ISD as a Basis for Set Identification in Strategically Monotonic Supermodular Games*

Francisco Garrido†

PRELIMINARY AND INCOMPLETE
DO NOT CIRCULATE OR CITE

February 6, 2023

Abstract

Using an Iterated Strict Dominance (ISD) argument, I build bounds on the (distribution of) outcomes of games and show that they pin down an identified set of the parameters of interest. These *ISD bounds* are robust to multiple equilibria in pure and mixed strategies, and apply to games of complete or incomplete information, with discrete, continuous, or discrete-continuous actions of any dimensionality. Furthermore, ISD bounds can account for unobserved heterogeneity and arbitrary informational structures (e.g., complete or incomplete information). To maximize the "bite" of the ISD bounds, I introduce Strategically Monotonic Supermodular Games, i.e., games of strategic complements/substitutes where players' payoffs are supermodular in their actions. I show that ISD rules out large swaths of the strategy set for this type of game via an easy-to-compute sequence of best-response iterations. Finally, in an application to the airline industry, I show that the identified set is economically informative about the parameters of interest.

Keywords: Partial identification, probability bounds, supermodularity, strategic substitutes, strategic complements.

JEL codes: L13, D43, C72.

*For useful comments and suggestions, I am indebted to Felipe Brugués, Adrian Rubli, Cristian Sánchez, José Tudón, Nathan Miller, Frank Verboven, as well as seminar participants at UT Austin, and the Latin American Meeting of the Econometric Society (Lima, 2022). I gratefully acknowledge financial support from Asociación Mexicana de Cultura, A.C.

†Department of Business Administration, ITAM, Río Hondo 1, Ciudad de México 01080, México. Email: fagarrido@gmail.com. Phone: + 52 55 5628 4000

1 Introduction

Industrial organization economists and policymakers are often interested in counterfactual questions that require structural models. Will a merger induce entry? Does entry affect quality decisions? Does competition affect technology adoption? However, any minimally realistic model that tries to answer one of these questions suffers from incompleteness in the sense of Tamer (2003). In a nutshell, due to equilibrium multiplicity, models do not generate a well-defined distribution over outcomes, rendering standard approaches to estimation, like maximum likelihood or GMM, infeasible.

In this paper I use an Iterated Strategic Dominance (ISD) argument to build bounds on the (distribution of) outcomes of games, and use them to pin down an identified set of the parameters of interest. These bounds (hereinafter, ISD bounds) are extremely general in that they are robust to multiple equilibria in pure and mixed strategies, can easily accommodate discrete, continuous, and discrete-continuous strategies of any dimensionality, and can account for arbitrary informational structures (i.e., complete information and incomplete information).

To understand how ISD bounds work, consider a complete information game where each player, f, independently chooses an action y_f. Also, suppose the game is indexed by an unobserved (to the econometrician) vector ξ, e.g., a player-specific payoff shifter. If for some ξ outcome $y = (y_f)_f$ *uniquely* survives ISD, then y is the *unique* Nash equilibrium of the ξ-game, and therefore it must be observed. Hence, aggregating over ξ: $Pr(y \text{ uniquely survives ISD}) \leq Pr(y \text{ is observed})$. Similarly, if for some ξ outcome y is observed, then y is one of possibly many Nash equilibria, and therefore it must survive ISD. Again, aggregating over ξ: $Pr(y \text{ is observed}) \leq Pr(y \text{ survives ISD})$.

To maximize the bite of ISD bounds I introduce Strategically Monotonic Supermodular Games (SMSGs), i.e., games where player's payoff are supermodular on their own actions, and best responses exhibit strategic complementarity/substitutability.[1] As I argue in the paper, SMSGs are a natural match for ISD bounds as in these games ISD is informative, i.e., it eliminates large swaths of the strategy set, and it is easy to compute through best response iterations.

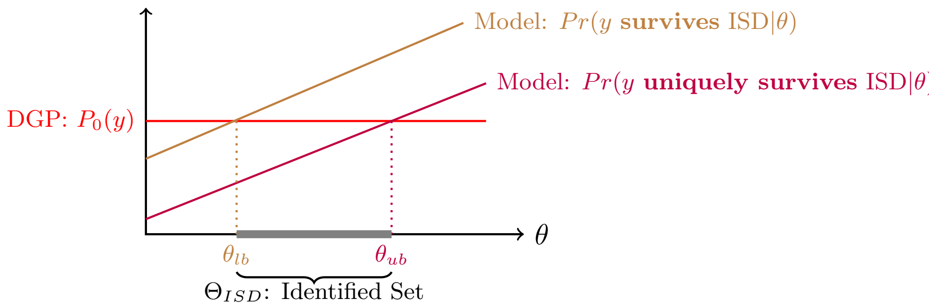
To show that ISD has bite on SMSGs, I generalize a classic result from the literature of supermodular games proposed by Milgrom and Roberts (1990), and generalized to games of incomplete information by Van Zandt and Vives (2007). In particular, while the standard results assume strategic complementarity, this paper allows for *strategic monotonicity*, i.e., for any pair of players, actions may exhibit either strategic complementarity or substitutability. In consequence, ISD bounds apply to many important strategic environments that are not covered by the standard theory of supermodular games, such as Cournot games, entry games, and capacity investment games, all of which exhibit strategic substitutability.[2]

ISD bounds generate a family of identified set which differ on their level of aggregation. In particular, going back to the notation from above, I propose an identified set based on bounds on the distribution of y, an identified set based on bounds on the distribution of *subsets* of y, and an

[1] For any two firms, f and f', f's payoffs exhibit increasing differences or decreasing differences on $(y_f, y_{f'})$.

