WARWICK

ECONOMICS

Estimation of Discrete Games with Weak Assumptions on Information

Lorenzo Magnolfi and Camilla Roncoroni

February 2020 No: 1247

Warwick Economics Research Papers

ISSN 2059-4283 (online)

ISSN 0083-7350 (print)

Estimation of Discrete Games with Weak Assumptions on Information*

Lorenzo Magnolfi[†] and Camilla Roncoroni[‡]

First Draft: November, 2015; This Draft: December, 2019

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 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 Niccolò Lomys, Ken Hendricks, Tibor Heumann and Dan Quint for numerous suggestions. Many others, to whom we are grateful, provided excellent comments including seminar participants at John's Hopkins, Northwestern, NYU Stern, Penn State, University of Chicago IO lunch, 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. We gratefully acknowledge financial support from the Cowles Foundation. All errors 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 us 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 information structures. We show that, under the assumption of BCE, 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

[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.

sharp identified set of parameters without explicitly modeling equilibrium selection. These results motivate the use of BCE to estimate the distribution of players' payoffs while being robust with respect to the information structure, thus avoiding misspecification bias due to strong assumptions on information.

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

[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.

formation, we find that the absence of the mall may or may not foster the emergence of a 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 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.[6] Their characterization, developed in the context of theoretical work, inspires our use of a similarly robust framework in empirical applications. More recently, Bergemann, Brooks and Morris (2019) show how to perform counterfactual analysis under a *fixed latent information structure.* We use their method to perform some of the counterfactuals in Section 7, showing that it can help to reduce the width of counterfactual prediction intervals in an applied context.

[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 establish necessary conditions to identify 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]Bergemann and Morris (2013) also discuss identification. While we highlight the relationship between BCE identified sets and BNE identified sets in a discrete game setup, they discuss properties of the BCE identified set for a particular parametrization of a game with a continuum of players, symmetric quadratic payoff functions, and normally distributed uncertainty.

Aradillas-Lopez and Tamer (2008) study identification for a less restrictive solution concept, rationalizability. 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. Yang (2009) performs estimation of discrete games of complete information under Nash equilibrium, using the non-sharp restrictions imposed by Correlated Equilibrium under complete information to simplify computation. The assumption of Correlated Equilibrium under complete information is nested in our approach. Beyond discrete games, which we consider in this article, Bergemann, Brooks and Morris (2017) characterize BCEs of first-price auctions, and Syrgkanis, Tamer and Ziani (2018) use BCE to perform inference in this class of models.

Our study of the effect of the presence of large malls on local supermarkets is related to several articles 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 article, 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. We present counterfactuals in Section 7. Section 8 concludes.

2 Model

We first outline the general class of discrete games that we consider in this article, and then develop our leading example: a two-player entry game.

2.1 A General Empirical Discrete Game

We consider a class of static games, indexed by realizations of covariates $x \in X$. Let $N = \{1, ..., n\}$ be the finite set of players; each player $i \in N$ chooses an action y_i from the finite set Y_i. Both the actions' space $Y = \times_{i \in N} Y_i$ and N are the same across different games. We outline the other primitives of the game in the next subsections, describing separately the payoff structure and the information structure that players face. The game is common knowledge among players.

2.1.1 Payoff Structure

Each player i is characterized by a payoff type $\varepsilon_i \in \mathcal{E}_i \subseteq \mathbb{R}$. Payoff types $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ are distributed according to the cdf $F(\cdot; \theta_\varepsilon)$,[7] parametrized by the finite dimensional vector $\theta_\varepsilon \in \Theta_\varepsilon$: we also refer to this cdf as players' prior. Payoffs to player i, denoted by π_i, depend on her payoff type and on action profiles. Observable covariates x and finite dimensional payoff parameters $\theta_\pi \in \Theta_\pi$ also affect payoffs. To sum up, the function π_i describes player i's payoff for every pair (x, θ_π):

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

A realization of x and a vector of parameters $\theta = (\theta_\pi, \theta_\varepsilon) \in \Theta$ pin down a payoff structure. Throughout the article we assume that ε is independent of x and discuss how to identify parameters θ from data on actions and market observable characteristics x. Although our description of payoff structures embeds some parametric restrictions to preserve the link with the applied literature, most of these assumptions are not necessary for the general discussion of robust identification in Section 3.[8]

Example 1. Consider a game of oligopoly entry such as the one proposed by Bresnahan and Reiss (1991a). Players are firms that can either "Enter" or "Not enter" a market; these actions correspond to $y_i = 1$ and $y_i = 0$, respectively. Economists observe the set of potential entrants making entry decisions in a cross-section of markets, each characterized by a realization of covariates x. Firms earn a profit of zero by not entering; when entering, firm i's profits are $\pi_i(y, \varepsilon_i; x, \theta_\pi) = \Pi_i(y_{-i}; x, \theta_\pi) + \varepsilon_i$. The additive ε_i represents factors that affect firms' (variable) profits or fixed costs and are unobservable to an outside analyst.

2.1.2 Information Structure

Every player i knows the realization of her payoff type ε_i, as well as parameters θ and covariates x. In addition, players receive a private random signal τ_i^x, which may be informative about the full vector of payoff types ε. For a game with covariates x an information structure S^x specifies the set of signals a player may receive and the probability of receiving them. Formally:

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

[7] In this model payoff types may be correlated across players; see also Xu (2014) and Wan and Xu (2015) for models with correlated payoff types.

[8] 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.

where T^x is the subset of a complete, separable metric space and represents the support of the vector of signals $\tau^x = (\tau_1^x, \ldots, \tau_n^x)$. The dimensionality of T^x is unrestricted: signal structures can be very complex objects. The probability kernel $\left\{P_{\tau|\varepsilon}^x \in \mathcal{P}_{T^x} : \varepsilon \in \mathcal{E}\right\}$ is a subset of $\mathcal{P}_{T^x}$, the set of all distributions over the space formed by T^x and its Borel $\sigma-$algebra. The kernel contains the probability distributions of signals τ^x conditional on every realization of ε and allows for correlation of signals across players.

The sets of signals and the distribution of signal vectors, which need not be symmetric across players, depend on x as the information structure may change with observable characteristics of the payoff structure. We denote as S the array that includes information structures corresponding to each realization of x, that is $S = (S^x)_{x \in X}$, and define $\mathcal{S}$ as the set of all possible information structures S:

$$\mathcal{S} = \left\{S : \ \forall x \in X,\ T^x \text{ complete, separable metric space, } P_{\tau|\varepsilon}^x \in \mathcal{P}_{T^x}\right\}.$$

Example 2. (Example 1 continued) In Bresnahan and Reiss (1991a) firm i not only observes its own payoff type ε_i, but also observes $\varepsilon_{-i} = (\varepsilon_1, \ldots, \varepsilon_{i-1}, \varepsilon_{i+1}, \ldots, \varepsilon_n)$, the payoff types of every other potential entrant. The information structure is hence *complete information*, denoted by $\overline{S}$: the signal space coincides with the type space, or $T_i^x = \mathcal{E}$ for every $x \in X$, and players observe perfectly informative signals: $P_{\tau_i|\varepsilon}^x([\tau_i = \varepsilon]) = 1$ for all $\varepsilon \in \mathcal{E}$, $x \in X$, $i \in N$.

2.1.3 Equilibrium

The parameter vector θ and the information structure S summarize the elements of the game that are unknown to the econometrician; a pair (θ, S) pins down a game $\Gamma^x(\theta, S)$ for every x. To specify the data generating process, linking primitives of the game to outcomes, we need an equilibrium notion. We describe strategies for player i as functions $\sigma_i : \mathcal{E}_i \times T_i^x \to \mathcal{P}_{Y_i}$, which map payoff types and signals into distributions over actions, and adopt as a solution concept for this game the standard notion of Bayes Nash Equilibrium.

Definition 1. (Bayes Nash Equilibrium) A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a Bayes Nash Equilibrium (BNE) of the game $\Gamma^x(\theta, S)$ if for every $i \in N$, $\varepsilon_i \in \mathcal{E}_i$ and $\tau_i \in T_i^x$ we have that, whenever for some $y_i \in Y_i$ the corresponding $\sigma_i(y_i \mid \varepsilon_i, t_i) > 0$, then :

$$E_{\sigma_{-i}}\left[\pi_i\left(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi\right) \mid \varepsilon_i, \tau_i\right] \geq E_{\sigma_{-i}}\left[\pi_i\left(y_i', y_{-i}, \varepsilon_i; x, \theta_\pi\right) \mid \varepsilon_i, \tau_i\right], \quad \forall\, y_i' \in Y_i,$$

where the expectation of y_{-i} is taken with respect to the distribution of equilibrium play $\sigma_{-i}(y_{-i} \mid \varepsilon_j, \tau_j) = \Pi_{j \neq i} \sigma_j(y_j \mid \varepsilon_j, \tau_j)$.

The information structure of the game has important implications for Bayes Nash equilibrium. 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 only base their equilibrium behavior on their prior beliefs. We denote as $BNE^x(\theta, S)$ the set of all BNE strategy profiles for the game $\Gamma^x(\theta, S)$.

In addition to BNE, we also introduce the notion of Bayes Correlated Equilibrium (BCE), due to Bergemann and Morris (2013, 2016).

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

1. *Consistent with the prior*: for all $\varepsilon \in \mathcal{E}$, $\tau \in T$,

$$\sum_{y \in Y} \int_{[t \leq \tau]} \int_{[e \leq \varepsilon]} \nu(y, e, t) \, \mathrm{d}t\mathrm{d}e = \int_{[t \leq \tau]} \int_{[e \leq \varepsilon]} P_{\tau|\varepsilon}(t) \mathrm{d}F(e; \theta_\varepsilon) \, \mathrm{d}t;$$

2. *Incentive Compatible*: for all $i, \varepsilon_i, \tau_i, y_i$ such that $\nu(y_i \mid \varepsilon_i, \tau_i) > 0$,

$$E_\nu[\pi_i(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi) \mid y_i, \varepsilon_i, \tau_i] \geq E_\nu[\pi_i(y_i', y_{-i}, \varepsilon_i; x, \theta_\pi) \mid y_i, \varepsilon_i, \tau_i], \qquad \forall y_i' \in Y_i,$$

   where the expectation operator $E_\nu[\cdot \mid y_i, \varepsilon_i, \tau_i]$ is taken with respect to the conditional equilibrium distribution $\nu(y_{-i}, \varepsilon_{-i}, \tau_{-i} \mid y_i, \varepsilon_i, \tau_i)$.

BCE is a generalization of Correlated Equilibrium (Aumann 1974, 1987) 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. Equilibrium is defined as a probability measure ν over outcomes, signals and payoff types.[9] This is in contrast to BNE, which represents equilibrium behavior through strategy functions. Whereas the product structure of BNE implies that correlation in players' actions must correspond to underlying correlation in payoffs or in signals, BCE may feature additional correlation in behavior.

The consistency property of BCE requires the equilibrium distribution ν (via its marginal over payoff types) to reflect common knowledge of the underlying distribution of ε. The incentive compatibility property may be illustrated with the usual mediator metaphor: players receive personalized (i.e. payoff type- and signal-dependent) recommendations from an omniscient mediator, and in equilibrium it is optimal to follow these recommendations.

Example 3. (Example 1 continued) In the oligopoly entry game of complete information, if a BNE strategy σ prescribes that firm i enters a market it must be that entry is optimal given

[9] Throughout the article we also use the symbol ν for conditional distributions derived from the joint measure. For instance, in the incentive compatibility condition of Definition 2, to keep notation light, we use the symbol ν to represent $\nu_{y_{-i}, \varepsilon_{-i}, \tau_{-i} | y_i, \varepsilon_i, \tau_i}$.

the firm's knowledge of ε and equilibrium expectations, or $E_\sigma \left[\Pi_i \left(y_{-i}; x, \theta_\pi\right) \mid \varepsilon\right] \geq -\varepsilon_i$. BCEs are instead distributions over the set of actions and types $(Y \times \mathcal{E})$ whose marginal over actions coincides with the common prior over payoff types, and such that whenever entry is recommended with a positive probability, or $\nu\left(y_i = 1 \mid \varepsilon\right) > 0$, then it must be $E_\nu \left[\Pi_i \left(y_{-i}, \varepsilon_i; x, \theta_\pi\right) \mid y_i = 1, \varepsilon\right] \geq -\varepsilon_i$ for all i, ε. In this case the entry decision must be optimal conditional on the mediator's recommendation of entering the market.

2.1.4 Alternative Sets of Information Structures

We conclude the description of the model with a remark on its generality. The set of information structures $\mathcal{S}$ embeds the assumption that players know *at least* their own payoff type and everything that is known to the econometrician. We adopt this restriction because it is plausible in the application we consider and strikes a balance between maintaining weak assumptions and providing identification power in practice.

Dropping the restriction that players know their own payoffs and that covariates are common knowledge results in a larger, more general set of admissible information structures.[10] Conversely we may assume that players have *at least* signals with a certain degree of informativeness about other players' payoff types: this would result in a smaller set of information structures than the one we consider. Our identification results in Section 3 can be adapted to both of these approaches; we leave a more formal discussion of these results to Appendix G in the Supplemental Materials online.

2.2 Illustration: the Two-player Entry Game

We introduce here a model that is both our leading example and is closely related to our application in Section 6: a two-player entry game. This game specializes the market entry model of Example 1 to the case of two firms, so that $N = \{1, 2\}$. 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 $(0, 0)$. In line with previous literature (e.g. Bresnahan and Reiss 1991a; Berry, 1992; Tamer, 2003) we specify payoffs as linear functions of covariates:

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

so that the payoff parameter vector is $\theta_\pi = (\beta_i, \Delta_i)_{i=1,2}$. The parameter Δ_i, called *competitive effect*, quantifies how entry by firm i affects firm $-i$'s payoffs. Payoffs are:

[10] For instance, 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.

Player 1 / Player 2:	0	1
0	(0, 0)	$\left(0,\ x_2^T\beta_2+\varepsilon_2\right)$
1	$\left(x_1^T\beta_1+\varepsilon_1,\ 0\right)$	$\left(\begin{array}{c} x_1^T\beta_1+\Delta_2+\varepsilon_1, \\ x_2^T\beta_2+\Delta_1+\varepsilon_2 \end{array}\right)$

A parametric assumption on the joint distribution of payoff types (e.g. iid Uniform) completes the definition of the payoff structure. In what follows we also use a simplified version of this payoff structure: the *one-parameter entry game.* In this two-player entry game there are no covariates, competitive effects are equal across firms, so that $\Delta_1 = \Delta_2 = \Delta$, and payoff types are iid Uniform over the interval $[-1, 1]$.

2.2.1 Information Structures and Bayes Nash Equilibrium

In addition to the *complete information* structure $\bar{S}$ of Example 2, many other salient information structures fit our framework: we describe some of these in the context of the two-player entry game. For instance, in the environment of *minimal information*, denoted by $\underline{S}$, payoff types are fully private information. This is because in this information structure signals are uninformative: $T_i^x = \mathcal{E}_i$ for all $x \in X$, $i \in N$ and $P^x_{\tau_i|\varepsilon}([\tau_i = \varepsilon_i]) = 1$ for all $\varepsilon = (\varepsilon_i, \varepsilon_{-i}) \in \mathcal{E}$, $x \in X$, $i \in N$.

Both the environments of minimal information and complete information are symmetric across players. An alternative information structure still contained in the set of information structures $\mathcal{S}$ is *privileged information* S^P, in which only one player knows the type of her opponent.[11] In this case signal spaces for all players are $T_i^x = \mathcal{E}$; for the informed player i, signals τ_i^x are distributed according to $P^x_{\tau_i|\varepsilon}([\tau_i = \varepsilon]) = 1$ for all $\varepsilon \in \mathcal{E}$, $x \in X$, whereas for the uninformed player j signals τ_j^x are distributed according to $P^x_{\tau_j|\varepsilon} = P^x_{\tau_j}$ for some distribution $P^x_{\tau_j}$.

Our model also accommodates the class of *flexible information* structures proposed in Grieco (2014). To recast his model in terms of our definitions of information and payoff structures, decompose payoff types in two additive parts, or $\varepsilon_i = \eta_i^1 + \eta_i^2$ so that private signals are $\tau_i = (\eta^1_{-i}, \eta^1_i)$. The vector (η^1_{-i}, η^1_i) is then the publicly observed component of the payoff type whereas η_i^2 is a privately known component, independent across players.[12]

Definition 1 specifies the condition for a strategy profile to be a BNE for a game with information structure S. As we vary the information structure, equilibria may vary considerably. We illustrate this point in Figure 1, which depicts equilibrium outcomes in the space of payoff types for the one-parameter entry game with $\Delta = -1/2$. In Panel (A) we represent equilibrium outcomes for a game of complete information such as the one in Bresnahan and

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

[12] See Appendix E in Supplementary Materials online for more discussion of this example.

Reiss (1991a) and Tamer (2003). For every realization of ε, common knowledge for players, there are one or two equilibrium outcomes. 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 action 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.]

2.2.2 Bayes Correlated Equilibrium

We further illustrate the properties of BCE for the one-parameter entry game with $\Delta = -1/2$ and minimal information $S = \underline{S}$. In this case BCE distributions are in the set $\mathcal{P}_{\{0,1\}^2 \times [-1,1]^2}$ since in the minimal information structure signals are uninformative.

