0.35, 0.34]	[-0.27, 0.34]	[-0.09, 0.33]
Mkt. 2	[-0.34, 0.07]	[-0.07, 0.06]	[-0.04, 0.14]	[-0.07, 0.14]	[-0.26, 0.31]	[-0.07, 0.19]
Mkt. 3	[-0.48, 0.14]	[-0.29, 0.05]	[-0.05, 0.07]	[-0.05, 0.08]	[-0.3, 0.29]	[-0.04, 0.09]
Mkt. 4	[-0.19, 0.02]	[-0.08, 0.23]	[-0.08, 0.28]	[-0.08, 0.29]	[-0.24, 0.32]	[-0.1, 0.31]
Mkt. 5	[-0.38, 0.08]	[-0.17, 0.08]	[-0.03, 0.13]	[-0.06, 0.11]	[-0.31, 0.31]	[-0.06, 0.18]
Mkt. 6	[-0.4, 0.07]	[-0.04, 0.05]	[-0.29, 0.13]	[-0.06, 0.1]	[-0.35, 0.36]	[-0.06, 0.15]
Mkt. 7	[-0.32, 0.07]	[-0.04, 0.06]	[-0.04, 0.17]	[-0.08, 0.13]	[-0.25, 0.31]	[-0.08, 0.19]
Mkt. 8	[-0.52, 0.25]	[-0.03, 0.02]	[-0.09, 0.03]	[-0.05, 0.04]	[-0.22, 0.21]	[-0.03, 0.04]
Average	[-0.35, 0.09]	[-0.1, 0.09]	[-0.09, 0.16]	[-0.1, 0.15]	[-0.28, 0.31]	[-0.07, 0.18]
Median	[-0.36, 0.07]	[-0.08, 0.06]	[-0.05, 0.14]	[-0.07, 0.12]	[-0.27, 0.31]	[-0.07, 0.18]

Table 2: Additional Counterfactual Results: Complete Information - Nash

Panel (A): Bounds on Difference in Upper Bound Probabilities

Outcome:	No Entry	Entry by Coop.	Entry by IT	Entry by FR	Entry by >=1	Entry by >=2
Mkt. 1	[-0.42, 0]	[-0.04, 0.77]	[-0.16, 0.9]	[-0.38, 0.57]	[0, 0.42]	[0.09, 0.81]
Mkt. 2	[-0.75, 0.03]	[0.06, 0.86]	[-0.13, 0.87]	[-0.19, 0.7]	[-0.03, 0.75]	[0.06, 0.52]
Mkt. 3	[-0.76, 0.13]	[0.11, 0.74]	[-0.19, 0.76]	[-0.1, 0.76]	[-0.13, 0.76]	[0.04, 0.21]
Mkt. 4	[-0.25, 0]	[-0.02, 0.96]	[-0.08, 0.83]	[-0.27, 0.62]	[0, 0.25]	[0.05, 0.79]
Mkt. 5	[-0.72, 0.05]	[0.03, 0.83]	[-0.11, 0.84]	[-0.17, 0.69]	[-0.05, 0.72]	[0.08, 0.52]
Mkt. 6	[-0.77, -0.06]	[-0.03, 0.8]	[-0.52, 0.72]	[-0.39, 0.67]	[0.06, 0.77]	[0.02, 0.32]
Mkt. 7	[-0.75, 0.03]	[0.06, 0.86]	[-0.13, 0.88]	[-0.2, 0.7]	[-0.03, 0.75]	[0.06, 0.53]
Mkt. 8	[-0.42, 0.11]	[0.09, 0.44]	[-0.18, 0.44]	[0, 0.38]	[-0.11, 0.42]	[0, 0.04]
Average	[-0.61, 0.04]	[0.03, 0.78]	[-0.19, 0.78]	[-0.21, 0.64]	[-0.04, 0.61]	[0.05, 0.47]
Median	[-0.74, 0.03]	[0.04, 0.82]	[-0.15, 0.83]	[-0.2, 0.68]	[-0.03, 0.74]	[0.05, 0.52]

Panel (B): Bounds on Difference in Lower Bound Probabilities

Outcome:	No Entry	Entry by Coop.	Entry by IT	Entry by FR	Entry by >=1	Entry by >=2
Mkt. 1	[-0.42, 0]	[-0.07, 0.67]	[-0.56, 0.82]	[-0.46, 0.71]	[0, 0.42]	[0.05, 0.8]
Mkt. 2	[-0.75, 0.03]	[-0.23, 0.34]	[-0.72, 0.38]	[-0.76, 0.11]	[-0.03, 0.75]	[0.02, 0.44]
Mkt. 3	[-0.76, 0.13]	[-0.03, 0.47]	[-0.67, 0.5]	[-0.44, 0.38]	[-0.13, 0.76]	[0.01, 0.2]
Mkt. 4	[-0.25, 0]	[-0.07, 0.7]	[-0.48, 0.71]	[-0.48, 0.45]	[0, 0.25]	[0.12, 0.78]
Mkt. 5	[-0.72, 0.05]	[-0.07, 0.6]	[-0.73, 0.59]	[-0.5, 0.41]	[-0.05, 0.72]	[0.05, 0.5]
Mkt. 6	[-0.77, -0.06]	[-0.13, 0.76]	[-0.38, 0.64]	[-0.48, 0.55]	[0.06, 0.77]	[0.01, 0.31]
Mkt. 7	[-0.75, 0.03]	[-0.23, 0.35]	[-0.72, 0.4]	[-0.76, 0.12]	[-0.03, 0.75]	[0.02, 0.46]
Mkt. 8	[-0.42, 0.11]	[-0.04, 0.06]	[-0.37, 0.03]	[-0.23, 0.03]	[-0.11, 0.42]	[0, 0.03]
Average	[-0.61, 0.04]	[-0.11, 0.49]	[-0.58, 0.51]	[-0.51, 0.35]	[-0.04, 0.61]	[0.04, 0.44]
Median	[-0.74, 0.03]	[-0.07, 0.54]	[-0.61, 0.54]	[-0.48, 0.4]	[-0.03, 0.74]	[0.02, 0.45]

Table 3: DESCRIPTION OF COUNTERFACTUAL MARKETS

Market:	POPULATION	SURFACE, IN KM^2	TAX INCOME PER CAPITA, IN EUR
Ancona	105,367	146.73	15,602.05
Aosta	77,822	1627.24	16,306.34
Cesenatico II	56,297	95.33	12,712.45
Empoli	105,156	340.28	13,994.89
Formia	74,402	255.11	11,586.54
Portogruaro	66,839	359.96	13,658.51
Rovereto	79,281	540.14	15,466.23
Trento II	35,768	223.67	14,845.97
Average	75,116.5	448.56	14,271.62
Median	76,112	297.69	14,420.43
Std. Dev.	23,341.05	495.85	1,592.85
Max	105,367	1627.24	16,306.34
Min	35,768	95.33	11,586.54