[2] I should note that the comparative statics results from Milgrom and Roberts (1990) do not generalize to the case of strategic monotonicity. These results, however, are irrelevant to the problem at hand.

Figure 1: Identified Set Example



Note: $Pr(y$ **survives** $ISD|\theta)$ and $Pr(y$ **uniquely survives** $ISD|\theta)$ are, respectively, the probability outcome y survives ISD, and the probability that outcome y uniquely survives ISD, according to the model, and given θ. These represent upper and lower bounds on the frequency with which the model predicts y given θ. If $Pr(y$ **uniquely survives** $ISD|\theta) > P_0(y)$, then y is observed less often than the lower bound predicted by the model given θ, so this cannot be the real θ. Similarly, if $P_0(y) > Pr(y$ **uniquely survives** $ISD|\theta)$ then y is observed more often than the upper bound predicted by the model given θ, so this cannot be the real θ.

identified set based on bounds on y itself. As a general rule, the identified sets built upon more aggregated ISD bounds (e.g., bounds on the level of y) are larger than identified sets built upon less aggregated ISD bounds (e.g., bounds on the distribution of y). This disadvantage, however, is compensated by the fact that more aggregated bounds impose lower computing demands.

The intuition behind the identified sets lies in the idea that there is a data generating process (DGP), $P_0(y)$, which corresponds to the "real world" probability of observing outcome y. Any model parameter, θ, such that the model predicts $Pr(y \text{ uniquely survives ISD}|\theta) > P_0(y)$ cannot be part of the DGP, as this would imply that the model predicts y with higher probability than the one with which it is observed in the real world. Similarly, any θ that generates $Pr(y \text{ survives ISD}|\theta) < P_0(y)$ implies that y is observed more frequently than the model can generate it. As a result, such θ cannot be part of the DGP. Figure 1, depicts this intuition.

To assess the performance of ISD bounds I perform several Montecarlo experiments on a standard entry game. First, I focus on the complete information case and compare ISD bounds to the bounds in by Ciliberto and Tamer (2009) (CT), and the ones proposed by Fan and Yang (2022) (FY), respectively. As expected CT bounds, which are built on the notion of Nash Equilibria, provide the tightest bounds, followed by outcome-ISD bounds, firm-ISD bounds, and FY bounds, which are based on the idea of level-1 rationality. Later, I perform Montecarlo experiments on a game of incomplete information and explore the performance of ISD bounds for this type of game. Interestingly, the ISD bounds are informative about the relative weight of unobserved heterogeneity and incomplete information.

The flexibility in informational structures is a major point of difference between this paper and previous papers in the literature. Previous research has proposed identified sets which are *robust* to any informational assumption, but not *informative* about it, e.g. Magnolfi and Roncoroni (2020). In other words, in these papers all informational structures generate the same identified set. In the present paper, in contrast, different informational structures imply different ISD bounds and therefore different identified sets. As a result, by applying ISD bounds researchers are able to learn about the underlying information or explore the identification power of different informational assumptions.

The focus on SMSGs has the additional benefit that these games provide a natural solution to the problem of performing counterfactual experiments in the presence of multiple equilibria.[3] In SMSGs, the set of strategies that survives ISD, and therefore, the set of equilibria, is pinned down by a largest and a smallest strategy. When conducting counterfactual experiments, then, rather than focusing on point predictions, researchers can report a range of predictions, i.e., the range of strategies that survives ISD in the counterfactual game.[4] Going back to the notation above, ISD pins down extreme outcomes in SMSGs, $y^L \leq y^H$, such that a strategy profile y survives ISD if and only if $y^L \leq y \leq y^H$. As a result, y^L and y^H bind the possible outcomes generated by the game. Furthermore, these extreme strategies also bound any monotonic function $W(y)$, i.e., if $W()$ is increasing, then for any equilibrium y, $W(y^L) \leq W(y) \leq W(y^H)$.

Finally, to study the identification power of ISD bounds I provide an empirical application to the airline industry, in the spirit of CT. Namely, I estimate an entry game in which carriers simultaneously decide whether to enter a market (airport pair) or not. As opposed to CT, my empirical model admits complete and incomplete information specifications, which allows me to explore the identifying power of these information assumptions.

The rest of the paper is organized as follows. Section 1.1 reviews the relevant literature and places the current paper within it. Section 2 presents the basic model, introduces SMSGs, and shows that ISD has bite in SMSGs. Section 3 derives ISD bounds and a family of ISD identified sets. Section 4 analyses the performance of ISD bounds through Montecarlo exercises, comparing it to other bounds proposed in the literature. Section 5 presents the airline data and model to be estimated, while section 6 presents the estimation results. Finally, section 7 provides some concluding remarks.

1.1 Literature Review

This paper is related to two strands that have mostly run on independent lanes. The literature on estimation of discrete games of *complete information*, and the literature focusing on estimation of discrete games of *incomplete information*.

1.1.1 Complete Information Games

The issue of model incompleteness as described by Tamer (2003), has been a common thread throughout the literature studying estimation/identification of discrete games of complete information. The early examples in this literature, such as Bresnahan and Reiss (1991), Berry (1992), and Mazzeo (2002), bypassed the problem of equilibrium multiplicity by making strong homogeneity assumptions

[3] This issue has is recognized as a major open problem in the literature. See Aradillas-López (2020).
[4] Reguant (2016) suggests a similar approach.

on firms' payoffs that guaranteed that all equilibria could be mapped into a single outcome (e.g., number of firms), for which the model makes a unique prediction.