For each $e \in [-1,1]^2$, BCE distributions in this game must satisfy

$$\sum_{y \in \{0,1\}^2} \int_{[e \leq \varepsilon]} \nu(y, e)\, \mathrm{d}e = \left(\frac{e_1 + 1}{2}\right)\left(\frac{e_2 + 1}{2}\right)$$

for consistency with the prior over payoff types. Moreover, in any BCE ν if player i receives the recommendation to enter with positive probability upon observing ε_i, then $\nu(y_{-i} = 1 \mid y_i = 1, \varepsilon_i) \leq 2\varepsilon_i$. Conversely player i will stay out if $\nu(y_{-i} = 1 \mid y_i = 0, \varepsilon_i) \geq 2\varepsilon_i$. Many BCEs satisfy these constraints; consider for instance the distribution ν':

$\varepsilon_1/\varepsilon_2$:	$\leq 1/5$	$> 1/5$
$\leq 1/5$	$\nu'(0,0,\varepsilon) = \frac{9}{25}$	$\nu'(0,1,\varepsilon) = \frac{6}{25}$
$> 1/5$	$\nu'(1,0,\varepsilon) = \frac{6}{25}$	$\nu'(1,1,\varepsilon) = \frac{4}{25}$

Checking consistency is immediate, as this BCE distribution prescribes pure strategies for every vector of payoff types ε. Incentive compatibility is also satisfied since

$$\nu'(y_{-i} = 1 \mid y_i = 1, \varepsilon_i) = \nu'(y_{-i} = 1 \mid y_i = 0, \varepsilon_i) = 2/5.$$

As another example, consider ν'':

$\varepsilon_1/\varepsilon_2$:	$\leq 1/8$	$> 1/8$
≤ 0	$\nu''(0,0,\varepsilon) = \frac{9}{32}$	$\nu''(0,1,\varepsilon) = \frac{7}{32}$
$> 0,\ \leq 1/2$	$\nu''(1,0,\varepsilon) = \frac{9}{64}$	$\nu''(0,1,\varepsilon) = \frac{7}{64}$
$> 1/2$	$\nu''(1,0,\varepsilon) = \frac{9}{64}$	$\nu''(1,1,\varepsilon) = \frac{7}{64}$

The behavior induced by ν' and ν'' is identical to the behavior described in Panel (B) and Panel (C) of Figure 1, respectively, which represent outcomes of BNE play in the games of incomplete information and privileged information. The correspondence between BCE behavior for the game with $\underline{S}$ and BNEs for different, possibly more informative, information structures anticipates a result we present in the next section: the set of outcomes of BNE play for any information structure can be represented as the set of outcomes of BCE play in the minimal information game. Our identification strategy builds on this result.

3 Identification

3.1 BNE Predictions and Identified Set

After describing a general class of games in the previous section, we investigate identification of payoffs when the econometrician observes cross-sectional data on outcomes y and covariates x. 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*

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

where $\mathcal{P}_Y$ is the set of probability distributions over the finite set Y.

For a game in the class described in Section 2 equilibrium strategies $\sigma \in BNE^x(\theta, S)$ result in predictions on observable behavior:

Definition 3. (BNE Prediction) A BNE prediction for an equilibrium σ of the game $\Gamma^x(\theta, S)$ is a distribution over outcomes q_σ such that

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

Since the set $BNE^x(\theta, S)$ may not be a singleton, i.e. the game $\Gamma^x(\theta, S)$ may have multiple equilibria, and we want to be agnostic about equilibrium selection, implications of equilibrium are summarized by sets of predictions. When data are generated by an *arbitrary* equilibrium selection mechanism, defined as a probability distribution over the set of equilibria $BNE^x(\theta, S)$, the set of predictions is convexified as in Beresteanu, Molchanov and Molinari (2011).[13] The prediction correspondence $Q^{BNE}_{\theta,S} : X \rightrightarrows \mathcal{P}_Y$ describes the set of

[13] Although the assumption that equilibrium selection rules are representable as probability distribution is not fully general (Epstein, Kaido and Seo, 2016), it is standard in the applied literature.

distributions over actions that correspond to BNE predictions in the game $\Gamma^x(\theta, S)$, and is defined as

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

where the operator co$[\cdot]$ takes the convex hull of a set.

Example 4. Consider the one-parameter entry game with $\Delta = -1/2$ and complete information $S = \overline{S}$ of Figure 1, Panel (A). There are three BNE strategies for this game, corresponding to the two pure and one mixed-strategy Nash equilibria, and the corresponding set of predictions (allowing for an arbitrary equilibrium selection) is:

$$Q_{\Delta=-1/2,\overline{S}}^{BNE} = \text{co}\left[\left\{\left(\frac{1}{4}, \frac{3}{8}, \frac{5}{16}, \frac{1}{16}\right), \left(\frac{1}{4}, \frac{5}{16}, \frac{3}{8}, \frac{1}{16}\right), \left(\frac{1}{4}, \frac{11}{32}, \frac{11}{32}, \frac{1}{16}\right)\right\}\right],$$

where vectors q_σ list the probabilities of outcomes $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$.

3.1.1 Data Generating Process and Identified Set

We assume that at least one equilibrium in $BNE^x(\theta_0, S_0)$ exists for every $x \in X$, and the data are generated by BNE play in games $\Gamma^x(\theta_0, S_0)$ characterized by a *true payoff and information structure* (θ_0, S_0). We also allow for $\Gamma^x(\theta_0, S_0)$ to have multiple equilibria, and assume that data are generated by an arbitrary equilibrium selection distribution. The properties of the data generating process are summarized by Assumption 2.

Assumption 2. (Data generating process) *For all $x \in X$, the set $BNE^x(\theta_0, S_0)$ is assumed to be non-empty and the outcomes y are generated by BNE play of the game $\Gamma^x(\theta_0, S_0)$, so that $P_{y|x} \in Q_{\theta_0,S_0}^{BNE}(x)$.*

Given this link between game-theoretic model and observables, we want to recover θ_0 but we do not know (nor attempt to recover) the true information structure S_0. Under Assumptions 1 and 2, the sharp identified set of parameters with *weak assumptions on information* is defined as:

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

The set $\Theta_I^{BNE}(\mathcal{S})$ captures all the information on parameters that can be obtained with the only restriction that the true information structure S_0 belongs to the set $\mathcal{S}$. 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 $Q_{\theta,S}^{BNE}$ rationalizing the observables.

We highlight the generality of this construction along two dimensions. First, for a given realization of x, we allow for general spaces of signals T^x that may contain rich structures of non-payoff-relevant signals and allow for arbitrary correlation in players' actions. Second, the restriction $S \in \mathcal{S}$ allows the information structure S^x to vary in an unrestricted way across different realizations of x. However, $\mathcal{S}$ also embeds the important restriction that players observe at least their payoff type.

3.1.2 Assumptions on Information and Identification

Our strategy, centered on the identification of $\Theta_I^{BNE}(\mathcal{S})$, is in contrast with the prevalent approach in the literature on estimation of games, based instead on performing identification after restricting the information structure for the data generating process. This is done by choosing an information structure $S' \in \mathcal{S}$ such that the set of equilibrium predictions $Q_{\theta,S'}^{BNE}$ is analytically tractable, and by focusing on the set:

$$\Theta_I^{BNE}(S') = \left\{\theta \in \Theta : \ P_{y|x} \in Q_{\theta,S'}^{BNE}(x), \ P_x - a.s.\right\}.$$

As examples of this approach, seminal articles such as Bresnahan and Reiss (1991a), Berry (1992) and Tamer (2003) assume complete information, or $S' = \overline{S}$. Conversely, other authors such as Sweeting (2009), Bajari et al. (2010), and de Paula and Tang (2012) restrict S' to be the minimal information structure $\underline{S}$, whereby signals τ^x are uninformative.

Example 5. (Example 4 continued) Suppose we observe data generated by the one-parameter entry game with $\Delta_0 = -1/2$ and $S_0 = \overline{S}$. If we perform identification of the competitive effect Δ under the *true* restriction $S' = \overline{S}$, the probability $P_y(1,1) = 1/16$ (which corresponds to the frequency of duopolies in the data) identifies

$$\Theta_I^{BNE}\left(\overline{S}\right) = \left\{\Delta : \Pr\left\{\varepsilon_i > -\Delta\right\}^2 = P_y(1,1)\right\} = \{-1/2\}.$$

Restrictions on information can have important consequences. Ideally, the restriction imposed on the information structure S' is *true*, that is $S' = S_0$ as in Example 5. Then

$$\Theta_I^{BNE}(S') = \Theta_I^{BNE}(S_0) \neq \emptyset.$$

In typical applications there is, however, little evidence on the nature of S_0. If the restriction maintained on information is not true, or $S_0 \neq S'$, the model is *misspecified* and one of the following three scenarios occurs. First, misspecification may have benign consequences and the true parameter θ_0 may still belong to the identified set $\Theta_I^{BNE}(S')$. A second case is when misspecification of the information structure results in a nonempty identified set that however does not contain the true parameter, or $\theta_0 \notin \Theta_I^{BNE}(S')$. Finally, misspecification may result in the is *falsification* of the model, that is $\Theta_I^{BNE}(S') = \emptyset$. In the latter two of these three scenarios, misspecification of the information structure may generate inconsistent estimates.

3.2 BCE and Robust Identification

Adoption of Bayes Correlated Equilibrium as solution concept solves the problem of characterizing the identified set $\Theta_I^{BNE}(\mathcal{S})$. For this purpose we focus on BCE distributions of games with the minimal information structure $\underline{S}$. Under this information structure BCE distributions for the game $\Gamma^x(\theta, \underline{S})$ are probability measures $\nu \in \mathcal{P}_{Y,\mathcal{E}}$ over actions profiles and types that are consistent with the prior and incentive compatible in the sense of Definition 2. To each BCE distribution corresponds a prediction on observed behavior, which can be obtained as the marginal with respect to players' actions.

Definition 4. (BCE Prediction) The BCE ν induces a BCE prediction in the form of a distribution over outcomes

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

The observable implications of BCE behavior in a structure characterized by $(\theta, \underline{S})$ are described by the prediction correspondence

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

Before proceeding with the identification results, we highlight here the assumptions on equilibrium selection embedded in our approach. Assumption 2 and our definition of $Q_{\theta,S}^{BNE}(x)$ restricts the data to be generated by BNE play and an *arbitrary* equilibrium selection distribution over $BNE^x(\theta_0, S_0)$. This assumption, allowing for all distributions (i.e. convex combinations) over equilibria, results in the convexification of the set $Q_{\theta,S}^{BNE}(x)$. For the set of BCE predictions $Q_\theta^{BCE}(x)$ convexification is not needed: because the set $BCE^x(\theta)$ is convex, any convex combination of BCE predictions is also a BCE prediction. Hence, $Q_\theta^{BCE}(x)$ captures not only the equilibrium predictions generated by a unique BCE, but also the predictions corresponding to any equilibrium selection.

3.2.1 Robust Prediction

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 $Q_{\theta,S}^{BNE}$ over all $S \in \mathcal{S}$.

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 $S \in \mathcal{S}$, $Q^{BNE}_{\theta,S}(x) \subseteq Q^{BCE}_{\theta}(x)$.*

Part 1 of the lemma states that, given a BCE, we can generate corresponding BNE predictions. This is done by constructing an information structure where signals correspond to BCE mediator recommendations. Conversely, in a BNE players receive signals on their opponents' payoffs and an equilibrium is selected. Adopting the mediator metaphor of BCE, for every payoff type and signal realization, the mediator suggests play according to the BNE strategies selected by the equilibrium selection rule. Hence, we can construct a BCE where each player i receives the suggestion to play action y_i if and only if realized payoff types and the selection rule are such that (y_i, y_{-i}) is a BNE.

Example 6. (Example 4 continued) Figure 2 depicts the set of BCE outcomes for the case with $\Delta_0 = -1/2$. Panel (A) shows that BCE imposes restrictions on equilibrium behavior that are weaker than those imposed by BNE for a specific information structure: the sets of BNE predictions are all contained in the set of BCE predictions, as stated in Lemma 1. 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.]

3.2.2 Robust Identification

We are most interested in the implications of adopting BCE for identification. Under the assumption of BCE, the identified set of parameters in this class of games is:

$$\Theta^{BCE}_{I} = \left\{\theta \in \Theta : \; P_{y|x} \in Q^{BCE}_{\theta}(x) \; P_x - a.s.\right\}. \tag{3.2}$$

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

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

$$\Theta^{BCE}_{I} = \Theta^{BNE}_{I}(\mathcal{S}).$$

This implies that the identified set under BCE contains the true parameter value, $\theta_0 \in \Theta^{BCE}_{I}$.

Proof. See Appendix B. □

Proposition 1 translates the robust prediction insight, due to Bergemann and Morris (2013, 2016) and summarized in Lemma 1, into a robust identification result and is the foundation for the use of BCE for identification. Adopting BCE enables us to characterize the set of parameters consistent with equilibrium behavior and a common prior, with weak assumptions on information. We do not use BCE as an alternative equilibrium assumption on the data generating process: our Assumption 2 maintains that data are generated by BNE play. Instead, in light of Proposition 1, the BCE identified set Θ_I^{BCE} is useful to relax assumptions on information and recover the set $\Theta_I^{BNE}(\mathcal{S})$.

There is an important difference between the Lemma 1 and Proposition 1. In a robust prediction perspective, the set of BCE predictions contains predictions corresponding to *all* BNEs for games with information structures $S \in \mathcal{S}$. However, the robust identification perspective of Proposition 1 does not imply that for *all* $S \in \mathcal{S}$ there exists some $\theta \in \Theta_I^{BCE}$ that is compatible with the data and with S. Instead, the set Θ_I^{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 parameter values exist for an information structure $S' \in \mathcal{S}$, then S' is falsified.

3.3 Illustration: Assumptions on Information and BCE Identification

We consider again identification in the one-parameter entry game, where the econometrician observes data generated by the payoff structure $\Delta_0 = -1/2$, and wants to identify the competitive effect $\Delta \in [-1, 0]$. As anticipated in section 3.1.2, restrictive assumptions on information have substantial impact on identification. For illustrative purposes, we consider the non-sharp identified set:[14]

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

obtained by using only the observable probability of the outcome $(1,1)$. 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.]

Table 1 shows 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.[15] 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 competitor also enters. In the model with complete information players know that the equilibrium outcome is $(1,1)$ whenever a duopoly is realized. Hence this model predicts, for a given parameter value, the lowest

[14] We use in this example the non-sharp identified set based on one moment of P_y – as opposed to the sharp identified set $\Theta_I^{BNE}(\mathcal{S}')$ – to build intuition on the direction of the misspecification bias: for this very parsimonious parametrization, the full set of moments always falsifies misspecified models.

[15] 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 (2019).

$P_y(1,1)$ across all information structures. On the other hand, a model with some level of incomplete information predicts a higher frequency of duopolies, as players are more likely to enter because of their uncertainty on the presence of a competitor.[16]

For models with richer spaces of actions and parameters it is harder to sign the direction of the bias resulting from misspecification of the information structure. Nevertheless, the example conveys the idea that misspecification of the information structure may result in significant bias: estimation of $\Theta_I^{BCE} = \Theta_I^{BNE}(\mathcal{S})$ avoids this risk. Corresponding to the robust identification intuition in our Proposition 1, the identified set of parameters under BCE behavior Θ_I^{BCE} always contains $\Delta_0 = -1/2$. However in this example BCE identified sets are large relative to the space of parameters. This highlights a potential trade off when using weak assumptions on information: robust identified sets Θ_I^{BCE} avoid bias from misspecified assumptions on information, but may be wide. We return to the issue of the informativeness of Θ_I^{BCE}- and of what variation in the data may help to shrink the identified set - in Section 5, after having introduced in Section 4 computational and inferential tools that make our approach applicable.

4 Computation and Inference

4.1 Support Function Characterization of the Identified Set

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

We already noticed that, for every $x \in X$, the set of BCE predictions $Q_\theta^{BCE}(x)$ is convex: this property follows from the definition of BCE. Hence, we can represent $Q_\theta^{BCE}(x)$ through its support function as in Beresteanu, Molchanov and Molinari (2011).[17] Let B denote the closed unit ball centered at zero 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 $Q_\theta^{BCE}(x)$:

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

[16]Attenuation bias is induced from misspecified complete information models even if we use for identification moments other than $P_y(1,1)$. For instance, if data are generated by a model with some incomplete information and we use a complete information model to estimate payoffs, the identified value of Δ is attenuated also when using $P_y(0,1)$ or $P_y(1,0)$ for identification.

[17]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.

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

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

We have then a new characterization of the identified set:

$$\begin{aligned} \Theta_I^{BCE} &= \left\{\theta \in \Theta: \; b^T P_{y|x} \leq h\left(b; Q_{\theta}^{BCE}(x)\right) \; \forall\, b \in B, \; P_x - a.s.\right\} \\ &= \left\{\theta \in \Theta: \; \max_{b \in B} \min_{q \in Q_{\theta}^{BCE}(x)} \left[b^T P_{y|x} - b^T q\right] = 0, \; P_x - a.s.\right\}. \end{aligned} \tag{4.1}$$

The computation of this object can be further simplified: because the inner program is a linear constrained minimization, we can consider its dual maximization program. This step makes it is possible to check whether θ belongs to the identified set Θ_I^{BCE} by solving a single constrained maximization problem. Appendix A provides discusses further computational details.

4.2 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 also assume that the set of covariates X is discrete.[18] We adopt an extremum estimation approach to perform inference. We redefine the identified set characterized in (4.1) as the set of minimizers of a non-negative criterion function G,[19] or

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

where

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

[18] 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 (2019) 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).

[19] 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:

$$\tilde{G}(\theta) = \int_X \sup_{\tilde{b} \in B^{|Y|-1}} \left[\tilde{b}^T \tilde{P}_{y|x} - h\left(\tilde{b}; \tilde{Q}_{\theta}^{BCE}(x)\right)\right] \mathrm{d}P_x,$$

where $B^{|Y|-1}$ is the $(|Y|-1)$-dimensional closed ball, $\tilde{P}_{y|x}$ is defined as the first $|Y|-1$ elements of $P_{y|x}$ and $\tilde{Q}_{\theta}^{BCE}(x)$ is the set of the first $|Y|-1$ elements of BCE predictions. With an argument analogous to Theorem B.1 in the online appendix of Beresteanu, Molchanov and Molinari (2011) it is immediate to establish that $\left\{\theta \in \Theta: \tilde{G}(\theta) = 0\right\} = \{\theta \in \Theta: G(\theta) = 0\}$.