Later papers dealt with this issue using two broad approaches. The first approach consists of completing the model with an equilibrium selection mechanism and either assuming that it is known (e.g., Jia (2008), Li et al. (2018)) or estimating it from the data (e.g., Bajari et al. (2010)). This strategy is attractive because it brings us back to the world where standard estimation techniques work and point identification holds. The problem, however, is that economic theory provides little guidance when it comes to equilibrium selection, making any assumption related to the equilibrium selection mechanism hard to justify.

The second approach, the one that this paper takes, gives up on point identification and rather focuses on identifying a set for the parameters of interest. This approach was pioneered by Tamer (2003) and Ciliberto and Tamer (2009) (CT), who build the identified set by putting bounds on the probability of observing an outcome. In particular, the probability of observing an outcome y must be higher than the probability that said outcome is the *unique* Nash equilibria, and lower than the probability that it is *a* Nash equilibria. In this strand, Fan and Yang (2020) (FY) propose building the identified set using one round ISD, and Aradillas-Lopez and Tamer (2008) study the identification power of rationalizability as a solution concept.

Aradillas-Lopez and Tamer (2008) is perhaps the closest paper to the present one. This paper studies identification of k-level rationality in 2×2 games of complete and incomplete information, while imposing no assumptions on player's beliefs (beyond common priors and what is implied by k-level rationality). The present paper can be seen as a generalization of these ideas to more much flexible settings.

More generally, the present paper contributes to this literature by proposing ISD based bounds that apply to games of complete and incomplete information, with discrete and/or continuous actions of any dimensionality. In this dimension, ISD bounds are more general than CT and FY bounds, which are customized for discrete games. Additionally, as opposed to CT bounds, ISD bounds have do not require one to solve the complete model to compute them. This makes inference easier, specially for games with very many players or large strategy sets, allowing one to estimate a broader class of games. Finally, as compared to the FY bounds, the ISD bounds are built on a stronger concept so they provide a (weakly) tighter identification set and should be preferred whenever ISD has bite.

The present paper is also similar to Aradillas-Lopez (2011) and Aradillas-López and Rosen (2022) in using shape restrictions on payoffs, and restrictions on the action set to pin down an identified set of the parameters of interest. In particular, by restricting their attention to ordered actions, and making appropriate concavity and increasingness assumptions, they are able to pin down an identified set based on Nash equilibrium conditions. The present paper, in contrast, makes much weaker assumptions on the game's structure, i.e., supermodularity in own actions and strategic monotonicity, which allow for non-ordered strategies of any dimensionality. This, however, comes at the cost of pinning down a wider identified set.

As mentioned above, I argue that ISD bounds are particularly useful in estimating SMSGs. Many of the static games estimated in the literature are instance of SMSGs, and therefore can be estimated using the method I advance here. For example, the models in Bresnahan and Reiss (1990), who estimate entry game for isolated retail and professional markets, and Berry (1992), Tamer (2003), Ciliberto and Tamer (2009) all of whom estimate entry games for the airline industry, are all instances of SMSGs with strategic substitutes. More recently, Wollmann (2018) estimates a two-stage model for the truck industry in which players can choose which truck varieties to offer and compete in prices. Although this model cannot be shown to be supermodular, as the payoffs depend on the reduced form variable profits in the pricing stage, economic intuition strongly suggest that strategic substitution should hold (i.e., the profit gain from introducing a variety is decreasing on the varieties of my competitors). Furthermore, supermodularity can be verified numerically from the pricing stage estimates.

A number of empirical papers explicitly exploit the theory of supermodular game to solve (and estimate) models with large strategy sets that would be computationally infeasible otherwise. Most prominently, Jia (2008) estimates an entry model for Wal-Mart and Kmart with spill over effects across markets. To solve this model, she shows that the duopolistic game can be written as a supermodular game, and proceeds with estimation assuming a known equilibrium selection mechanism. This trick, however, applies only to two-player games, so her methodology does not generalize to games with three or more players. Other empirical papers that exploit supermodularity are Uetake and Watanabe (2020) who study entry and merger decisions in a supermodular matching model, and Ackerberg and Gowrisankaran (2006) who study study technology adoption with network externalizes. Both in the banking industry.

In all these papers the underlying model can be thought of as an SMSG, and therefore can be estimated using the approach I outline here. Furthermore, the approach I outline makes it feasible to relax some strong assumptions these papers made on equilibrium selection or information structures.

The idea of exploiting supermodularity to estimate empirical models is not new to this paper. Molinari and Rosen (2008) and Uetake and Watanabe (2013) both proposed using the theory of supermodular games for set identification. The current paper, however, represents a major step forward with respect to these two papers in at least two dimensions. First, I show that an ISD argument in general, not just applied to supermodular games, generates a family of bounds which differ in their identification power. Second, and most importantly, while these papers constrain their focus to games of strategic complementarity, I am able to consider the much broader class of games of strategic monotonicity, expanding the applicability of this method to a much broader class of games. This generalization is particularly important since strategic substitutability is likely more common than complementarity in empirical research.

1.1.2 Incomplete Information Games

As opposed to the complete information case, the literature on estimation of discrete games of incomplete information has, until recently, largely ignored the problem of model incompleteness in estimation. The reason for this asymmetry is that in games of incomplete information, from the perspective of each player and the econometrician, actions are probabilistic. As a result, by estimating the conditional choice probabilities, for any player f, the econometrician learns the distribution over f's competitors actions that f is facing, and can use this to estimate f's payoffs as a single agent problem using the methods developed by Hotz and Miller (1993) and Aguirregabiria and Mira (2002) for single agent dynamic settings.