The sample analogue of the population criterion function is:

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

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

Proposition 2. (Consistent estimator) *Assume that:*

1. *The function $\theta_\pi \to \pi_i(y, \varepsilon_i; x, \theta_\pi)$ 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 function $\theta_\varepsilon \to F(\cdot; \theta_\varepsilon)$ 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. *For all $\theta \in \Theta_I$ we have $nG_n = O_p(1)$.*

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. □

Conditions 1. and 2. of the proposition are necessary to establish that the population criterion function G is lower semicontinuous over a compact domain. Condition 3. is similar to the uniform convergence conditions usually maintained in consistency proofs for estimators in point identified models (e.g. Newey and McFadden, 1994), and ensures that G_n gets uniformly (i.e. over the whole Θ) closer to its population equivalent G as n gets larger. Additionally, since by definition of identified set $G(\theta) = 0$ if and only if $\theta \in \Theta_I^{BCE}$, condition 4. serves to guarantee that G_n is close to zero for $\theta \in \Theta_I$ at a faster rate than specified in condition 3.

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:

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

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

5 Identifying Power of BCE

We address in this section the issue of the informativeness of Θ_I^{BCE}, the identified set under BCE behavior. When relaxing identifying restrictions there is, in principle, a trade-off between robustness and informativeness of identified sets. This is true not only for assumptions on information, the focus of this article, but more generally. For instance, the assumption that play is according to BNE or BCE could be weakened to non-equilibrium concepts such as k-level rationality (Aradillas-Lopez and Tamer, 2008). We consider in this example the *two-parameter entry game*, a variant of the two-player entry game described in Section 2.2. In this game there are no covariates, but competitive effects Δ_i are firm-specific, and payoff types are iid standard Normal, or $\varepsilon_i \sim N(0,1)$. In Figure 3 we explore how assumptions on equilibrium and information affect the identification of competition effects in this game. Each panel represents identified sets under different assumptions on players' behavior for data generated by Nash equilibrium play under complete information. Our method is a compromise between the goals of robustness and informativeness. On the one hand, the identified set Θ_I^{BCE} in red (Panels C, D) is much larger than the set obtained under the (correct) assumption of Nash Equilibrium with complete information (in yellow, Panel D). On the other hand, identified sets obtained under weaker assumptions on behavior, such as level-1 and level-2 rationality (Panels A, B) are hardly informative.[20]

[Figure 3 about here.]

5.1 The Role of Covariates: Point and Set Identification

The previous example shows that the assumption of Bayes Correlated Equilibrium results in tighter identification than non-equilibrium restrictions do. 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

[20]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. We also note that the identified set under rationalizability (corresponding to level-2 rationality for this model) is not a superset of Θ_I^{BCE}: although this equilibrium assumption relaxes the assumption of equilibrium play, it is only defined for a complete information environment.

the data generating process, coincides with $\overline{S}$. In turn, this may raise concerns on the informativeness of Θ_I^{BCE}.

To shrink the identified set Θ_I^{BCE} we introduce a key source of identifying power: variation in exogenous x. In particular, full-support variation of covariates player-specific covariates yields point identification of payoff parameters θ_π under the assumption of BCE, as it does under more restrictive informational assumptions. The full-support assumption ensures the existence of values of covariates for which players have a dominant strategy for almost all payoff types: identification of payoffs can then proceed as in single-agent discrete choice models. This identification strategy was first proposed by Tamer (2003) for games of complete information under the assumption of pure Nash Equilibrium play, but it still applies without restrictions on information and equilibrium selection.[21] In fact, BCE guarantees that players have equilibrium beliefs and do not play dominated actions: this is sufficient for point identification of θ_π.[22] A formal statement of this intuition in a simple setting (two-player entry games with linear index payoffs) is in Proposition 3, part 1. in Appendix B.

5.1.1 Identification with Finite Support: the Price of Robustness

Although we do not expect the large support assumptions to always hold in applications, the identification at infinity argument points to a source of variation that helps identification also in the case of covariates with finite support. To illustrate the identifying power of BCE in the latter case, we compute identified sets for a two-player entry game with linear index payoffs. We present in Table 2 projections of Θ_I^{BCE} for three different data generating processes, characterized by information structures $S_0 = \bar{S}$, $S_0 = \underline{S}$, and $S_0 = S^P$ respectively,[23] and for two examples of uniformly distributed covariates with finite support, X' and X''. The set of covariates $X' = X_1' \times X_2' \times X_C'$ is characterized by $X_i' = X_C' = \{-1, 0, 1\}$; the set $X'' = X_1'' \times X_2'' \times X_C''$ is instead characterized by player-specific $X_i'' = \{-3, 0, 3\}$ for $i = 1, 2$ and $X_C'' = X_C'$.

[Table 2 about here.]

[21] Several other articles 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.

[22] 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 equilibrium notions may suffice.

[23] 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 do not change qualitatively the informativeness of our identified sets.

Results indicate that discrete sets of covariates have 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. In particular, the projection of the identified sets along Δ_1 and Δ_2 shrinks by a factor of about 5 to 7 (depending on the assumptions on the data generating process) when the support of covariates gets larger from X' to X''.

5.2 Identification of Correlation among Payoff Types

Whereas full-support variation of covariates generates point identification of payoff parameters θ_π, we do not have a corresponding point identification result for the payoff type parameters θ_ε. Identification of θ_ε is challenging because in BCE correlation in actions may arise from information (i.e. correlated signals). For instance, if signals may induce highly correlated play despite low correlation in payoff types, it is hard to identify the bounds for the correlation in payoff types. There are however limits to the extent to which this can occur: players know their payoff type, and BCE play has to be consistent with incentive compatibility and with a common prior. BCE distributions that systematically induce duopolies that are not profitable (based on players' incorrect beliefs that the opponent would stay out) are unlikely to satisfy incentive compatibility.

In this subsection we further investigate the identifying power of BCE with respect to θ_ε. The general characterization of the sharp identified set Θ_I^{BCE} in Section 4 uses an infinite number of moment inequalities to (set) identify θ_ε so that the mapping between data and parameters is not transparent. To get concrete intuition on how the parameter θ_ε is identified, we assume that θ_π is point identified and we construct non-sharp bounds using a few moment inequalities and a simple implication of equilibrium behavior. Any BCE distribution gives zero probability to dominated actions: any positive weight on a dominated action for player i would violate her incentive compatibility constraint.

For each value of θ_ε and x this observation implies thus a lower bound $LB_y(\theta_\varepsilon; x)$ on the probability of observing any outcome $y \in Y$, constructed as the integral of the cdf of ε over the region of $\mathcal{E}$ where y is a *dominant outcome* (i.e. y_i is dominant for every player i). Similarly we construct an upper bound $UB_y(\theta_\varepsilon; x)$ for the probability of each outcome y by integrating over all areas of $\mathcal{E}$ where y is a *non dominated outcome,* that is for all players i no other $y_i' \neq y_i$ is dominant. We can then construct bounds for θ_ε. For each outcome $y \in Y$ we define a set $BD(y, x)$ which includes values of θ_ε such that $P_{y|x}$ falls within the bounds $LB_y(\theta_\varepsilon; x)$ and $UB_y(\theta_\varepsilon; x)$. Finally, the sets $BD(y) = \cap_{x \in X} BD(y, x)$ summarize the identification power of the bounds constructed using outcome y. Variation in x naturally shrinks the sets $BD(y)$.[24]

[24]The construction of the bounds outlined above is described more formally in Proposition 3, Part 2 in Appendix B for a two-player entry game with point identified payoff parameters θ_π.

In Figure 4 we show the bounds on outcome probabilities LB and UB, and the sets BD of parameters θ_ε compatible with these bounds for a two-player entry game with point identified payoffs. For this figure we assume that payoff types are jointly Normal so that we can focus on the identification of the parameter $\theta_\varepsilon = \rho$ that represents the correlation of players' payoff types. Panel (A) depicts bounds on outcome probabilities when all covariates are zero. Although the bounds are wide, encompassing a range of realizations of $P_{y|x}$, they are non-trivial. As the correlation in ρ increases, players are more likely to choose the same action: bounds on the probabilities of outcomes $(0,0)$ and $(1,1)$ increase with ρ, whereas bounds on outcomes $(0,1)$ and $(1,0)$ decrease with ρ.

[Figure 4 about here.]

Panel (B) depicts sets $\cap_{y \in Y} BD(y,x)$ of parameters ρ that generate bounds compatible with the data for a given value of x. To understand what variation in covariates is most helpful in identifying ρ we plot these sets as vertical segments for different values of x. In this example, where the upper bound of BD is sharp,[25] values of covariates that generate the largest dispersion in payoffs across players are the most informative about the lower bound of BD. This is because if the observed level of correlation in actions is high even if the deterministic part of players' payoffs is very different, then the value of correlation in payoff types cannot be too low. Symmetrically, values of covariates that generate identical payoffs are most informative about the upper bound on the correlation parameter ρ.

Panel (C) shows bounds on parameters implied by the inequalities LB and UB for different values of correlation ρ_0 in the data generating process. The upper bound on ρ in $\cap_{y \in Y} BD(y)$ is sharp (it coincides with ρ_0), but only the moment $P_{(1,1)}$ generates a non-trivial lower bound for most values of ρ_0. This is not surprising: if we observe a certain frequency of duopolies, it must be the case that correlation in payoff types is not too low. Panel (D) exemplifies how the bounds on parameters implied by the inequalities shrink as the amount of variation in x increases. As x takes values on a wider support, the moments $P_{(0,1)}$ and $P_{(1,0)}$ start being informative on the lower bound for ρ, and the set of values compatible with the inequalities becomes reasonably small.

We remark that the set $\cap_y BD(y)$ need not be a subset of the projection of Θ_I^{BCE} onto the direction of ρ: to obtain $\cap_y BD(y)$ we have assumed a point identified θ_π and we only use part of the information contained in the model, whereas Θ_I^{BCE} considers joint sharp identification of the full vector of parameters. However, the figure provides reassurance that

[25] This is a because in the DGP we chose (complete information, $S_0 = \overline{S}$) the probability of observing firms doing the same action (hence, selecting either $(0,0)$ or $(1,1)$) is the lowest across all possible S_0: every other information structure implies a higher probability of observing $(1,1)$ or $(0,0)$ for any given ρ. In turn, this means that there is no level of correlation in unobservables $\rho > \rho_0$ that is compatible with the data, since for such ρ the value of LB_y for $y \in \{(0,0),(1,1)\}$ would exceed the corresponding $P_{y|x}$.

the structure of the model - together with moments of the *joint* distribution of outcomes - have significant identifying power with respect to the parameter θ_ε that summarizes the distribution of payoff types.

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 lead to 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. According to this view, 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. 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).[27]

The empirical methods developed in the previous sections of this article are well suited for this application. The institutional features of the industry offer limited guidance on the information available to players, and firms base 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.[28] Moreover, industry players are likely to be heterogeneous in their ability to collect and process private information.[29] We also estimate the game under the assumption of complete information and minimal information, and discuss the consequences of using standard methods that impose these

[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/`.

[27] 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.

[28] 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.

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

more restrictive assumptions.

Results from our method are consistent with a substantial degree of differentiation between the grocery stores in malls and local supermarkets. In particular, although we do not reject high values (in absolute value) of competitive effects among supermarket chains, we reject high values (in absolute value) for the effect of malls on supermarkets. Adopting weak assumptions on information is key for this finding: models with minimal and complete information generate confidence sets for parameters that are not nested into those produced by the more general model. We obtain lower bounds for competitive effects that are closer to zero in the case of the complete information model, and point estimates of positive competitive effects for the minimal information model.

As a consequence, in the counterfactual analysis of Section 7 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 models with complete and minimal information predict an increase of the probability of observing two or more local stores upon removing the mall from small 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 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.

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)[30] as the relevant market for our study.

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

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.[31] The geographic extension of these markets is consistent with industry sources and previous studies.[32] 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.

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.

[31] 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.

[32] 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.

The Italian supermarket industry is subject to extensive regulation, and entry in local markets may be delayed significantly by zoning and other laws.[33] 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.[34] The dependent variable in these specifications is either the number of supermarkets in a 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 heterogeneity in the impact of large malls on the decisions of different industry groups. In addition, the counterfactual market structure that would emerge if malls were not present in some markets also depends on the competitive effect that supermarket industry groups have on each other's entry decisions: credible counterfactuals require estimates of these parameters.

6.2 Game-theoretic Model

To capture strategic interaction among players in the supermarket industry we estimate a static model of entry that is within the class described in Section 2, and similar to our illustration of Section 2.2. Each player chooses whether to be present in each local market. 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:

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

whereas payoffs from staying out of the market are normalized to zero.[35]

Market level covariates x_{im} include a measure of market size (product of population and logarithm of income), an indicator for the presence of large malls in the market, and a player-specific home-region indicator. The coefficient measuring the effect of market size

[33] 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.

[34] The ordered probit model is equivalent to the specification of Bresnahan and Reiss (1991b). It may be interpreted as a game-theoretic model with homogeneous players, complete information, and iid Normal payoff types.

[35] 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).

on profits is constant across players. The coefficients that measure the effect of malls on supermarket players, the home-region indicator and competitive effects Δ_i are instead heterogeneous across players. 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 ρ between any pair $(\varepsilon_{i,m},\varepsilon_{j,m})$.

In principle supermarket groups may choose to enter a market with several stores, or to vary store format. To reduce the complexity of the model, we assume instead that player's actions y_i are binary. Moreover, we consider a game with three players, lumping together cooperatives, independent Italian groups and French groups.[36] Hence player i (for example, independent Italian groups) can take a binary action $y_{im}\in\{0,1\}$ in market m, so that $y_{im}=1$ 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 interesting counterfactuals.[37]

We also assume that the presence of large malls is exogenous to outcomes in the supermarket industry.[38] This is a strong assumption, but not unrealistic in our data. Malls have a much larger catchment area than supermarkets, as they can attract consumers from a region that only partly overlaps with the local grocery market. Moreover, restrictive regulation and the limited availability of 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.

We estimate the model under weak assumptions on the information structure: according to Proposition 1, this is equivalent to estimating the set Θ_I^{BCE}. 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 estimates under two more restrictive assumptions: complete information (Ciliberto and Tamer, 2009), and minimal information (Su, 2014).[39]

[36] 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.

[37] Previous studies of market structure in retail industries have explored other 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 area where no pair of markets is more than a few hundred miles apart.

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

[39] In both of these case the additional restrictions imposed on estimation go beyond assumptions on information. To be in line with Ciliberto and Tamer (2009) we also restrict equilibrium selection assuming that data are generated by pure-strategy Nash equilibrium. To follow Su (2014) we restrict in an important way the payoff structure (by

Proposition 3 guides our intuition on what variation in the observables identifies the parameters. Although our model includes a firm specific covariate, the home-region indicator variable, this variable does not have full-support thus our parameters are set identified. Bounds on the β parameters are identified by covariation of observable characteristics and entry patterns. Identification of Δ_j stems 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. As discussed in Section 5.2, high correlation between entry decisions across firms in markets that have different profitability across firms (based on data and other parameters) is particularly informative about the lower bound on ρ. Similarly, low correlation between entry decisions across firms in markets that have uniform profitability across firms helps establish an upper bound on ρ.

6.3 Estimation Results

Column (I) in Table 4 presents projections of the estimated 95% confidence set for parameters in the identified set under the assumptions of BCE. We report, for each parameter of the model, the lowest and highest value it takes in the confidence set. Below the projection of the confidence set we also report parameter values that correspond to $\hat{\theta}_0$, the minimizer of the criterion function G_n. Coefficient magnitudes are not immediately interpretable in this class of models, but the counterfactuals in the next section give a sense of their quantitative impact on outcomes.

[Table 4 about here.]

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.

In column (II) 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 equilibrium and complete information. For the constant, market size parameters, and home-region

assuming that payoff types are iid) and equilibrium selection (by assuming that a *unique* BNE is played in the data generating process).

parameters the confidence sets corresponding to the weak assumptions on information and complete information models are largely similar. Assuming 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 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. Under complete information the competitive effects are mostly milder than those obtained with weak assumptions on information. This finding is in line with our discussion in Section 3.3: 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 equilibrium expectations incorporate uncertainty about competitors' actions. 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. This is intuitive: 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.

In column (III) of Table 4 we report parameter estimates obtained under the assumption of minimal information.[40] This method finds negative and precisely estimated effects of malls, and very weak or positive competitive effects. This latter finding seems inconsistent with economic intuition, suggesting that the strong restriction of minimal information - together with the assumption of iid payoff types that is necessary estimation with standard methods - does not fit the data well.[41] For this reason, we mostly focus on the models with weak assumptions on information and with complete information in the next section.

The fact that the set we estimate under restrictive assumption on information is not nested in the set estimated under the weaker BCE assumption deserves further discussion. Our robust identification result establishes that the complete information identified set is a subset of the BCE identified set. However, when going from an identification argument to finite-sample estimates, sampling variation could cause the sets to be non-nested. Additionally, misspecification driven by more restrictive assumption could lead the complete information model to be falsified by the data, and hence have an empty identified set. In this case, there would be no reason to expect estimates obtained under that assumption to

[40] We compute these estimates using the method of Su (2014); in Appendix D we discuss the details of this estimation procedure, and compare it with alternative methods such as Bajari et al. (2010).

[41] This finding is similar to Grieco (2014), who in a different empirical context finds the minimal (private) information model rejected by the data.

lie inside the robust estimated set.[42] [43]

7 Counterfactuals

We consider the counterfactual scenario in which regulation prevents the construction of large shopping malls in small markets. This counterfactual quantifies how market structure is affected by the presence of large malls. We examine in particular the eight small markets 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. In a setting where estimation yields a confidence set containing many parameter vectors, and where the model has multiple equilibria, there are several possible ways of defining and computing the counterfactuals of interest. We give a formal definition of the counterfactual objects in the next subsection.

7.1 Model Predictions and Counterfactual Objects

To quantify counterfactual changes in market structure we focus on two classes of outcomes: the probabilities of observing certain market structures and the expected number of players operating in a market. We define these outcomes as real-valued functions of equilibrium distributions ν, parameters θ and covariates x. In particular, the probability of specific market structure outcomes $\hat{Y} \subset Y$ is

$$W_{\hat{Y}}\left(\nu, \theta, x\right) = \sum_{y \in Y} \int 1\left\{y \in \hat{Y}\right\} \nu\left(y, \varepsilon\right) \mathrm{d}\varepsilon,$$

and the expected number of players is

$$W_N\left(\nu, \theta, x\right) = \sum_{y \in Y} \int \left[\sum_{n=1}^{|N|} n \cdot 1\left\{\text{number of entrants in } y = n\right\}\right] \nu\left(y, \varepsilon\right) \mathrm{d}\varepsilon.$$

More generally, functions W such as those defined above may be adapted to consider more counterfactual outcomes of interest, such as expected firm value or different measures of consumer welfare (which would require a demand system).

Several approaches are possible to summarize counterfactual predictions across the ad-

[42] A similar result is observed in Haile and Tamer (2003) and in Dickstein and Morales (2018).

[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 article. 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 equilibrium in complete information games.

missible equilibrium distributions and parameter values in the set

$$\{(\nu, \theta) : \nu \in BCE^x(\theta), \theta \in C_n\}.$$

The first, and most conservative, approach to reporting counterfactual outcomes looks at the predicted intervals whose upper (lower) bounds involve maximization (minimization) over both equilibria and parameter values in the confidence set. Such intervals are:

$$I_W^x = \left[\min_{\{(\nu,\theta):\, \nu \in BCE^x(\theta),\, \theta \in C_n\}} W(\nu, \theta, x), \max_{\{(\nu,\theta):\, \nu \in BCE^x(\theta),\, \theta \in C_n\}} W(\nu, \theta, x) \right].$$

Furthermore, when the function W describes outcomes that are desirable from the point of view of regulators, and that the counterfactual policy seeks to foster, it is interesting to focus on upper bound probabilities

$$\overline{W}(\theta, x) = \max_{\nu \in BCE^x(\theta)} W(\nu, \theta, x).$$

This is because regulators may naturally be seeking to encourage such outcomes, thus helping firms to select equilibria that maximize their probabilities. Focusing on upper bounds is also in line with important articles in the literature (Ciliberto and Tamer, 2009). Hence, we also compute changes in upper bounds of the probability of market outcomes of interest, and then average these upper bounds across markets to summarize the overall direction of change.

Our counterfactual exercise - evaluating the effect of large malls in small grocery markets - involves a change in x: we denote x^{pre} as the market-level covariates in our data, and x^{post} as the covariates in the counterfactual scenario that removes large malls. For each parameter value in the confidence set $\theta \in C_n$ we obtain the difference in average upper bounds as:

$$\Delta\overline{W}(\theta) = \left(\frac{1}{|\hat{X}|} \sum_{m \in \hat{X}} \overline{W}\left(\theta, x_m^{post}\right) - \frac{1}{|\hat{X}|} \sum_{x} \overline{W}\left(\theta, x_m^{pre}\right) \right),$$

where $\hat{X}$ denotes the set of markets affected by the counterfactual, and then report its bounds across parameter values in the confidence set:

$$I_{\overline{W}} = \left[\min_{\{\theta \in C_n\}} \Delta\overline{W}(\theta), \max_{\{\theta \in C_n\}} \Delta\overline{W}(\theta) \right].$$

Finally, we may want to separate the uncertainty in prediction due to the multiplicity of parameters in the confidence set from the uncertainty arising from the equilibrium multiplicity.[44] To do so, we fix the parameter value $\hat{\theta}_0$ that minimizes the empirical criterion function G_n, and compute intervals:

$$I_W^x\left(\hat{\theta}_0\right)=\left[\min_{\left\{\nu\in BCE^x\left(\hat{\theta}_0\right)\right\}} W\left(\nu,\hat{\theta}_0,x\right),\ \max_{\left\{\nu\in BCE^x\left(\hat{\theta}_0\right)\right\}} W\left(\nu,\hat{\theta}_0,x\right)\right].$$

This last exercise is subject to an important caveat. In our set-identified model there is no rigorous sense in which $\hat{\theta}_0$ is more likely to be the true parameter value θ_0 than any other parameter in our confidence set C_n. Hence this exercise is intended more as an illustration of the properties of the model than an empirical evaluation of the counterfactual.

7.2 Variable and Fixed Latent Information Structure

When performing estimation under weak assumptions on information we are assuming that the data are generated by BNE in the game $\Gamma^{x^{pre}}\left(\theta_0,S_0\right)$ for some unspecified $S_0\in\mathcal{S}$ (see our Assumption 2). Then, we compute counterfactual objects such as $\overline{W}\left(\theta,x^{post}\right)$ by considering all outcomes that correspond to BCE distributions $\nu\in BCE^{x^{post}}\left(\theta\right)$. In doing so, we are implicitly allowing the counterfactual outcomes to be generated by BNEs of the game $\Gamma^{x^{post}}\left(\theta,S\right)$, where the information structure can be *any* $S\in\mathcal{S}$ (see also Lemma 1). Hence, the counterfactual objects outlined in the previous subsection let the latent information structure in the counterfactual be different from S_0, the original (unknown) information structure in the data generating process. We refer thus to this approach as the *variable latent information structure* approach. Although appropriate in some contexts, the assumptions implicit in this approach may not fit well other empirical settings. For instance in our application it is plausible that removing malls does not affect the information structure of the game between supermarket chains.

This important point is originally made in Bergemann, Brooks and Morris (2019), who also suggest an alternative procedure to compute counterfactuals for a *fixed latent information structure*. In an example where the counterfactual involves changing covariates from x^{pre} to x^{post} their approach suggests that, if $\Gamma^{x^{pre}}\left(\theta_0,S_0\right)$ generated the data, the analyst should compute counterfactual outcomes compatible with BNEs of the game $\Gamma^{x^{post}}\left(\theta_0,S_0\right)$, keeping S_0 fixed. To implement this procedure, they examine BCEs of the *double game* where players choose actions y of the factual game and actions y' of the counterfactual game in a way that is compatible with common knowledge of the primitives and incentives.

[44] An alternative exercise would be to abstract from equilibrium multiplicity altogether, and use in the counterfactual the equilibrium distributions that best fit the data, as in Grieco (2014). However, in our setting the set of equilibria for the game with covariates x^{pre} need not be the same as the set of equilibria for the game with covariates x^{post}, so that it may not be possible to fix the equilibrium selection to simulate counterfactual outcomes. Instead, we compute counterfactuals imposing restrictions on how information may change across factual and counterfactual games - see Subsection 7.2.

We adapt their framework by defining the set of BCEs of the double game $B\tilde{C}E^{x^{pre},x^{post}}(\theta)$ which contains augmented equilibrium distributions $\tilde{\nu} \in \Delta(Y \times Y \times \mathcal{E})$. These equilibrium distributions, just like BCE distributions as defined in Section 3, have to be consistent with the prior and incentive compatible. The latter condition has to hold for both factual and counterfactual actions. Moreover, BCE distributions of the double game $\tilde{\nu}$ have to be consistent with the observed outcomes of the factual game.[45] We can then modify outcome functions W as $\tilde{W}(\tilde{\nu}, \theta, x^{pre}, x^{post})$, so that counterfactuals depend on augmented equilibrium distributions $\tilde{\nu}$.

For any generic value (x, x') of factual and counterfactual covariates (x^{pre}, x^{post}) intervals for counterfactual objects under fixed latent information are thus:

$$\tilde{I}_{\tilde{W}}^{x,x'} = \left[\min_{\left\{(\tilde{\nu},\theta):\, \tilde{\nu} \in B\tilde{C}E^{x,x'}(\theta),\, \theta \in C_n\right\}} \tilde{W}(\tilde{\nu}, \theta, x, x'), \max_{\left\{(\tilde{\nu},\theta):\, \tilde{\nu} \in B\tilde{C}E^{x,x'}(\theta),\, \theta \in C_n\right\}} \tilde{W}(\tilde{\nu}, \theta, x, x') \right],$$

upper bound probabilities are

$$\overline{\tilde{W}}(\theta, x, x') = \max_{\tilde{\nu} \in B\tilde{C}E^{x,x'}(\theta)} \tilde{W}(\tilde{\nu}, \theta, x, x'),$$

and counterfactual intervals for the parameter $\hat{\theta}_0$ are

$$I_{\tilde{W}}^{x,x'}\left(\hat{\theta}_0\right) = \left[\min_{\left\{\tilde{\nu} \in B\tilde{C}E^{x,x'}(\hat{\theta}_0)\right\}} \tilde{W}\left(\tilde{\nu}, \hat{\theta}_0, x, x'\right), \max_{\left\{\tilde{\nu} \in B\tilde{C}E^{x,x'}(\hat{\theta}_0)\right\}} \tilde{W}\left(\tilde{\nu}, \hat{\theta}_0, x, x'\right) \right].$$

7.3 Counterfactual Results

We report in Figure 5 the average across markets of intervals I_W^x for counterfactual probabilities of market structure outcomes. Panel (A) represents probabilities of the outcome "no entrants," and Panel (B) represents the outcome "at least two entrants." Intervals for x^{post} (without mall) are solid lines, whereas intervals for x^{pre} (with mall) are dashed lines. The two intervals in green at the top of the two panels represent intervals for the model with weak assumptions on information, whereas the two intervals in red at the bottom refer to the complete information model.

[Figure 5 about here.]

[45] See Appendix D for a more formal treatment of this approach and for computational details.

In the model with weak assumptions on information, the effect of removing malls from the markets we study is ambiguous: the average upper bound probability of observing no entrants goes up, but also the average upper bound probability of observing at least two entrants increases.[46] The data and the model, which generate a large identified set and allow for a rich set of counterfactual predictions, do not yield intervals I_W^x that give a definitive answer to our question of interest. In contrast, the complete information model generates a much sharper conclusion: the intervals I_W^x obtained under the assumption of complete information show that removing a large mall from a small market generates an overall decrease in the probability of observing no entrants, and an increase of the probability of observing at least two entrants.

We do not report intervals for fixed latent information in Figure 5, as these are similar to those constructed with variable latent information. In Figure 6 instead we turn to intervals I_W^x for a different outcome of interest: the expected number of entrants. Here, we compare directly counterfactual results obtained under variable and fixed latent information. The intervals I_W^x obtained under variable latent information (on the left) confirm the overall message of Figure 5: the model with weak assumptions on information suggests that removing malls from eight small markets would generate uncertain effects. The lower bound on the expected number of entrants decreases in all markets when removing the mall, but the upper bound increases in most of the markets we consider.

The fixed latent information approach (Bergemann, Brooks and Morris, 2019), instead, gives a more consistent picture: lower bounds on the expected number of entrants decrease when removing the mall, *and upper bounds also decrease* for a majority of the markets we consider, as well as on average. This leaves open the possibility that the opposite of what is suggested by the complete information model in Figure 5 may happen: removing the mall from a small market may generate a market structure with fewer entrants.

[Figure 6 about here.]

The figure also highlights the potential of the fixed latent information methods for sharpening counterfactual conclusions from models with weak assumptions on information. Because of the reduction in the uncertainty of counterfactual predictions, in this case the fixed latent information approach gives a sharper policy evaluation, which potentially contradicts the one obtained under a stronger assumption on information (complete information).

We turn in Table 5 to the analysis of the changes in upper bound probabilities $I_{\overline{W}}$. We also report this counterfactual object for the complete information model, for the minimal information model and for the reduced form models used in Table 3.[47]

[46] This result is not an artifact of averaging across markets, and holds true for most individual markets.

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

[Table 5 about here.]

In line with the confidence sets (where the sign of the effect of malls is not identified) and with the evidence on I^x_W in Figure 5, counterfactual predictions on the effect of removing malls as summarized by $I_{\overline{W}}$ are mostly inconclusive for the model with weak assumptions on information. This is true both when using a fixed latent information structure and when using a variable latent information structure in counterfactuals (column II).[48] We notice however that intervals for outcomes such as entry of at least two player, or entry of individual supermarket groups, mostly lie in the negatives: this means that the upper bound probabilities of these events may *decrease* upon removing malls. This finding is consistent with our discussion of the results on number of entrants in Figure 6: it is possible that a policy that removes malls would have adverse effects on market structure.

Models with more restrictive assumptions on information yield a different conclusion. The model with complete information in column (III) predicts a decrease of the probability of no entry for most parameter values, and an increase in the probabilities of having at least two players operating in a market, or of observing entry by specific players. The predictions of the game-theoretic model with minimal information (column IV) and of the reduced form model (column V) are similar to the predictions of the complete information model. The models in columns (IV) and (V) are point identified, so that they yield point predictions of $I_{\overline{W}}$. These are quite precisely estimated and in most cases close to the midpoints of the intervals produced by the complete information model.

The similarity between complete information predictions in column (III) and reduced form predictions in column (V) is not surprising: the ordered probit model we use to make counterfactual predictions on the probability of no entrants or at least two entrants is analogous to a Bresnahan and Reiss (1991b) specification with homogeneous payoffs and complete information. In this specification, payoff types are perfectly correlated across players, closely mimicking the high correlation estimated in the context of the complete information model. Whereas reduced form models can be very useful to describe correlations in the data, using them to extrapolate in counterfactual predictions entails strong (and often not obvious) assumptions. Predictions from the BCE model rely on strictly weaker assumptions.

7.4 The Empirical Content of BCE: Informativeness of Counterfactuals

A recurring theme in our counterfactual results is that the model with weak assumptions on information yields wide prediction intervals. Although this may seem undesirable, it is

[48] Counterfactual intervals obtained under fixed latent information need to be subsets of the intervals obtained under fixed latent information. However, the same inclusion relationship does need not be true of the change in upper bound probabilities that we report in this table.

also useful, as it enables the analyst to distinguish between those results that are robust and those that are driven by assumptions. Moreover, the uncertainty in the model's counterfactual predictions deserves further investigation: in principle, intervals may be wide because of the the size of the confidence set of parameters, or because of the equilibrium multiplicity. The former source of uncertainty may be addressed by better variation in the data (see also our discussion in Section 5) or by adopting less conservative inferential methods. The latter source instead, equilibrium multiplicity, is a fundamental feature of the model and is linked to our weak assumptions on information. The multiplicity of BCEs reflects in fact the many information structures that may have generated the data. Anchoring counterfactual predictions to the *unobserved* information structures that generated the data, as in the fixed latent information procedure, may help reduce the uncertainty in prediction stemming from multiple equilibria. Hence, in this subsection we investigate to what extent the uncertainty in prediction is due to equilibrium multiplicity: this ultimately clarifies the empirical content of BCE and its power to predict counterfactual outcomes in models of discrete games.

We first compare intervals I_W^x with $I_W^x\left(\hat{\theta}_0\right)$, the predictions that are obtained fixing the parameter value at $\hat{\theta}_0$, the parameter that minimizes the criterion function G_n. Fixing a parameter value in the identified set allows us to abstract from the uncertainty stemming from partial identification. Table 6 reports ratios $\frac{|I_W^x(\hat{\theta}_0)|}{|I_W^x|}$ and $\frac{|I_{\tilde{W}}^{x,x'}(\hat{\theta}_0)|}{|I_{\tilde{W}}^{x,x'}|}$ that represent the relative width of counterfactual prediction intervals computed at $\hat{\theta}_0$ with respect to general intervals (incorporating uncertainty due to $\mathcal{C}_n$) for the variable and fixed latent information approaches, respectively. The table shows that, although the ratio varies across markets and values of x, there is still considerable uncertainty in prediction in our model even when we fix payoff parameters at $\hat{\theta}_0$: counterfactual prediction intervals shrink by about 30 percent on average for the variable latent information approach, and by 40 percent on average for the fixed latent information approach. Despite the potential of sharper inference and better data to reduce the width of prediction intervals, multiple equilibria and uncertainty in prediction seem to be an unavoidable cost associated with our approach.

[Table 6 about here.]

The fixed latent information approach, however, mitigates this cost. In fact, we already found that fixing the latent information structure in the counterfactual delivers meaningful improvements in sharpness in the exercise of Figure 6. At the same time, fixed latent information does not contribute to sharpening the analysis of the change in counterfactual upper bound probabilities in Table 5. To better understand the role of fixing the latent information structure in our application, we compare directly counterfactual prediction intervals $I_W^x\left(\hat{\theta}_0\right)$ and $I_{\tilde{W}}^{x,x'}\left(\hat{\theta}_0\right)$ obtained by fixing the parameters at $\hat{\theta}_0$. We report in Table 7 the ratio $\frac{|I_{\tilde{W}}^{x,x'}(\hat{\theta}_0)|}{|I_W^x(\hat{\theta}_0)|}$ which represents how much the fixed latent information approach shrinks counterfactual intervals relative to the variable latent information approach.

[Table 7 about here.]

Overall, the fixed latent information method delivers an average reduction of 20 percent in the width of counterfactual intervals. For some outcomes the effect is even larger: for instance, when looking at counterfactual probabilities of entry of Italian groups, intervals are shrunk (on average) by 35 percent. At best, counterfactual prediction intervals shrink up to more than 60 percent. The fixed latent information approach of Bergemann, Brooks and Morris (2019) is thus a very useful tool to deliver sharper counterfactual predictions in this class of models.

Three main conclusions emerge from our counterfactual exercises. The first concerns the answer to our empirical question in this application: what is the effect of large malls on supermarkets? The question has no clear theoretical answer, and findings in previous literature are not conclusive (Grieco, 2014). Our data - read through the lens of a flexible model - do not dispel this uncertainty. However, if we maintain the assumption that the information structure stays constant in a counterfactual scenario where we remove malls from small markets, the model offers sharper conclusions in some dimensions. In particular, the range of the expected number of players in a market decreases if we remove the mall in most of the markets we consider. This suggests that the presence of malls in small markets may have positive spillovers for the supermarket industry. There is hence no clear cut case to adopt any policy that limits the presence of malls.