This approach, which is widely used in the literature (e.g. Seim (2006), Draganska et al. (2009), Atal et al. (2022)), rests on the assumption that all the data available comes from the same equilibrium, and that there is no unobserved heterogeneity. However, de Paula and Tang (2012), for static games, and Otsu et al. (2016) and Otsu and Pesendorfer (2022), for dynamic environments, propose tests for this assumptions and find that, in commonly used datasets, the assumption fails.

The problem of model incompleteness in games of incomplete information is an area of active research. Two prominent efforts to deal with this issue are Aguirregabiria and Mira (2019), who study the problem of (point) identification in games with incomplete information and unobserved heterogeneity while estimating an equilibrium selection mechanism, and Otsu and Pesendorfer (2022) who treat equilibrium multiplicity as a market specific correlated latent variable. As in this paper, they provide results for set identification. As compared to these papers, the current paper deals with the problem of equilibrium multiplicity in a more tractable way, by imposing bounds on (the distribution of) outcomes, and making fewer assumptions on the distribution of private shocks.

On the informational structure point, the ISD bounds proposed in this paper require no assumptions on the informational structure of the game, being able to accommodate games of complete information, games of correlated private information, or games where one (or more) party is better informed than others. For example, one party may have full information, while others may only observe their private shocks. In this sense, this paper joins Magnolfi and Roncoroni (2020), in relaxing the informational assumptions required for identification.

1.1.3 Revealed preferences

A third popular route to estimation in discrete games was proposed by Pakes et al. (2015). Their approach is based on the idea that, if the data are generated by a Nash Equilibrium, then unilateral deviations from the observed actions should be unprofitable for the deviating firm. This reasoning generates profit inequalities that lend themselves for set identification, as any parameter vector that violates these inequalities cannot have generated the data.

Although the profit inequality approach has gained traction in the empirical literature due to its relative simplicity and tractability (e.g., Ellickson et al. (2013) and Wollmann (2018)), it suffers from a number of drawbacks that can make it a less than ideal candidate. First, it assumes that observed outcomes are produced by equilibrium behavior, which may be amount to a strong assumption in games with large and complex strategy sets. Second, it leaves the informational structure of the game largely unspecified, which makes it difficult to justify these when performing counterfactual experiments. Relative to this approach, ISD bounds do not suffer from either of this problems which but may be harder to implement.

2 The Model, SMSGs, and ISD

In this section I provide the building blocks of a Bayesian game and introduce a Strategically Monotonic Supermodular Games (SMSGs). Then I show that ISD has bite in SMSGs.

2.1 Model Set-Up

Consider a finite set of players (firms), $\mathcal{F}$, indexed by f, who simultaneously choose a vector, y_f, from a compact action set $\mathcal{Y}_f \subseteq \mathbb{R}^{\dim(\mathcal{Y}_f)}$, after receiving a private signal/shock, $\epsilon_f \in \mathcal{E}_f \subseteq \mathbb{R}^{\dim(\mathcal{E}_f)}$. Letting $\epsilon_{-f} = (\epsilon_{f'})_{f' \neq f}$, ex-post profits are:

$$\pi_f(y_f, y_{-f}, \epsilon_f, \epsilon_{-f}; x, \theta, \xi)$$

where, as is standard, $y_{-f} = (y_{f'})_{f' \neq f}$ is a vector containing f's competitors' actions, and where the vector of private shocks, $\epsilon = (\epsilon_f)_f$ follows a distribution $G(\epsilon|x, \theta, \xi)$, which is common knowledge.

Each tuple (x, θ, ξ) indexes a different realization of the game, which I refer to as the (x, θ, ξ)-game. Here, $x \in \mathcal{X} \subseteq \mathbb{R}^{\text{dims}(\mathcal{X})}$ represents a vector of observables, $\theta \in \Theta \subseteq \mathbb{R}^{\text{dims}(\Theta)}$ is the vector of parameters of interest, and $\xi \in \Xi \subseteq \mathbb{R}^{\text{dims}(\Xi)}$ is a vector of common knowledge variables, unobservable to the econometrician. For brevity, in what follows I omit dependence of profits, equilibrium strategies, and other variables on (x, θ, ξ) unless doing so is likely to result in confusion.

Given an (x, θ, ξ)-game, a strategy for player f is any function $\sigma_f \in \Sigma_f$ mapping f's private information, ϵ_f, to an action y_f, where Σ_f represents the set of strategies of f.[5] A strategy profile is a collection of strategies, one for each player: $\sigma = (\sigma_f)_f \in \Sigma \equiv \times_f \Sigma_f$. Given any strategy for f's competitors, $\sigma_{-f} = (\sigma_{f'})_{f' \neq f}$, f's interim payoff is:

$$\Pi_f(y_f, \sigma_{-f}, \epsilon_f) = \int_{\mathcal{E}_{-f}} \pi_f(y_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) dG(\epsilon_{-f}|\epsilon_f) \tag{1}$$

where $G(\epsilon_{-f}|\epsilon_f)$ is the conditional distribution of ϵ_{-f}.[6]

I follow Van Zandt and Vives (2007) and use interim (rather than ex ante) payoffs to define an equilibrium. In particular, a Bayes Nash Equilibrium (BNE) for the (x, θ, ξ)-game corresponds to

[5] For simplicity, throughout the paper I stick to pure strategies. It is straightforward to extend results to the mixed strategy case.