The second takeaway is that assumptions on information (maintained both in estimation and in counterfactual prediction) are a key primitive in empirical discrete games. The method developed in this article allows us to considerably weaken assumptions on information, thus transparently showing what counterfactual conclusions are robust and which ones are driven by strong assumptions. In our application, models with either complete information or perfectly private information suggest that removing malls has considerable scope to improve market outcomes in the supermarket industry. This conclusion however does not stand when we remove strong restrictions on information, and is - at least in part - overturned when counterfactual predictions are generated with a fixed latent information structure. Finally, we found that the robustness provided by our approach comes at the cost of uncertainty in counterfactual prediction. This uncertainty, although partly due to the lack of variation in the data of our application, is mostly an unavoidable by-product of the weak assumptions on information that we maintain.

8 Conclusion

We present in this article 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 misspecified. 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 us to avoid restrictive assumptions.

There are several avenues for future research left open by this article. First, we 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 us to identify the information structure of the game that generates the observable outcomes. Second, estimation of games under the BCE solution concept may be interesting beyond providing robustness to assumptions on information - the angle we explored in this article. Results in the theory of learning in games (see e.g. Hart and Mas-Colell, 2013) show a deep connection between regret-minimizing dynamics and correlated equilibrium. In turn, this suggests that BCE may capture well outcomes of long-run interaction in incomplete information games, thus providing a connection between dynamic play and a static solution concept. We plan to explore this in future work.

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 (2019): "Identification and Estimation of Dynamic Games When Players' Beliefs Are Not in Equilibrium," *The Review of Economic Studies*, Forthcoming.

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

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

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 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.

Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

——— (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 1–18.

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., B. A. Brooks, and S. Morris (2019): "Counterfactuals with Latent Information," *Working Paper*.

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*.

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.

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 (2018): "What do Exporters Know?" *The Quarterly Journal of Economics*, 133, 1753–1801.

Ellickson, P. B. and P. L. E. Grieco (2013): "Wal-Mart and the Geography of Grocery Retailing," *Journal of Urban Economics*, 75, 1–14.

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

Epstein, L. G., H. Kaido, and K. Seo (2016): "Robust Confidence Regions for Incomplete Models," *Econometrica*, 84, 1799–1838.

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.

Hart, S. and A. Mas-colell (2013): *Simple Adaptive Strategies: from Regret-Matching to Uncoupled Dynamics*, vol. 4, World Scientific.

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 (2019): "Confidence Intervals for Projections of Partially Identified Parameters," *Econometrica*, 87, 1397–1432.

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*.

Newey, W. K. and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, 4, 2111–2245.

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

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.

Su, C.-L. (2014): "Estimating discrete-choice games of incomplete information: Simple static examples," *Quantitative Marketing and Economics*, 12, 167–207.

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.

Syrgkanis, V., E. Tamer, and J. Ziani (2018): "Inference on Auctions with Weak Assumptions on Information," Tech. rep., Working Paper.

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*.

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 perform inference we need to compute the functions G and G_n. In this appendix we describe the steps necessary to compute these functions defined in Section 4.2. At the core of both G and G_n there is the maxmin program

$$\max_{b\in B}\min_{q\in Q_\theta^{BCE}(x)}\left[b^T P_{y|x}-b^T q\right], \qquad (P0)$$

which must be computed for every value of x.

Step 1 - Discretization: To make $(P0)$ feasible we approximate the infinite dimensional object ν by discretizing the set $\mathcal{E}=\times_i\mathcal{E}_i$. Let $\mathcal{E}^r\subset\mathcal{E}$ be the discretized set, with $|\mathcal{E}^r|=r$. We construct $\mathcal{E}^r$ as the product space of $\mathcal{E}_i^r\subset\mathcal{E}_i$, where every set $\mathcal{E}_i^r$ contains $r_i=\frac{r}{|N|}$ equally spaced quantiles of F_{ε_i}.[49] We also define $f^r(\cdot;\theta_\varepsilon)$ as the probability mass function over $\mathcal{E}^r$, where the mass of each $\varepsilon\in\mathcal{E}^r$ is generated by F_{ε_i} and a Normal copula with correlation parameter $\rho=\theta_\varepsilon$. The program $(P0)$ is then approximated by the feasible program

$$\begin{array}{rcr}
\max\limits_{b\in\mathbb{R}^{|Y|}}\ \min\limits_{q\in\mathbb{R}^{|Y|},\,\nu\in\mathbb{R}^{|Y|\times r}} & b^T\left(P_{y|x}-q\right) & (P1)\\
s.t. & b^Tb-1 & \leq 0\\
\forall y\in Y & q(y)-\sum_\varepsilon\nu(y,\varepsilon) & =0\\
\forall\varepsilon\in\mathcal{E}^r & \sum_y\nu(y,\varepsilon)-f^r(\varepsilon;\theta_\varepsilon) & =0\\
 & \sum_{y,\varepsilon}\nu(y,\varepsilon)-1 & =0\\
\forall i,y_i,y_i',\varepsilon_i & \sum_{y_{-i}}\sum_{\varepsilon_{-i}}\nu(y,\varepsilon_i,\varepsilon_{-i})\left(\pi_i(y_i',y_{-i},\varepsilon_i;x,\theta)-\pi_i(y,\varepsilon_i;x,\theta)\right) & \leq 0.
\end{array}$$

Although in $(P0)$ the minimum is taken over $q\in Q_\theta^{BCE}(x)$ only, here we minimize over both a vector of predictions $q\in\mathcal{P}_Y$ and a distribution $\nu\in\mathcal{P}_{Y\times\mathcal{E}^r}$ whose marginal on Y corresponds to q. The restriction that q must be a BCE prediction is now incorporated by imposing that ν must satisfy the constraints that characterize BCE distributions, as specified in Definition 2.

Step 2 - Vectorization: The discretized ν is a matrix with dimensions $|Y|\times r$; we define $v=\text{vec}(\nu)$, the vectorized ν that stacks the columns of ν in a vector with $d_\nu=|Y|\cdot r$ rows. We then transform $(P1)$ by defining new variables $\tilde{p}=P_{y|x}-q$ and $z=\begin{bmatrix}z_1\\z_2\end{bmatrix}=\begin{bmatrix}\tilde{p}\\v\end{bmatrix}$. As the set of predictions is a subset of the $(|Y|-1)$-dimensional simplex, we modify the objective

[49] We have experimented with other discretization techniques (e.g. Halton sets, random draws) and have found negligible impact on our results as long as $\mathcal{E}^r$ includes at least some relatively extreme (both positive and negative) payoff types. Including such values of payoff types is important because for them the incentive compatibility constraint of BCE is more likely to be binding.

of the program to $\begin{bmatrix} \tilde{b} \\ 0 \end{bmatrix}^T \left(P_{y|x} - q\right)$, where $\tilde{b}$ is a vector in the $(|Y|-1)-$dimensional closed ball. As argued in footnote 19, this modified objective yields a value of zero if and only if the original program has a value of zero. The transformed program is

$$\begin{array}{lll} \max\limits_{\tilde{b}\in\mathbb{R}^{|Y|-1}} \min\limits_{z_1\in\mathbb{R}^{|Y|}, z_2\in\mathbb{R}_+^{d_\nu}} & \begin{bmatrix} \tilde{b} \\ 0_{d_\nu+1} \end{bmatrix}^T z, & (P2) \\ s.t. & \tilde{b}^T\tilde{b} & \leq 1 \\ & A_{eq}z & = a \\ & A_{ineq}z & \leq 0_{d_{ineq}}, \end{array}$$

where A_{eq} and A_{ineq} are matrices that stack, respectively, linear equality constraints and linear inequalities. These matrices have $d_{eq} = |Y| + r + 1$ and $d_{ineq} = \sum_{i\in N}(|Y_i| \cdot |Y_i - 1| \cdot r_i)$ rows, respectively. The object a is a vector of constants, and we use 0_d, 1_d and I_d to denote, respectively, the $d-$vector of zeros and ones, and the $d \times d$ identity matrix. To construct the matrix A_{eq}, notice that the equality constraints in $(P1)$ can be written as

$$\begin{aligned} I_{|Y|}\tilde{p} + A_{eq}^1 v &= P_{y|x} \\ A_{eq}^2 v &= f^r(\theta_\varepsilon) \\ 1_{d_v}^T v &= 1, \end{aligned}$$

where A_{eq}^1 is a matrix of r copies of a $I_{|Y|}$, or $A_{eq}^1 = 1_r^T \otimes I_{|Y|}$, and A_{eq}^2 is a block-diagonal matrix with r rows and $1_{|Y|}^T$ on the diagonal, or $A_{eq}^2 = I_r \otimes 1_{|Y|}^T$. The $d_{eq} \times d_z$ matrix A_{eq} is then

$$A_{eq} = \begin{bmatrix} I_{|Y|} & A_{eq}^1 \\ 0_{(r\cdot|Y|)} & A_{eq}^2 \\ 0_{|Y|}^T & 1_{d_v}^T \end{bmatrix}$$

with $d_z = |Y| \cdot (r+1)$; a is a $d_{eq}-$vector defined as

$$a = \begin{bmatrix} P_{y|x} \\ f^r(\theta_\varepsilon) \\ 1 \end{bmatrix}.$$

The incentive compatibility inequality constraints in $(P1)$ are also linear, so that the matrix A_{ineq} can be constructed in a similar way.

Step 3 - Duality and Maximization Program: Although $(P2)$ is in the form of a maxmin problem, it can be transformed into a maximization problem by considering the dual of the inner minimization:

$$\begin{array}{lccc} \max\limits_{\tilde{b}\in\mathbb{R}^{|Y|-1},\lambda_{eq}\in\mathbb{R}^{d_{eq}},\lambda_{ineq}\in\mathbb{R}_{+}^{d_{ineq}}} & -\left[\begin{array}{c} a \\ 0_{d_{ineq}} \end{array}\right]^{T}\left[\begin{array}{c} \lambda_{eq} \\ \lambda_{ineq} \end{array}\right] & & (P3) \\ s.t. & \tilde{b}^{T}\tilde{b} & \leq 1 & \\ & \left(A^{T}\right)_{1:|Y|}\left[\begin{array}{c} \lambda_{eq} \\ \lambda_{ineq} \end{array}\right] & =-\left[\begin{array}{c} \tilde{b} \\ 0 \end{array}\right] & \\ & \left(A^{T}\right)_{|Y|+1:d_{z}}\left[\begin{array}{c} \lambda_{eq} \\ \lambda_{ineq} \end{array}\right] & \geq 0_{d_{\nu}}, & \end{array}$$

where $A=\left[\begin{array}{c} A_{eq} \\ A_{ineq} \end{array}\right]$, the vectors λ_{eq} and λ_{ineq} are the dual variables associated to the constraints of $(P2)$, $\left(A^{T}\right)_{1:|Y|}$ and $\left(A^{T}\right)_{|Y|+1:d_{z}}$ denote the first $|Y|$ and the last $r\cdot|Y|$ rows of the matrix A^{T}. By strong duality and existence of BCE, the program $(P3)$ has the same value than $(P2)$ and we compute it using the solver KNITRO in the modeling environment AMPL.

Computational Burden: - Due to the tractable nature of the the program $(P3)$, the computational burden of mapping BCE identified sets and confidence intervals is manageable. For example, computation of $G(\theta)$ for the two-player game of Table 2 with $r=50^{2}$ takes less than 30 seconds of CPU time on a 3.4Ghz processor. Computation times for the function $G_{n}(\theta)$ in our application, with $r=10^{3}$, are similar. The total time necessary to map the identified set or confidence sets depends - for a fixed dimension of the problem - on the extent to which parallelization is implemented. [50]

Computing time also depends on the dimension of the game (i.e. number of players and number of strategies) and on the discretization adopted. The dimension of the program $(P3)$ - which needs to be solved for every value of x - is determined by the number of variables

$$\begin{aligned} \left|\left(\tilde{b}^{T},\lambda_{eq}^{T},\lambda_{ineq}^{T}\right)\right| &= |Y|-1+d_{eq}+d_{ineq} \\ &= 2|Y|+r\left(1+\sum_{i\in N}(|Y_{i}|\cdot|Y_{i}-1|)\right), \end{aligned}$$

and the number of equality constraints $|Y|$, and inequality constraints $r\cdot|Y|$. The exact relation between computing time and the scalars r, $|Y|$ and $|N|$ depends on the specific computing environment, but (as it is common in empirical models of discrete games) the

[50] Although parallel computation of $G(\theta)$ for different values of θ is not natively supported by AMPL, it can be efficiently implemented using the script Parampl, available at `www.parampl.com`. We thank Arthur Olszak for kind and patient support with Parampl.

computational burden grows fast with the dimensionality of the game. 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 $\left\{P_{\tau|\varepsilon}^{x} : \varepsilon \in \mathcal{E}\right\}$[51] such that:

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

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

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:

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

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

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

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

This implies that the identified set under BCE contains the true parameter value, $\theta_0 \in \Theta_I^{BCE}$.

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

Proof. 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, $\Theta_I^{BCE} \subseteq \Theta_I^{BNE}(\mathcal{S})$. Moreover, by Assumption 2, $P_{y|x} \in Q_{\theta_0,S_0}^{BNE}(x)$ almost surely with respect to P_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 $\theta_0 \in \Theta_I^{BCE}$. □

Proposition 2. *Assume that:*

1. *The function $\theta_\pi \to \pi_i(y, \varepsilon_i; x, \theta_\pi)$ 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 function $\theta_\varepsilon \to F(\cdot;\theta_\varepsilon)$ 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. *For all $\theta \in \Theta_I$ we have $nG_n = O_p(1)$.*

Then, the set $\hat{\Theta}_I = \{\theta \in \Theta : nG_n(\theta) \le a_n\}$ 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 function $\theta_\varepsilon \to F(\cdot;\theta_\varepsilon)$ 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 can be shown by contradiction, as in Milgrom and Weber, 1985). Therefore the correspondence Q_θ^{BCE} is upper hemi-continuous.

Then, the function

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

is upper semicontinuous (Lemma 17.30 in Aliprantis and Border, 1994), for all values of x, b. It follows that the function $\theta \to -h\left(b; Q_\theta^{BCE}(x)\right)$ 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:

$$P_{y_2=1|x} = \int_{\left\{\varepsilon_2 : \varepsilon_2 \geq -x_c^T\beta^C - x_2^T\beta_2^E\right\}} \mathrm{d}F_2(\cdot;\theta_\varepsilon), \tag{8.1}$$

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 $\nu^x \in BCE^x(\theta)$:

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

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, $\nu^x(y_1 = 1 \mid y_2, \varepsilon_1) = 0$ whenever $\pi_1(1, y_2, \varepsilon_1; x, \theta_\pi) < 0$. It follows that:

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

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 (8.1) 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 $\theta_\pi=(\beta,\Delta)$ be identified. We can derive (non-sharp) bounds on the distribution of observable outcomes, and thus on the joint distribution of payoff types $F\left(\varepsilon;\theta_\varepsilon\right)$. To construct lower bounds on the probabilities of outcomes, we need to define regions where such outcomes are the product of dominant strategies. For instance, 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 $\nu^x\left([y=(1,1)]\mid\varepsilon\right)=1$ for every $\nu^x\in BCE^x(\theta)$. We can similarly define a region $\underline{\mathcal{E}}^y\left(x,\theta\right)$ for any action profile y.

For each y, we can also construct upper bounds by defining regions where for any i, no y_i is dominated. Hence, let

$$\overline{\mathcal{E}}^y\left(x,\theta\right)=\left\{\varepsilon:\max_{\nu^x\in BCE^x(\theta)}\nu^x\left(y\mid\varepsilon\right)>0\right\};$$

for the outcome $y=(1,1)$ for instance,

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

We can then construct the bounds:

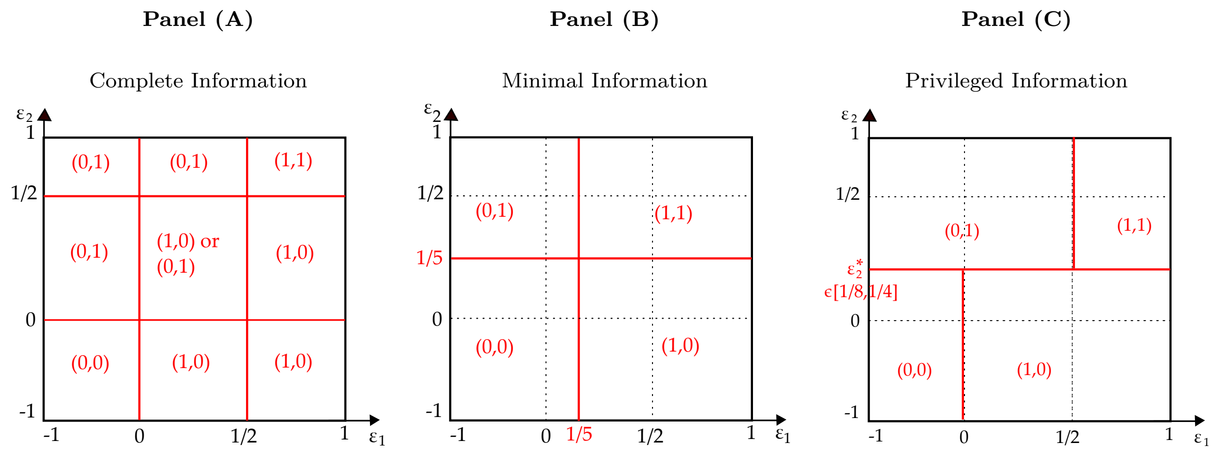
$$LB_y\left(\theta_\varepsilon;x\right)=\int_{\underline{\mathcal{E}}^y}\mathrm{d}F\left(\cdot;\theta_\varepsilon\right)\leq P_{y|x}\leq\int_{\overline{\mathcal{E}}^y(x,\theta)}\mathrm{d}F\left(\cdot;\theta_\varepsilon\right)=UB_y\left(\theta_\varepsilon;x\right),$$

and define the set

$$BD\left(y\right)=\left\{\theta_\varepsilon:LB_y\left(\theta_\varepsilon;x\right)\leq P_{y|x}\leq UB_y\left(\theta_\varepsilon;x\right),\ a.e.-x\right\}.$$