[6] A standard assumption in theoretical work, to guarantee that Π_f is well defined, is for π_f to be bounded. Although empirical models routinely violate this assumption by imposing, say, extreme value distributed error terms, the primitives of the game are sufficiently well behaved to guarantee the Π_f is always well defined. For the purpose of this paper, I will simply assume that Π_f exists for any σ_{-f}.

a strategy profile (σ_f, σ_{-f}) such that:

$$\Pi_f(\sigma_f(\epsilon_f), \sigma_{-f}, \epsilon_f) \geq \Pi_f(\sigma'_f(\epsilon_f), \sigma_{-f}, \epsilon_f), \forall \epsilon_f \in \mathcal{E}_f, \forall \sigma'_f \in \Sigma_f, \forall f \in \mathcal{F} \tag{2}$$

Furthermore, the set of BNEs is:

$$\mathcal{B} = \{\sigma \in \Sigma : \sigma \text{ satisfies (2)}\} \tag{3}$$

It is well known that even in simple settings the model above is incomplete in the sense of Tamer (2003), i.e., there exist (x, θ, ξ) for which the (x, θ, ξ)-game has a non-singleton equilibrium set $\mathcal{B}$, therefore the model does not yield a well defined prediction. I complete the model with an equilibrium selection mechanism, ρ,[7] i.e., a function that selects a strategy profile from the set of BNE, and use $\sigma^\rho \in \mathcal{B}$ to represent the strategy selected by the equilibrium selection mechanism ρ.

Before introducing the SMSGs, a quick comment regarding the informational structure of the model is in order. The model allows for an arbitrary informational structure through the private shocks/signals, ϵ, and their distribution, G. The *complete information* case, for example, can be represented by a degenerate distribution G. In this case, the randomness of the outcomes is driven by the randomness (from the perspective of the econometrician) of the common knowledge unobservable, ξ.

Other informational structures can be represented by letting $\epsilon_f = (\nu_f, \tau_f)$ where ν_f is the payoff relevant shock and τ_f is a, payoff irrelevant, signal about other player's private information, as in Magnolfi and Roncoroni (2020). For example, the *independent private information* case corresponds to $\nu_f \perp \nu_{-f}$ and $\tau_f = \varnothing$ for all f. The *privileged information* case, where one player is perfectly informed and the rest only observe their private shocks, can be represented by $\nu_f \perp \nu_{-f}$ and $\tau_f = \varnothing$ for all f except the privileged party whose signal is $\tau_f = \nu_{-f}$. Similarly, the case with *independent partially observed information* corresponds to the case where $\nu_f \perp \nu_{-f}$ and $\tau_f = \nu_{-f} + \varsigma_{-f}$, where ς_{-f} is noise. In this dimension, the present paper can easily accommodate many more informational structures than previous research has allowed for.[8]

2.2 Strategically Monotonic Supermodular Games and ISD

Here I introduce a class of games which I call *Strategically Monotonic Supermodular Games* (SMSGs) and show that for this type of games ISD is informative, in that it rules out large swaths of the strategy set, and practical, in that it is easy to compute. As a result, an estimation approach based on ISD is particularly promising for SMSGs.

The main result of this section, Theorem 1, says that in SMSGs there exist strategies, $\sigma^{i,L}$ and $\sigma^{i,H}$, such that any strategy σ that survives i rounds of ISD, for $i = 0, \ldots, \infty$, lies between $\sigma^{i,L}$ and

[7] The Equilibrium Selection Mechanism may depend on (x, θ, ξ) as well as other additional unobservables, i.e., sun spots. I omit this dependence here for simplicity, and because it plays no role in the estimation procedure I put forward.

[8] Other papers that allows for flexible information structures are Magnolfi and Roncoroni (2020), Aradillas-Lopez (2010), and Aradillas-Lopez and Tamer (2008).

$\sigma^{i,H}$ in the sense that $\sigma^{i,L}(\epsilon) \leq \sigma(\epsilon) \leq \sigma^{i,H}(\epsilon)$ for all ϵ, where "$\leq$" represents the standard vector inequality. This result is the main building block for the ISD bounds from in Section 3. To move in this direction, let us begin by investing in some definitions.

Definition 1 (Increasing Differences and Decreasing Difference)**.** *Let $h(z_1, z_2)$ be a function mapping from $\mathcal{Z}_1 \times \mathcal{Z}_2$ to $\mathbb{R}$, where $\mathcal{Z}_j \subseteq \mathbb{R}^{dim(\mathcal{Z}_j)}$ for $j = 1, 2$.*

1.a. ***Increasing Differences (ID):*** *h has increasing differences in (z_1, z_2) if, for any distinct $z_1' \geq z_1$, and distinct $z_2' \geq z_2$:*

$$h(z_1', z_2') - h(z_1, z_2') > h(z_1', z_2) - h(z_1, z_2)$$

1.b. ***Decreasing Differences (DD):*** *h has decreasing differences in (z_1, z_2) if, for any distinct $z_1' \geq z_1$, and distinct $z_2' \geq z_2$:*

$$h(z_1', z_2') - h(z_1, z_2') < h(z_1', z_2) - h(z_1, z_2)$$

Definition 2 (Complements and Substitutes)**.** *Pick an arbitrary (x, θ, ξ)-game, and let $y_{-\{f,f'\}} = (y_t)_{t \neq f,f'}$. Define (omitting dependence on (x, θ, ξ) for brevity):*

2.a. ***Complements:*** *f' is f's complement if $\pi_f(y_f, y_{f'}, y_{-\{f,f'\}}, \epsilon)$ has ID in $(y_f, y_{f'})$ for all $(y_{-\{f,f'\}}, \epsilon)$. The set of f's complements is denoted by $C(f)$.*

2.b. ***Substitutes:*** *f' is f's substitute if $\pi_f(y_f, y_{f'}, y_{-\{f,f'\}}, \epsilon)$ has DD in $(y_f, y_{f'})$ for all $(y_{-\{f,f'\}}, \epsilon)$. The set of f's substitutes is denoted by $S(f)$.*

In Definition 1, ID and DD are notions of complementarity and substitutability, respectively. Intuitively, ID implies that the marginal return of y_f is increasing in $y_{f'}$, hence the optimal y_f is increasing in $y_{f'}$. Many games exhibit ID, such as games with complementary investments. Similarly, DD implies that the marginal return of y_f is decreasing in $y_{f'}$, so the optimal y_f is decreasing in $y_{f'}$. In IO settings, DD is more common than ID. Games of entry, capacity investment, and Cournot competition, for example, typically exhibit DD.

In Definition 2, a complement (substitute) of firm f is a firm, f', whose actions are strategic complements (substitutes) to f's actions. Note that if f' is f's complement, this does not imply that π_f is in increasing in $y_{f'}$,[9] nor does it imply that f is f''s complement (i.e., the complement relation is not necessarily symmetric). Similarly, if f' is f's substitute, this does not imply that π_f is decreasing in $y_{f'}$,[10] nor does it imply that f is f''s substitute (i.e., the substitute relation is not necessarily symmetric).

[9] Say f and f' produce differentiated goods, engage in Bertrand competition, and have to decide whether to adopt a cost-saving technology or not. If f' adopts the technology it makes f worse off (f is harmed by the lower cost of f'). Nevertheless, f' adopting the technology may increase f's incentive to adopt, so that adoption decisions are strategic complements.

[10] For example, in a public good financing game, π_f might be increasing in $y_{f'}$ (the more f' invests in the public good the higher the benefit for f), and $(y_f, y_{f'})$ may be strategic substitutes, i.e., the more f' invests in the public good the lower the marginal return for f to do so.

Before moving to the definition of SMSGs, let us define the concept of a lattice, which is central to the theory of supermodular games which I exploit in this paper.

Definition 3 ((Complete) Lattice)**.** *A set $\mathcal{Z}$ together with a partial order, $\leq$, constitute a lattice if for any $z, z' \in \mathcal{Z}$, $\sup\{z, z'\} \in \mathcal{Z}$ and $\inf\{z, z'\} \in \mathcal{Z}$. Furthermore, the tuple $(\mathcal{Z}, \leq)$ is a* complete lattice *if for every $Z \subseteq \mathcal{Z}$, $\inf\{Z\} \in \mathcal{Z}$ and $\sup\{Z\} \in \mathcal{Z}$.*

Definition 4 (SMSG)**.** *The (x, θ, ξ)-game is a Strategically Monotonic Supermodular Game if (omitting dependence on (x, θ, ξ) for brevity):*

4.a. **Complete Lattice Action Set:** *The action set, $\mathcal{Y}_f \subseteq \mathbb{R}^{\dim(\mathcal{Y}_f)}$, together with the standard vector inequality, "$\geq$", conform a complete lattice for all $f \in \mathcal{F}$.*[11] *Furthermore, $\mathcal{Y}_f$ is compact for all $f \in \mathcal{F}$.*

4.b. **Order Upper Semi-Continuity** *The profit function, π_f, is order upper semi-continuous in y_f. Formally, for any totally ordered set $O \subset \mathcal{Y}_f$:*[12]

$$\limsup_{y_f \in O, y_f \downarrow \inf(O)} \pi_f(y_f, y_{-f}, \epsilon_f, \epsilon_{-f}) \leq \pi_f(\inf(O), y_{-f}, \epsilon_f, \epsilon_{-f})$$
$$\limsup_{y_f \in O, y_f \uparrow \sup(O)} \pi_f(y_f, y_{-f}, \epsilon_f, \epsilon_{-f}) \leq \pi_f(\sup(O), y_{-f}, \epsilon_f, \epsilon_{-f})$$

for all $y_{-f} \in \mathcal{Y}_{-f}$, all $f \in \mathcal{F}$, and all $\epsilon \in \mathcal{E}$.

4.c. **Supermodularity:** *The profit function, π_f, is supermodular in y_f, i.e., for any $y_f, y'_f \in \mathcal{Y}_f$:*

$$\pi_f(\sup\{y_f, y'_f\}, y_{-f}, \epsilon) + \pi_f(\inf\{y_f, y'_f\}, y_{-f}, \epsilon) \geq \pi_f(y_f, y_{-f}, \epsilon) + \pi_f(y'_f, y_{-f}, \epsilon)$$

for all $y_{-f} \in \mathcal{Y}_{-f}$, all $f \in \mathcal{F}$, and all $\epsilon \in \mathcal{E}$.

4.d. **Pairwise Strategic Monotonicity:** *For all $f, f' \in \mathcal{F}$, either f' is f's complement, i.e., $f \in C(f)$. or f' is f's substitute, $f' \in S(f)$.*

As I argue below, SMSG's have properties that make them particularly good candidates for estimation using ISD bounds. Point 4.a. of the definition is necessary to exploit the supermodular games infrastructure advanced by Milgrom and Roberts (1990) for games of complete information, and Van Zandt and Vives (2007) for games of incomplete information. Although most empirical studies satisfy this assumption, it is easy to construct games in which it is violated. For example, consider an entry game with location choice as in Seim (2006). Firms have to choose between not entering a market, entering in location A, or entering in location B. Letting 1 (0) represent the case where f does (does not) enter a given location, the strategy set is $\mathcal{Y}_f = \{(0,0), (0,1), (1,0)\}$, and it is easy to see that $\sup\{(0,1), (1,0)\} = (1,1) \notin \mathcal{Y}_f$. Point 4.b. is a technical condition

[11] Note that this definition allows $\mathcal{Y}_f$ to include $\{-\infty, +\infty\}$. Naturally, for this to work payoffs need to be well defined at infinity.