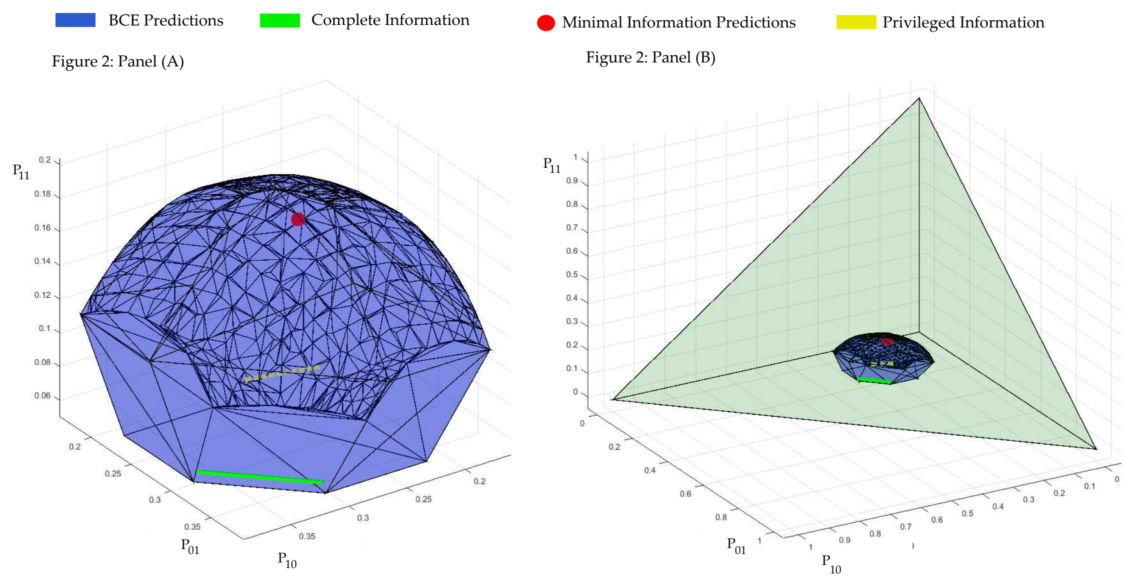
Variation in x shifts the regions $\underline{\mathcal{E}}$ and $\overline{\mathcal{E}}$, thus providing useful restrictions on θ_ε and shrinking the set of parameters $\cap_{y\in Y} BD\,(y)$ that are compatible with the bounds. $\square$

Figure 1: Information and Equilibrium Predictions



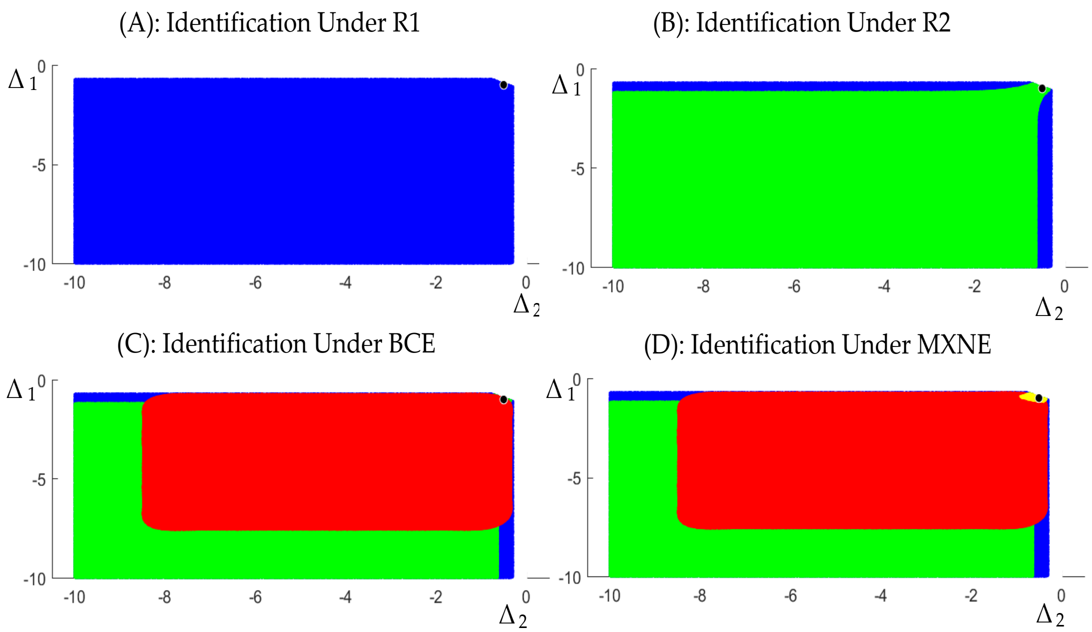
Note: - We represent BNE outcomes in the space $(\varepsilon_1, \varepsilon_2)$ for the one-parameter entry game with payoffs $\pi_i(y, \varepsilon) = y_i\left(-\frac{1}{2}y_j + \varepsilon_i\right)$ for $i = 1, 2$ and $\varepsilon_i \overset{iid}{\sim} U[-1, 1]$. (A) represents complete information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

Figure 2: BCE Predictions



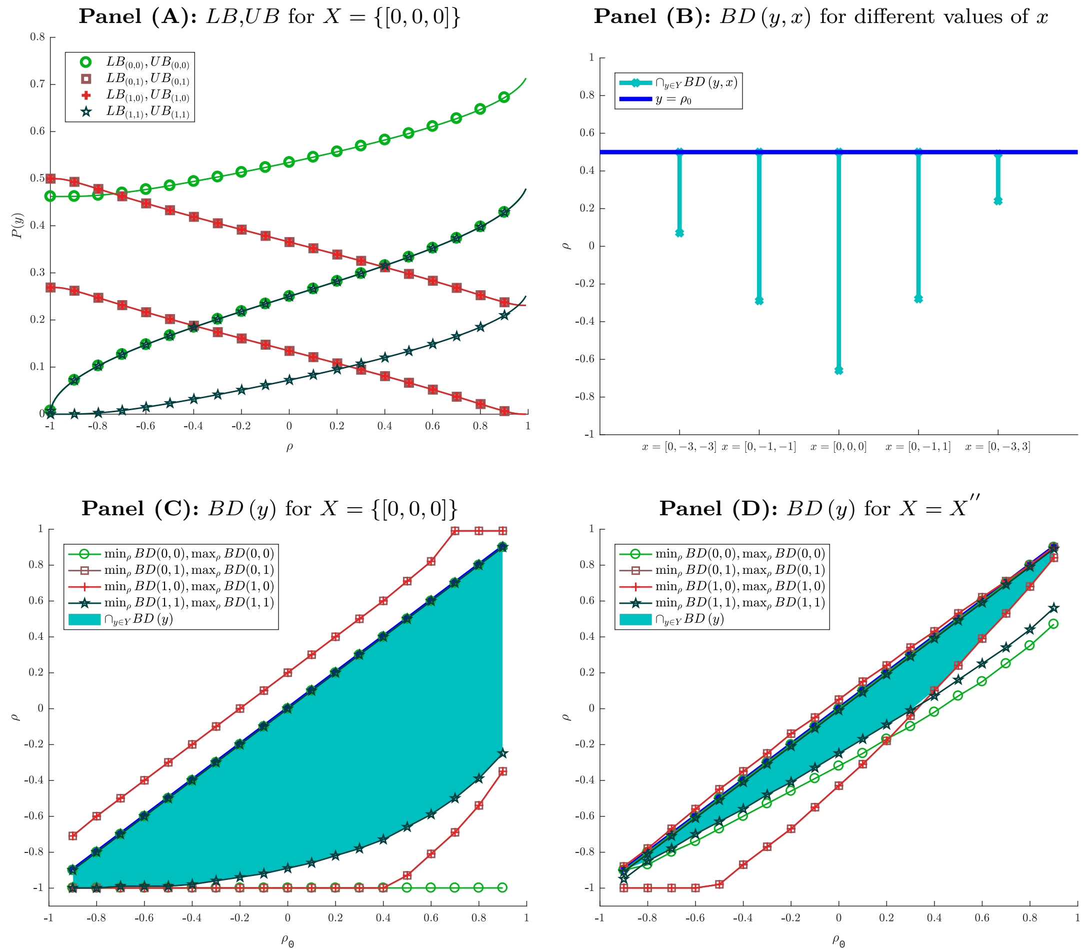
Note: - We compare BCE predictions Q_{θ}^{BCE} with the BNE predictions $Q_{\theta,S}^{BNE}$ obtained under different information structures S for the one-parameter entry game with payoffs $\pi_i(y,\varepsilon) = y_i\left(-\frac{1}{2}y_j + \varepsilon_i\right)$ for $i = 1,2$ and $\varepsilon_i \overset{iid}{\sim} U[-1,1]$. 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: Equilibrium Assumptions and Identification



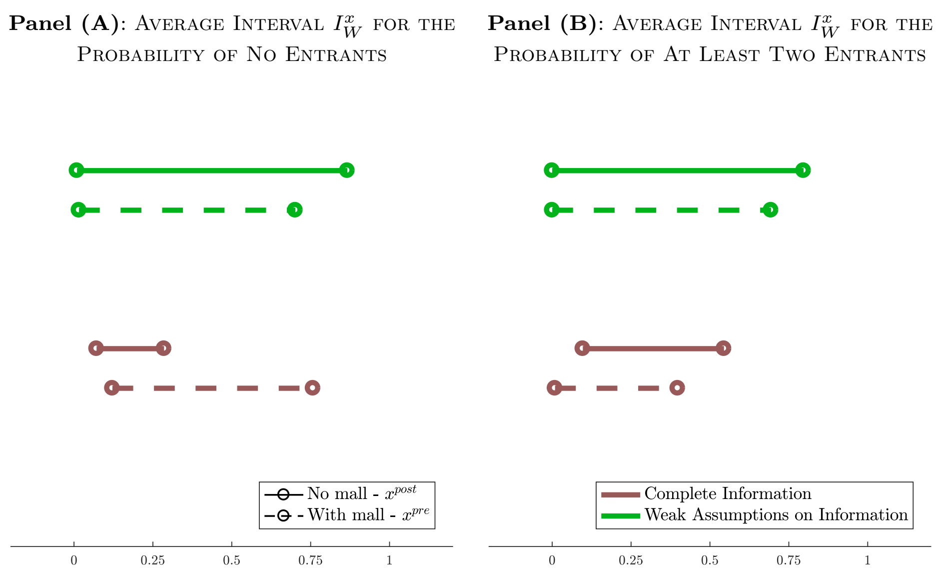
Note: - We represent the identified sets for Δ_1, Δ_2 under different restrictions on behavior in a two-player game with payoffs $\pi_i = y_i\left(y_j\Delta_i + \varepsilon_i\right)$, $\varepsilon_i \sim N\left(0,1\right)$. Data are generated by Nash Equilibrium play with complete information. The black 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 $\Theta_I^{BNE}(\bar{S})$.

Figure 4: Identification of Correlation among Payoff Types



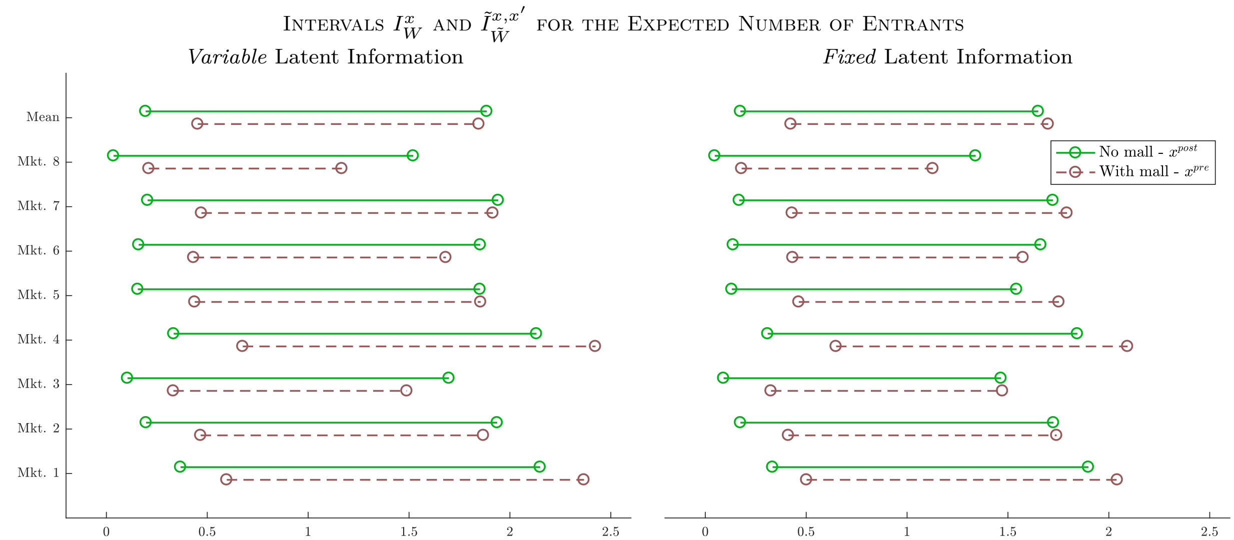
Note: - We represent bounds on probabilities of outcomes LB_y and UB_y (Panel A) and bounds $BD(y)$ on the parameter ρ (Panels B-D) for the two-player game with payoffs $\pi_i(y, \varepsilon_i; x, \theta_\pi) = y_i \left(x_c^T \beta^C + x_i^T \beta_i^E + \Delta_{-i} y_{-i} + \varepsilon_i \right)$ for $i = 1, 2$. Payoff types are distributed $\varepsilon \sim N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, and payoff parameters θ_π are the same as those in Table 2. The vector x takes values in $\{[0,0,0]\}$ for Panels (A) and (C), and in X'' in Panel (D). See Section 5.1.1 for the definition of X''. Values of x are indicated on the horizontal axis in Panel (B).

Figure 5: Probabilities of Market Structure Outcomes With and Without Malls



Note: - We represent the average across markets of counterfactual intervals I_W^x for two different outcomes of interest: the probability of observing no entry (Panel A), and the probability of observing at least two entrants (Panel B). The two green lines at the top depict intervals obtained for the model with weak assumptions on information. The two red lines at the bottom refer to the model with complete information. In each panel average intervals I_W^x are represented as solid line segments for x^{post} and as dashed lines for x^{pre}.

Figure 6: Expected Number of Entrants



Note: - We represent counterfactual intervals I_W^x (on the left) and $\tilde{I}_{\tilde{W}}^{x,x'}$ (on the right) for the expected number of entrants. Both intervals are computed for the model with weak assumptions on information. The interval I_W^x is computed with variable latent information, while $\tilde{I}_{\tilde{W}}^{x,x'}$ is computed with fixed latent information. Each figure represents intervals I_W^x and $\tilde{I}_{\tilde{W}}^{x,x'}$ as green solid line segments for x^{post} and as red dashed lines for x^{pre}. Segments for different markets and average values are stacked vertically.

Table 1: Information and Identification

True Information Structure :	$S_0 = \overline{S}$	$S_0 = S^P$	$S_0 = \underline{S}$
Projection of $\tilde{\Theta}_I^{BNE}(S')$:			
$S' = \overline{S}$	$\{-0.50\}$	$\{-0.36\}$	$\{-0.2\}$
$S' = S^P$	$[-0.82, -0.72]$	$[-0.54, -0.47]$	$[-0.29, -0.26]$
$S' = \underline{S}$	$\emptyset$	$\emptyset$	$\{-0.50\}$
Projection of Θ_I^{BCE}:	[-1,-0.47]	[-1,-0.31]	[-1,-0.33]

Note: - We report the identified sets for the one-parameter 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 $\tilde{\Theta}_I^{BNE}(\{S\})$ are obtained under restrictive assumptions on information 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 $\Delta_0 = -1/2$. For $S_0 = S^P$, we generate the data with the equilibrium corresponding to the threshold $\bar{\varepsilon}_2 = 3/16$. For further details on the computation of Θ_I^{BCE}, see Section 4 and Appendix A.

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 $\pi_i(y, \varepsilon_i; x, \theta_\pi) = y_i \left(x_c^T \beta^C + x_i^T \beta_i^E + \Delta_{-i} y_{-i} + \varepsilon_i\right)$ for $i = 1, 2$. Payoff types $\varepsilon_i \sim N(0,1)$ in (A) (B), and $\varepsilon \sim N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ 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 Appendices 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	# of Supermarkets		# of Players in Market	
	(I)	(II)	(III)	(IV)
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 (A) reports market-level descriptive statistics for the 484 markets included in our analysis. Panel (B) reports coefficient estimates and standard errors (in parenthesis) from linear regressions (columns (I) and (III)) and ordered probit models (columns (II) and (IV)). The dependent variable is the number of supermarkets of at least 1500 m^2 in column (I) and (II), or the number of supermarket players in column (III) and (IV). Market size is the product of population and log of tax income per capita. All regressions include fixed effects for 13 administrative regions. Values of R^2 refer to McFadden's pseudo-R^2 for the ordered probit regressions.