[12] A totally ordered (sub)set $C \subseteq \mathcal{Y}_f$ is a subset of $\mathcal{Y}_f$ such that for any $y_f, y'_f \in \mathcal{C}$ either $y_f \geq y'_f$ or $y_f \leq y'_f$.

necessary to guarantee that f's problem has a solution. Order upper semi-continuity is satisfied if π_f is continuous or if the strategy set is discrete.

In point 4.c. of the definition, supermodularity of π_f represents a notion complementarity between the elements of y_f. If y_f is uni-variate then this assumption is trivially satisfied. Otherwise, supermodularity is likely satisfied in cases where there are positive spill over effects between the different elements of y_f. Jia (2008) provides a prominent example of an empirical game exhibiting supermodularity. In her model, opening a Wal-Mart store in any location increases the profitability of opening a store in neighboring locations due to economies of scope in inventory management. Finally, point 4.d. says that there is *Strategic Monotonicity* meaning that each of f's competitors is either f's substitute or f's complement. For $f' \in C(f)$ this implies that f's optimal behavior is increasing in $y_{f'}$, whereas for $f' \in S(f)$ this implies that f's optimal behavior is decreasing in $y_{f'}$. Either way, the pairwise strategic relation is monotonic.

Before moving to Theorem 2 it is worth formalizing the notion of strict dominance I use throughout the paper.

Definition 5. *Strategy σ_f strictly dominates strategy σ'_f in the (x, θ, ξ)-game if:*

$$\Pi_f(\sigma_f(\epsilon_f), \sigma_{-f}, \epsilon_f) \geq \Pi_f(\sigma'_f(\epsilon_f), \sigma_{-f}, \epsilon_f), \forall \epsilon_f \in \mathcal{E}_f, \forall \sigma_{-f} \in \Sigma_{-f}$$

with strict inequality for at least one ϵ_f, where Π_f is the interim payoff defined in (1)*.*

Defining strict dominance in terms of interim payoffs, rather than ex-ante payoffs, has the advantage that it allows us to distinguish between strategies that are ex-ante equally attractive. To see this, consider two strategies, σ_f and σ'_f, equal everywhere except for a zero-measure subset of $\mathcal{E}_f$, in which σ_f is preferred to σ'_f. Ex-ante, these two strategies would be evaluated as equally good, however an interim evaluation will say σ_f is preferred to σ'_f because there are values of ϵ_f for which, σ_f fares strictly better, even if this contingencies have zero probability.

Theorem 1. *Let the (x, θ, ξ)-game be an SMSG, and let Σ^i_{ISD} denote the set of strategies that survive i ISD rounds. Furthermore, let $\sigma \leq \sigma'$ if and only if $\sigma(\epsilon) \leq \sigma'(\epsilon)$ for all ϵ. The following holds (omitting dependence on (x, θ, ξ) for brevity):*

1.a For all $i = 0, 1, 2, \ldots$, there exists $\sigma^{i,L}, \sigma^{i,H} \in \Sigma$ such that $\sigma^{i,L} \leq \sigma^{i,H}$, and such that the set of strategies that survive i rounds of ISD is:

$$\Sigma^i_{ISD} = \{\sigma \in \Sigma : \sigma^{i,L} \leq \sigma \leq \sigma^{i,H}\}$$

1.b Both $\sigma^{i,L}$ and $\sigma^{i,H}$ result from a sequence of best response iterations.

1.c As $i \to \infty$, $(\sigma^{i,L}, \sigma^{i,L}) \to (\sigma^L, \sigma^H)$, with $\sigma^L \leq \sigma^H$.

Proof. See Appendix A. □

Theorem 1 is a generalization of Theorem 5 in Milgrom and Roberts (1990), which assumes increasing differences (i.e., strategic complementarity). The present generalization from strategic complementarity to strategic monotonicity is crucial to the practical relevance of the approach to estimation I propose in this paper, as it implies that ISD bounds can be applied to a much broader class of games than the classical theory of Supermodular Games considers. Namely, ISD bounds can be applied to games of strategic substitutability which are likely the norm in industrial organization.

To get an intuition of how the proof operates, consider an entry game between competing firms, so that profits exhibit decreasing differences in entry decisions, and let $y_f = 1$ denote entry, and $y_f = 0$ denote no-entry. The worst case scenario for firm f occurs when all other players choose an "always enter" strategy, i.e., $\sigma_{f'}(\epsilon_{f'}) = 1$ for all $\epsilon_{f'}$ and $f' \neq f$. Let $\sigma_f^{1,L}$ be f's best response to this strategy profile. Decreasing differences implies that even if f's competitors choose less aggressive strategies, $\sigma_f^{1,L}$ will still be preferred to $\sigma_f < \sigma_f^{1,L}$. Hence, $\sigma_f^{1,L}$ strictly dominates $\sigma_f < \sigma_f^{1,L}$. Since this holds for every f, all $\sigma < \sigma^{1,L}$ are strictly dominated.

Similarly, the best case scenario for firm f occurs when every competitor chooses a "never enter" strategy, i.e., $\sigma_{f'}(\epsilon_{f'}) = 0$ for all $\epsilon_{f'}$ and $f' \neq f$. An analogue argument shows that the best response to this strategy, $\sigma_f^{1,H}$, strictly dominates all $\sigma_f > \sigma_f^{1,H}$. Furthermore, because this is true for every f, then all $\sigma > \sigma^{1,H}$ are strictly dominated by $\sigma^{1,H}$.