Table 4: Confidence Sets

Parameter	Weak Assumptions on Info - BCE (I)	Complete Info - Nash (II)	Minimal Info - BNE (III)
Constant	[-2.15 , -0.21] -1.46	[-3.26, -1.51] -2.08	[-3.51, -3.16] -3.32
Market Size	[3.00, 7.64] 3.66	[3.67, 6.23] 4.28	[2.59, 3.94] 3.06
Home-region:			
Cooperatives	[-0.91, 1.95] 0.61	[-0.21, 1.16] 0.64	[1.36, 1.81] 1.60
Italian Groups	[-0.39, 2.62] 1.34	[-0.14, 1.66] 0.97	[1.54, 1.99] 1.72
French Groups	[-1.46, 1.96] 1.10	[-0.50, 1.15] 0.62	[1.10, 1.58] 1.32
Presence of Large Malls:			
Cooperatives	[-3.26, 1.79] 1.35	[-2.37, 0.45] -1.19	[-2.03, -1.19] -1.61
Italian Groups	[-3.77, 1.49] 0.87	[-2.63, -0.53] -1.41	[-1.46, -0.42] -1.04
French Groups	[-2.94, 1.02] -1.04	[-4.39, -0.19] -1.31	[-1.30, -0.46] -0.80
Competitive Effects:			
Cooperatives	[-5.30, -1.11] -2.70	[-2.40, -0.73] -1.76	[-0.55, 0.47] 0.02
Italian Groups	[-6.11, -1.69] -2.46	[-2.45, -1.34] -1.84	[-1.08, -0.29] -0.66
French Groups	[-7.12, -1.55] -5.61	[-3.46, -0.39] -1.49	[1.73, 3.51] 2.88
ρ - Correlation Of Unobservable Profitability	[0.36, 0.96] 0.69	[0.90, 0.99] 0.96	–

Note: - We report estimates for the game-theoretic model of Section 6.2 under different assumptions on information. Models in columns (I) and (II) are set identified: for each individual parameter we report projections of C_n, the .95 confidence set for identified the identified set, as well as the value $\hat{\theta}_0$ that minimizes the empirical criterion function below. The model in section (III) is point identified, and we report for each parameter point estimates and 95% confidence intervals. See Appendices A and C for computational details.

Table 5: Counterfactual Change in Probability of Outcomes

Outcome	Weak Assumptions on Info - BCE		Complete Info	Minimal Info	Reduced Form
	Var. Info	Fix. Latent Info			
	(I)	(II)	(III)	(IV)	(V)
No Entry	[-0.24, 0.27]	[-0.22, 0.29]	[-0.56, 0.04]	[-0.34, -0.23] -0.29	[-0.22, -0.01] -0.12
At least 2 Players	[-0.42, 0.24]	[-0.41, 0.21]	[0.07, 0.40]	[0.20,0.27] 0.23	[0.02, 0.28] 0.15
Entry by:					
Cooperatives	[-0.40, 0.19]	[-0.39, 0.17]	[0.04, 0.70]	[0.22, 0.33] 0.28	[0.07, 0.37] 0.22
Italian Groups	[-0.60, 0.15]	[-0.59, 0.30]	[-0.17, 0.70]	[0.14, 0.29] 0.23	[-0.07, 0.28] 0.10
French Groups	[-0.56, 0.12]	[-0.55, 0.21]	[-0.16, 0.51]	[0.05, 0.14] 0.08	[-0.10, 0.12] 0.01

Note: - We report in this table counterfactual change in upper bound probabilities $I_{\overline{W}}$ of market structure outcomes for different models. Columns (I) to (III) correspond to sets $I_{\overline{W}}$ for the model with weak assumptions on information, under variable and fixed latent information, and for the complete information model, respectively. In column (IV) we report the 95% confidence interval and point estimate for the change in probability of market structure outcome for the model with minimal information. In column (V) we report 95% confidence intervals and point estimates for changes in outcome probabilities obtained from simple parametric models. To compute these we use an ordered probit to predict probabilities of no entry, Entry by at least 1 Player and Entry by at least 2 Players, and a probit specifications to predict the binary outcomes Entry by Cooperatives, Entry by Italian Groups, Entry by French Groups. We include as explanatory variables in each of these models the presence of large malls, as well as interactions and levels of region dummies, market size, and square of market size.

Table 6: Relative Size of I_W^x And $I_W^x\left(\hat{\theta}_0\right)$

	Variable latent info ratio				Fixed latent info ratio			
	No Entrants		Two+ Entrants		No Entrants		Two+ Entrants	
	x^{pre}	x^{post}	x^{pre}	x^{post}	x^{pre}	x^{post}	x^{pre}	x^{post}
Average	**0.55**	**0.74**	**0.87**	**0.71**	**0.53**	**0.68**	**0.63**	**0.57**
Mkt. 1	0.60	0.67	0.77	0.68	0.59	0.65	0.56	0.62
Mkt. 2	0.59	0.80	0.86	0.67	0.56	0.75	0.53	0.52
Mkt. 3	0.47	0.78	0.86	0.76	0.44	0.74	0.71	0.53
Mkt. 4	0.54	0.78	0.89	0.70	0.48	0.71	0.63	0.60
Mkt. 5	0.49	0.79	0.89	0.76	0.48	0.67	0.68	0.65
Mkt. 6	0.55	0.73	0.87	0.69	0.53	0.71	0.71	0.64
Mkt. 7	0.60	0.81	0.87	0.68	0.61	0.82	0.58	0.50
Mkt. 8	0.55	0.56	0.88	0.74	0.53	0.37	0.66	0.55

Note: - We report ratios of intervals $\frac{|I_W^x(\hat{\theta}_0)|}{|I_W^x|}$ and $\frac{|I_{\tilde{W}}^{x,x'}(\hat{\theta}_0)|}{|I_{\tilde{W}}^{x,x'}|}$ for two outcomes of interest: observing no entrants, and observing at least two entrants. The ratios are reported for factual (x^{pre}) and counterfactual $\left(x^{post}\right)$ values of covariates. Intervals $I_W^x\left(\hat{\theta}_0\right)$ and $I_{\tilde{W}}^{x,x'}\left(\hat{\theta}_0\right)$ are computed for the value $\hat{\theta}_0$ which minimizes the empirical criterion function.

Table 7: Relative Size of Variable and Fixed Latent Information Intervals

Panel (A):	Fixed to Variable latent info ratio $\frac{\lvert I^{x,x'}_{\tilde{W}}(\hat{\theta}_0)\rvert}{\lvert I^{x}_{W}(\hat{\theta}_0)\rvert}$							
	No Entrants		Two+ Entrants		Coop Entry		Ita Entry	
	x^{pre}	x^{post}	x^{pre}	x^{post}	x^{pre}	x^{post}	x^{pre}	x^{post}
Average	**0.95**	**0.89**	**0.64**	**0.77**	**0.90**	**0.90**	**0.76**	**0.52**
Mkt. 1	0.99	0.96	0.66	0.86	0.97	0.98	0.91	0.72
Mkt. 2	0.93	0.88	0.55	0.75	0.92	0.86	0.70	0.47
Mkt. 3	0.93	0.93	0.69	0.66	0.88	0.96	0.74	0.52
Mkt. 4	0.89	0.92	0.59	0.84	0.90	0.98	0.76	0.54
Mkt. 5	0.97	0.77	0.65	0.79	0.85	0.92	0.76	0.48
Mkt. 6	0.95	0.95	0.70	0.91	0.96	0.96	0.83	0.77
Mkt. 7	1.00	0.98	0.57	0.73	0.89	0.92	0.70	0.38
Mkt. 8	0.93	0.61	0.69	0.73	0.86	0.70	0.72	0.46

Panel (B):	Variable latent info ratio for Expected Number of entrants	
	x^{pre}	x^{post}
Average	**0.70**	**0.81**
Mkt. 1	0.78	0.91
Mkt. 2	0.60	0.82
Mkt. 3	0.71	0.70
Mkt. 4	0.60	0.85
Mkt. 5	0.77	0.83
Mkt. 6	0.70	0.88
Mkt. 7	0.62	0.86
Mkt. 8	0.79	0.69

Note: -We report ratios of intervals $\frac{\lvert I^{x,x'}_{\tilde{W}}(\hat{\theta}_0)\rvert}{\lvert I^{x}_{W}(\hat{\theta}_0)\rvert}$ for each market and on average. All intervals are computed for the value $\hat{\theta}_0$ which minimizes the empirical criterion function. Panel (A) reports ratios for four market structure outcomes, whereas Panel (B) reports ratios for the expected number of entrants.

Supplementary Materials - For Online Publication

Appendix C - Further Computational Details: Estimation

Computation of Identified Sets Θ_I^{BCE}

We describe in this appendix how compute 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, 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 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:

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

where Dir denotes the core-determining class (Galichon and Henry, 2011) and $Q_{\theta}^{MXNE}\left(x_{j}\right)$ 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 simple for games with a small number of players and actions. Then, we have:

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

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 $Q_{\theta}^{R1}\left(x\right)$ and $Q_{\theta}^{R2}\left(x\right)$ 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 $Q_{\theta}^{R1}\left(x\right)$ and of $Q_{\theta}^{R2}\left(x\right)$) does not involve numerical simulation of the values of ε.

Computation of Confidence Sets C_n for Θ_I^{BCE}

We begin by discretizing the space of covariates in three steps. First, we compute the median of market size and code a binary variable $D_m=1\left\{\text{market size}_m\geq Median\right\}$. Second, we consider the set $\tilde{M}$ of all combinations of D_m and of the other four discrete regressors in our model (home-region dummies for each player, and presence of large malls), and classify each market m as one such combination $\tilde{m}$. Out of 2^5 such bins $\tilde{m}$, 20 contain a positive number of markets m. Finally, for each $\tilde{m}$ we compute discretized values of market size as

$$\text{market size}_{\tilde{m}}=\frac{1}{|\tilde{m}|}\sum_{m\in\tilde{m}}\text{market size}_m.$$

We end up with a discretization of the market size variable with 20 distinct values. This procedure preserves the correlations of entry patterns with the exogenous variables in the data.

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 compute the 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:

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

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 $\hat{c}_0 = n g_n \times 1.25$, 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}.

Computation of Minimal Information Estimates

In the context of the model of our application in Section 6, we compute parameter estimates for a minimal information model $\hat{\theta}\left(\underline{S}\right)$. To apply standard methods in the literature we maintain the following assumptions:

1. the information structure is $S = \underline{S}$,

2. payoff types are iid type-1 EV,

3. the data are generated by an unique equilibrium σ^x for each $x \in X$.

Assumptions 2 and 3 impose strong restrictions on the payoff structure and on equilibrium selection that we do not maintain in either the general model of Section 2, or the application of Section 6 of the paper.

Under these assumptions, suppose that behavior is defined by the BNE strategy profile σ^x; for each player i, let $\tilde{\sigma}_i^x$ be the equilibrium probability of entry derived from σ_i^x. Define moreover the deterministic part of expected payoffs as:

$$\begin{aligned}\Pi_i^E(x, \theta_\pi, \sigma) &= E_\sigma \Pi_i(y_{-i}; x, \theta_\pi) \\ &= \sum_{y_{-i}} \Pi_i(y_{-i}; x, \theta_\pi)\, \tilde{\sigma}_{-i}^x(y_{-i}) \\ &= x_{im}^T \beta_i + \sum_{j \neq i} \Delta_j \tilde{\sigma}_j^x.\end{aligned}$$

The definition of BNE implies that, for all y_i such that $\tilde{\sigma}_i^x(y_i) > 0$ we have that

$$\begin{aligned}\tilde{\sigma}_i^x &= \int_{\left\{\varepsilon_i | \Pi_i^E(x, \theta_\pi, \sigma) + \varepsilon_i > 0\right\}} 1\{\varepsilon_i = e_i\}\, dF(e_i), \\ &= \Phi\left(x_{im}, \tilde{\sigma}_{-i}^x; \beta, \Delta\right).\end{aligned}$$

and using the EV distributional assumption, this becomes:

$$\tilde{\sigma}_i^x = \frac{\exp\left\{x_{im}^T \beta_i + \sum_{j \neq i} \Delta_j \tilde{\sigma}_j^x\right\}}{1 + \exp\left\{x_{im}^T \beta_i + \sum_{j \neq i} \Delta_j \tilde{\sigma}_j^x\right\}}. \tag{.2}$$

This expression motivates two estimation strategies: the first adopts a Maximum Likelihood approach, and is described in Su (2014); the second adopts instead a two-step approach, and is developed by Bajari et al. (2010). We use the former to produce the estimates used in the paper.

Maximum Likelihood Estimation - Su (2014) - We can reinterpret equation (.2) as the equilibrium map:

$$\sigma^x = \Phi(\sigma^x; x, \theta_\pi),$$

and can thus form the likelihood function of the data:

$$L(\sigma; x, y) = \sum_{i,m} \left\{y_{im} \times \log\left(\sigma_i^{x_m}\right) + (1 - y_{im}) \times \log\left(1 - \sigma_i^{x_m}\right)\right\},$$

and we adopt an MPEC approach by recovering:

$$\hat{\theta}_\pi = \arg\max_{\theta_\pi,\sigma} L(\sigma; x, y)$$
$$s.t. \qquad \sigma^x = \Phi(\sigma^x; x, \theta_\pi) \ \forall x.$$

Standard errors can be derived analytically or obtained via bootstrap. Notice that for this method we need a discrete set of covariates X; we proceed to discretize the set as we do for the estimation of confidence sets under weak assumption on information.

Two Step Estimation - Bajari et al. (2010) - Assume that there are firm-specific covariates: hence, for each player i, we have that $\Pi_i(y_{-i}; x_m, \theta_\pi) = \Pi_i(y_{-i}; x_{im}, \theta_\pi)$. Then, we can first recover estimates of (marginals of) equilibrium strategies; in our context this is equivalent to recovering how entry probabilities vary as a function of x. This can be done by estimating for each player a function $\hat{\sigma}_i(x) = \Pr(y_i = 1|x)$, for instance by fitting a linear model with OLS. In a second step, we plug the first-step estimates into equation (.2) and obtain:

$$\Pr\{y_i = 1\} = \frac{\exp\left\{x_{im}^T \beta_i + \sum_{j\neq i} \Delta_j \hat{\sigma}_j^x\right\}}{1 + \exp\left\{x_{im}^T \beta_i + \sum_{j\neq i} \Delta_j \hat{\sigma}_j^x\right\}}. \tag{.3}$$

This equation can then be estimated as a Logit model. Standard errors need to be recovered by bootstrap or with a two-step correction. As opposed to the Maximum Likelihood method, this method does not require discretization of the covariates.

Comparison of Results - In Table 1 we report estimation results for both methods. The estimates are qualitatively similar, although (bootstrap) standard errors are systematically smaller for the Maximum Likelihood method, reflecting its greater efficiency.

[Table 1 about here.]

Appendix D - Further Computational Details: Counterfactuals

Computation of Counterfactuals for the model with Weak Assumptions on Information: the Variable Latent Information Approach

All of the counterfactual objects described in Section 6.4.1 - I_W^x, $I_{\overline{W}}$ and $I_W^x\left(\hat{\theta}_0\right)$ - can be easily obtained from the computation of

$$\overline{W}(\theta, x) = \max_{\nu \in BCE^x(\theta)} W(\nu, \theta, x) \qquad (PC0)$$
$$\underline{W}(\theta, x) = -\max_{\nu \in BCE^x(\theta)} -W(\nu, \theta, x),$$

for all values of $\theta \in C_n$, where W is a function such as $W_{\hat{Y}}(\nu, \theta, x)$ or $W_N(\nu, \theta, x)$ which is *linear in* ν. For simplicity, we focus on the computation of $\overline{W}(\theta, x)$ since $\underline{W}(\theta, x)$ can be obtained with minimal changes.

With the same discretization applied in Appendix A, the program $(PC0)$ can be approximated by the feasible program

$$
\begin{array}{rcr}
\max\limits_{\nu \in \mathbb{R}^{|Y| \times r}} & W(\nu, \theta, x) & (PC1) \\
s.t. & \sum_{y,\varepsilon} \nu(y, \varepsilon) - 1 & = 0 \\
\forall \varepsilon \in \mathcal{E}^r & \sum_{y} \nu(y, \varepsilon) - f^r(\varepsilon; \theta_\varepsilon) & = 0 \\
\forall i, y_i, y_i', \varepsilon_i & \sum_{y_{-i}} \sum_{\varepsilon_{-i}} \nu(y, \varepsilon_i, \varepsilon_{-i}) \left(\pi_i(y_i', y_{-i}, \varepsilon_i; x, \theta) - \pi_i(y, \varepsilon_i; x, \theta)\right) & \leq 0.
\end{array}
$$

This is a linear programming problem which can be easily solved with commercial solvers; we compute it using the solver KNITRO in the modeling environment AMPL. Experimenting with alternative solvers (e.g. CPLEX) gave us similar results.

Computation of Counterfactuals for the model with Weak Assumptions on Information: the Fixed Latent Information Approach

We first introduce more formally the notion of the BCE for the double game, encompassing both the factual game with covariates x, and the counterfactual game with covariates x':

Definition. (BCE of the Double Game) A Bayes Correlated Equilibrium $\tilde{\nu} \in \mathcal{P}_{Y,Y,\mathcal{E},T}$ for the double game $\Gamma^{x,x'}(\theta, S)$ is a probability measure $\tilde{\nu}$ over *factual* and *counterfactual* actions profiles, payoff types, and signals that is:

1. *Consistent with the prior*: for all $\varepsilon \in \mathcal{E}$, $\tau \in T$,

$$\sum_{y,y' \in Y} \int_{[t \leq \tau]} \int_{[e \leq \varepsilon]} \tilde{\nu}(y, y', e, t)\, \mathrm{d}t\mathrm{d}e = \int_{[t \leq \tau]} \int_{[e \leq \varepsilon]} P_{\tau|\varepsilon}(t) \mathrm{d}F(e; \theta_\varepsilon)\, \mathrm{d}t;$$

2. *Incentive Compatible*: for all $i, \varepsilon_i, \tau_i, y_i, y_i'$ such that $\tilde{\nu}(y_i \mid \varepsilon_i, \tau_i, y_i') > 0$,

$$E_{\tilde{\nu}}\left[\pi_i(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi) \mid y_i, \varepsilon_i, \tau_i\right] \geq E_{\tilde{\nu}}\left[\pi_i(\tilde{y}_i, y_{-i}, \varepsilon_i; x, \theta_\pi) \mid y_i, \varepsilon_i, \tau_i\right], \qquad \forall \tilde{y}_i \in Y_i,$$

   and for all $i, \varepsilon_i, \tau_i, y_i, y_i'$ such that $\tilde{\nu}(y_i' \mid \varepsilon_i, \tau_i, y_i) > 0$,

$$E_{\tilde{\nu}}\left[\pi_i(y_i', y_{-i}', \varepsilon_i; x', \theta_\pi) \mid y_i', \varepsilon_i, \tau_i\right] \geq E_{\tilde{\nu}}\left[\pi_i(\tilde{y}_i, y_{-i}', \varepsilon_i; x', \theta_\pi) \mid y_i', \varepsilon_i, \tau_i\right], \qquad \forall \tilde{y}_i \in Y_i,$$

where the expectation operators $E_{\tilde{\nu}}\left[\cdot \mid y_i, \varepsilon_i, \tau_i\right]$ are taken with respect to the conditional equilibrium distributions $\tilde{\nu}\left(y_{-i}, \varepsilon_{-i}, \tau_{-i} \mid y_i, y_i', \varepsilon_i, \tau_i\right)$ and $\tilde{\nu}\left(y_{-i}', \varepsilon_{-i}, \tau_{-i} \mid y_i, y_i', \varepsilon_i, \tau_i\right)$, respectively in the two inequalities.

3. *Consistent with factual outcomes*: equilibrium behavior in the factual game must be consistent with factual outcomes, so that

$$\sum_{y'} \int_{\tau} \int_{\varepsilon} \tilde{\nu}\left(y, y', e, t\right) \mathrm{d}t\mathrm{d}e = P_{y|x}\left(y\right), \qquad \forall y \in Y.$$

This definition mirrors closely the one in Bergemann, Brooks and Morris (2019). Let the set $\widetilde{BCE}^{x,x'}(\theta)$ be the set of all BCE of the double game. To compute counterfactual objects $\tilde{I}_{\tilde{W}}^{x,x'}$, $\tilde{I}_{\overline{W}}$ and $\tilde{I}_{\tilde{W}}^{x,x'}\left(\hat{\theta}_0\right)$ and implement the fixed latent information approach we need to compute

$$\overline{\tilde{W}}\left(\theta, x, x'\right) = \max_{\tilde{\nu} \in \widetilde{BCE}^{x,x'}(\theta)} \tilde{W}\left(\tilde{\nu}, \theta, x, x'\right), \qquad (PC2)$$

where the function $\tilde{W}$ adapts the corresponding W in a natural way, that is:

$$\begin{array}{rcl}
\max\limits_{\tilde{\nu} \in \mathbb{R}^{|Y| \times |Y| \times r}} & \tilde{W}\left(\tilde{\nu}, \theta, x, x'\right) & (PC3) \\
s.t. & \sum_{y,y',\varepsilon} \tilde{\nu}\left(y, y', \varepsilon\right) - 1 & = 0 \\
\forall y \in Y & P_{y|x}\left(y\right) - \sum_{\varepsilon, y' \in Y} \tilde{\nu}\left(y, y', \varepsilon\right) & = 0 \\
\forall \varepsilon \in \mathcal{E}^r & \sum_{y,y'} \tilde{\nu}\left(y, y', \varepsilon\right) - f^r\left(\varepsilon; \theta_\varepsilon\right) & = 0 \\
\forall i, y_i, y_i', \tilde{y}_i, \varepsilon_i & \sum_{y_{-i}, y_{-i}'} \sum_{\varepsilon_{-i}} \tilde{\nu}\left(y, y', \varepsilon_i, \varepsilon_{-i}\right)\left(\pi_i\left(\tilde{y}_i, y_{-i}, \varepsilon_i; x, \theta\right) - \pi_i\left(y, \varepsilon_i; x, \theta\right)\right) & \leq 0 \\
\forall i, y_i, y_i', \tilde{y}_i, \varepsilon_i & \sum_{y_{-i}, y_{-i}'} \sum_{\varepsilon_{-i}} \tilde{\nu}\left(y, y', \varepsilon_i, \varepsilon_{-i}\right)\left(\pi_i\left(\tilde{y}_i, y_{-i}', \varepsilon_i; x', \theta\right) - \pi_i\left(y', \varepsilon_i; x', \theta\right)\right) & \leq 0.
\end{array}$$

To compute these counterfactuals, we need first to operationalize ($PC2$), which we do with the usual discretization of $\mathcal{E}$.

There is however another consideration when implementing this method. At the identification level, if $\theta \in \Theta_I^{BCE}$ there exists a latent information structure such that BCE predictions can match $P_{y|x}$. However, the inferential procedure that we employ implies only that for $\theta \in C_n$ BCE predictions need to match $P_{y|x}$ *approximately.* In fact, $\theta \in C_n$ only implies that θ is *close* to the minimizer of the empirical criterion function G_n built using a finite sample $\{x_i, y_i\}_{i=1}^{\infty}$.

Moreover, a feasible implementation of our empirical strategy involves several approximations: we estimate the set C_n by relying on a discretized set of covariates X, and we discretize the support of $\mathcal{E}$ in order to compute ν. Hence, finite-sample and computational error may make the restriction $marg_{\tilde{\nu}}\left(y\right) = P_{y|x}$ for some parameters $\theta \in C_n$ impossible to satisfy exactly, thus rendering ($PC3$) an unfeasible linear program. To address this problem, we compute $\tilde{P}_{y|x}(\theta)$, the distribution of the observables that best fits the data for a given parameter θ. More formally, $\tilde{P}_{y|x}(\theta)$ is equal to q that solves the program:

$$\max_{b\in B} \min_{q\in Q_{\theta}^{BCE}(x)} \left[b^T P_{y|x} - b^T q\right].$$

Intuitively, $\tilde{P}_{y|x}(\theta)$ is the distribution of the observables corresponding to the BCE that *best fits the data* $P_{y|x}$ and the parameter value θ. In our experience, these $\tilde{P}_{y|x}(\theta)$ are reasonably close to the data for parameters θ in the confidence set. We can thus compute the counterfactual quantity of interest $\overline{\tilde{W}}(\theta, x, x')$ as the solution to the program:

$$\begin{array}{rcl}
\max\limits_{\tilde{\nu}\in\mathbb{R}^{|Y|\times|Y|\times r}} & \tilde{W}(\tilde{\nu}, \theta, x, x') & (PC4) \\
s.t. & \sum_{y,y',\varepsilon} \tilde{\nu}(y, y', \varepsilon) - 1 & = 0 \\
\forall\varepsilon \in \mathcal{E}^r & \sum_{y,y'} \tilde{\nu}(y, y', \varepsilon) - f^r(\varepsilon; \theta_\varepsilon) & = 0 \\
\forall y \in Y & |\tilde{P}_{y|x}(y; \theta) - \sum_{\varepsilon, y'\in Y} \tilde{\nu}(y, y', \varepsilon)| & \leq \epsilon \\
\forall i, y_i, y_i', \tilde{y}_i, \varepsilon_i & \sum_{y_{-i}, y'_{-i}} \sum_{\varepsilon_{-i}} \tilde{\nu}(y, y', \varepsilon_i, \varepsilon_{-i}) (\pi_i(\tilde{y}_i, y_{-i}, \varepsilon_i; x, \theta) - \pi_i(y, \varepsilon_i; x, \theta)) & \leq 0 \\
\forall i, y_i, y_i', \tilde{y}_i, \varepsilon_i & \sum_{y_{-i}, y'_{-i}} \sum_{\varepsilon_{-i}} \tilde{\nu}(y, y', \varepsilon_i, \varepsilon_{-i}) (\pi_i(\tilde{y}_i, y'_{-i}, \varepsilon_i; x', \theta) - \pi_i(y', \varepsilon_i; x', \theta)) & \leq 0.
\end{array}$$

Notice that the constraint of consistency with the (pseudo-) factual outcomes is enforced with some slack, to maintain feasibility in light of the multiple approximations involved in the program. We set $\epsilon = 0.001$ in our computation; experimenting with tighter and looser tolerances did not alter the results substantially.

Computation of Counterfactuals for Models with more Restrictive Assumptions on Information

Complete Information - Under the assumption of complete information and Nash equilibrium in pure strategies, the lower and upper bound probabilities of market structure outcomes have analytical expressions for our three-player entry game (Tamer, 2003; Ciliberto and Tamer, 2009). We obtain in this way $\overline{W}^{\overline{S}}(\theta, x)$ and $\underline{W}^{\overline{S}}(\theta, x)$, and we can compute from these the complete information intervals in the bottom part of Figure 5 and the results in column (III) of Table 5.

Minimal Information - Market structure outcomes $\hat{Y}$ can be expressed as functions of a vector of strategy profiles σ as $W_{\hat{Y}}^{\underline{S}}(\sigma)$. For instance, the probability that there are no entrants is

$$W_{\{(0,0,0)\}}^{\underline{S}}(\sigma) = \Pi_{i\in I}(1 - \sigma_i).$$

Hence, upper bound probabilities for market structure outcome under the minimal information model can be obtained as

$$\overline{W}_{\hat{Y}}^{S}(x) = \arg\max_{\sigma} W_{\hat{Y}}^{S}(\sigma)$$
$$s.t. \qquad \sigma = \Phi\left(\sigma; x, \hat{\theta}_{\pi}^{S}\right),$$

where $\hat{\theta}_{\pi}^{S}$ is the parameter estimate obtained under the assumption of minimal information (see above Computation of Minimal Information Estimates) and Φ represents the equilibrium mapping in equation (.2). The average changes in upper bounds probabilities for the minimal information model reported in column (IV) of Table 5 are computed from $\overline{W}_{\hat{Y}}^{S}(x)$. Confidence intervals for the counterfactual prediction are based on 200 bootstrap samples, and account for uncertainty in $\hat{\theta}_{\pi}^{S}$.

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:

$$\begin{pmatrix} \eta_1^1 \\ \eta_2^1 \\ \eta_1^2 \\ \eta_2^2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2\rho & 0 & 0 \\ \sigma^2\rho & \sigma^2 & 0 & 0 \\ 0 & 0 & 1-\sigma^2 & 0 \\ 0 & 0 & 0 & 1-\sigma^2 \end{pmatrix}\right). \tag{.4}$$

The realizations of (η_1^1, η_2^1) are publicly observable, so that player i observes $(\eta_1^1, \eta_2^1, \eta_i^2)$. 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:

$$\varepsilon_{-i} \mid \left(\eta_i^1, \eta_{-i}^1, \eta_i^2\right) \sim N\left(\eta_{-i}^1, 1-\sigma^2\right). \tag{.5}$$

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_{-i}^1 as the signal that player i gets on ε_{-i}, and η_i^1 as what player i knows that

$-i$ knows about her payoff, so that $(\tau_i^1, \tau_i^2) = (\eta_i^1, \eta_{-i}^1)$. It follows that $(\tau_i^1, \tau_i^2) = (\tau_{-i}^2, \tau_{-i}^1)$, 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 (.4) is thus:

$$\left(\begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \tau_1^1 \\ \tau_1^2 \end{array}\right) \sim N\left(\left(\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array}\right), \left(\begin{array}{cccc} 1 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \\ \sigma^2\rho & 1 & \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 & \sigma^2 & \sigma^2\rho \\ \sigma^2 & \sigma^2\rho & \sigma^2\rho & \sigma^2 \end{array}\right)\right). \tag{.6}$$

Notice that (.6) implies that the belief of player i about ε_{-i} conditional on her information set is:

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

which is identical to the belief (.5).

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 $\pi_i(y, \varepsilon_i; x, \theta_\pi)$ 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 $\text{Sel}(BCE_\theta)$ 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:

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

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

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

Then the conditional Aumann expectation of this random set is:

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

Notice however that:

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

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.

Appendix G - A More General Model

The model in Section 2 of the paper embeds an important restriction on information: our definition of the class of information structures $\mathcal{S}$ maintains the assumption that players know the realization of market-level covariates x and their own payoff type ε_i. This restriction in turn is important for the definition of $\Theta_I^{BNE}(\mathcal{S})$ and the equivalence result in Proposition 1. In this appendix we discuss identification under more general assumptions.

We assume that players first receive a private random signal $\tilde{\tau}_i^x$ that is part of their *baseline information structure* $\tilde{S}^x$ defined as:

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

This definition of *baseline* information structure allows for both non-informative signals and perfectly informative signals on ε. The model in section 2 specifies this baseline to be perfectly informative about ε_i; we consider here cases where the baseline could be either more or less informative.

In addition to the baseline signal, every player receives an *extra* private random signal τ_i^x, which may also be informative about the full vector of types $\tilde{\tau}$ and the full vector of ε. An information structure S^x specifies, for a game with covariates x, the set of extra signals a player may receive and the probability of receiving them, given the realization of the vector of payoff types and baseline signals. Formally:

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

Whereas $\tilde{S}$ describes the baseline information, the information structure S denotes the extra information players might receive. The game $\Gamma^x(\theta, \tilde{S}, S)$ is then analogous to the game $\Gamma^x(\theta, S)$ defined in the main text, except that players observe both baseline signals according to $\tilde{S}$ and extra signals according to S. We use $\mathcal{S}(\tilde{S})$ to denote the class of admissible extra information structures S when the baseline information structure is $\tilde{S}$.

We also redefine the BNE concept used in the paper:

Definition. (Bayes Nash Equilibrium) A strategy profile $\sigma = \times_{i\in N}\sigma_i,\ \sigma_i : \tilde{T}_i^x \times T_i^x \to \mathcal{P}_{Y_i}$ is a Bayes Nash Equilibrium (BNE) of the game $\Gamma^x(\theta, \tilde{S}, S)$ if for every $i \in N$, $\tilde{\tau}_i \in \tilde{T}_i^x$ and $\tau_i \in T_i^x$ we have that, for every $y_i \in Y_i$ such that $\sigma_i(y_i \mid \tilde{\tau}_i, \tau_i) > 0$:

$$E_\sigma\left[\pi_i\left(y_i, y_{-i}, \varepsilon_i; x, \theta_\pi\right)|\tilde{\tau}_i, \tau_i\right] \geq E_\sigma\left[\pi_i\left(y_i', y_{-i}, \varepsilon_i; x, \theta_\pi\right)|\tilde{\tau}_i, \tau_i\right], \quad \forall\, y_i' \in Y_i.$$

Based on this modified notion of BNE, which defines equilibrium strategies as functions of both baseline signals $\tilde{\tau}$ and extra signals τ, it's immediate to redefine the set of BNE predictions $Q^{BNE}_{\theta,\tilde{S},S}(x)$ for the game $\Gamma^x(\theta, \tilde{S}, S)$ and the BNE identified set $\Theta_I^{BNE}\left(\mathcal{S}(\tilde{S})\right)$.

Our main result relies on the following Lemma, a re-statement of Lemma 1 in the article:

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

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

We may now extend our Proposition 1 to this environment:

Proposition. (Robust identification) *Let Assumptions 1 and 2, aptly modified for the game* $\Gamma^x(\theta, \tilde{S}, S)$*, hold. Then* $\Theta_I^{BCE}(\tilde{S}) = \Theta_I^{BNE}\left(\mathcal{S}(\tilde{S})\right)$*. This implies that the identified set under BCE behavior contains the true parameter value,* $\theta_0 \in \Theta_I^{BCE}(\tilde{S})$*.*

For $\tilde{S} = \underline{S}$, this Proposition coincides with our Proposition 1 in Section 3 of the article. However, the explicit reference to the baseline information structure allows us to generalize the result to environments where players do not observe their own ε_i, or observe not only ε_i but also other components of the vector of payoff types ε.

It's interesting to investigate how identified sets vary for different assumptions on $\tilde{S}$. The set $\Theta_I^{BCE}\left(\mathcal{S}(\tilde{S})\right)$ is certainly not invariant to $\tilde{S}$; intuitively, as information structures get more informative, the set of BCE predictions gets smaller, and hence fewer parameters are compatible with the observables, so that the identified set shrinks. This notion of "more informative" can be made precise (as in Bergemann and Morris, 2016).

Having established that $\Theta_I^{BCE}\left(\tilde{S}\right)$ is not invariant to $\tilde{S}$, we still maintain that using $\tilde{S} = \underline{S}$ is the best compromise given the goal of the paper. Baseline information structures $\tilde{S}$ that are less informative than $\underline{S}$ are likely to result in limited identifying power, whereas baselines that are more informative than $\tilde{S}$ are unlikely to be well justified in most empirical applications.

References

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010): "Estimating static models of strategic interactions," *Journal of Business & Economic Statistics*, 28, 469–482.

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

Bergemann, D., B. A. Brooks, and S. Morris (2019): "Counterfactuals with Latent Information," *Working Paper*.

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

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.

Su, C.-L. (2014): "Estimating discrete-choice games of incomplete information: Simple static examples," *Quantitative Marketing and Economics*, 12, 167–207.

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

Table 1: Minimal Information Estimation - Two Methods

Parameter	Minimal Info ML - Su (2014)	Minimal Info Two Step - Bajari et al. (2010)
Constant	-3.32 [-3.51, -3.16]	-2.89 [-3.30, -2.49]
Market Size	3.06 [2.59, 3.94]	2.62 [1.63, 3.62]
Home-region:		
Cooperatives	1.60 [1.36, 1.81]	1.95 [1.48, 2.42]
Italian Groups	1.72 [1.54, 1.99]	1.60 [1.03, 2.18]
French Groups	1.32 [1.10, 1.58]	1.66 [1.13, 2.19]
Presence of Large Malls:		
Cooperatives	-1.61 [-2.03, -1.19]	-0.96 [-1.93, 0.011]
Italian Groups	-1.04 [-1.46, -0.42]	-0.50 [-1.54, 0.53]
French Groups	-0.80 [-1.30, -0.46]	-0.28 [-1.15, 0.59]
Competitive Effects:		
Cooperatives	0.02 [-0.55, 0.47]	-0.069 [-1.22, 1.08]
Italian Groups	-0.66 [-1.08, -0.29]	-1.57 [-2.62, -0.52]
French Groups	2.88 [1.73, 3.51]	2.93 [0.69, 5.17]

Note: - We report estimates for the game-theoretic model in Section 6, obtained using the two methods for the estimation of minimal information games described in this appendix. Bootstrap standard errors are in parenthesis, and are calculated from 200 bootstrap samples.