Finally, the sequence of sets that survive i rounds of ISD, i.e., Σ^i_{ISD} results from letting $\sigma^{1,L}$ and $\sigma^{1,H}$ become the new best and worst case scenarios, and iterating over best responses as described above. In appendix A.2 I show how to build and apply this sequence for the case of pure complements, i.e., $C(f) = \mathcal{F} \setminus \{f\}$ for all f; the case of pure substitutes, i.e., $S(f) = \mathcal{F} \setminus \{f\}$ for all f; and the general case.

2.3 Two Entry-Game Examples

Here I show the implications of Theorem 1 to two archetypal entry games. The independent private information case, and the complete information case.

2.3.1 Independent Private Information Entry Game

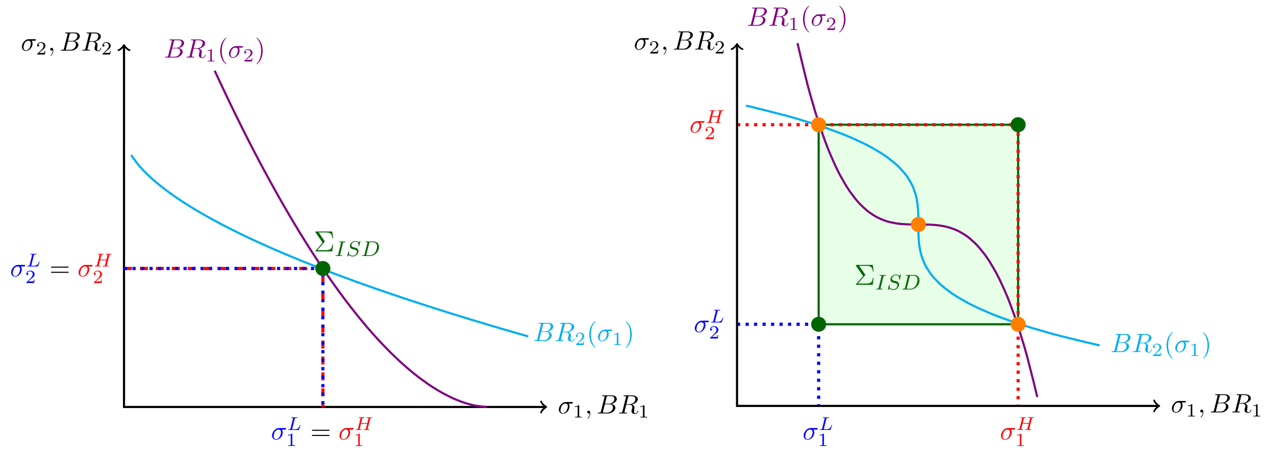
Two firms simultaneously choose whether to enter a market ($y_f = 1$) or not ($y_f = 0$). Firm f's profit is:

$$\pi_f(y_f, y_{-f}, \epsilon_f; x_f, \theta, \xi_f) = y_f \left(x_f \beta - \delta y_{f'} + \xi_f + \epsilon_f \right)$$

where ϵ_f is an independently distributed, privately observed shock, i.e., $\epsilon_f \perp \epsilon_{f'}$ for all $f \neq f'$. It is easy to see that f's optimal strategy will take the form of a threshold strategy, where the threshold corresponds to the lowest value of ϵ_f such that the profit of entry is positive, conditional on σ_{-f}.

One can show that this is an SMSG. To see this, note that $\mathcal{Y}_f = \{0, 1\}$ is a complete lattice, π_f order upper semi-continuous and supermodular in y_f (trivially so, since y_f is discrete and uni-

Figure 2: Two Player Game Best Responses in Probability Space



Note: Slightly abusing notation, I use σ_f to represents f's entry probability. $BR_f(\sigma_{-f})$ represents the optimal entry probability of firm f given the entry probability of firm f'. The left panel exhibits an (x, θ, ξ)-game with a single equilibrium, while the right panel exhibits an (x, θ, ξ)-game with multiple equilibria.

variate), and π_f has DD in $(y_f, y_{f'})$. This is:

$$\pi_f(1, y_{-f}, \epsilon_f, \epsilon_{-f}) - \pi_f(0, y_{-f}, \epsilon_f, \epsilon_{-f}) = x_f\beta - \delta y_{f'} + \xi_f + \epsilon_f$$

is decreasing in $y_{f'}$.

Figure 2 depicts the implications of Theorem 1.c for this game. Slightly abusing notation, it uses σ_f to represent the entry probability of firm f, and it depicts the best response function, BR_f, as the optimal entry probability of firm f given an entry probability for its competitor. The left panel shows the case with a unique equilibrium. In this case $\sigma^H = \sigma^L$, so the set of strategies that survives ISD, Σ_{ISD}, is singleton. The right panel shows the case with multiple equilibria. Here, $\sigma^L < \sigma^H$, so the set of strategies that survives ISD is non singleton, and is represented by green box.

2.3.2 Complete Information Entry Game

Consider the same example as above, only now $\epsilon_f = 0$ for all f, i.e., players are completely informed. Clearly the resulting game is still an SMSG.

The best response function of firm f can take three "values" depending on the realization of ξ_f. One where *entry* is dominant, i.e., $BR_f(\sigma_{-f}) = 1$ if $x_f\beta - \delta + \xi_f > 0$. One where *no entry* is dominant, i.e., $BR_f(\sigma_{-f}) = 0$ if $x_f\beta + \xi_f < 0$. And one where entry is only profitable as a monopolist:

$$BR_f(\sigma_{-f}) = \begin{cases} 1 & \text{if} \quad y_{-f} = 0 \\ 0 & \text{if} \quad y_{-f} = 1 \end{cases}$$

if $x_f\beta - \delta + \xi_f < 0 < x_f\beta + \xi_f$.

Each realization of (ξ_1, ξ_2) triggers one of nine possible game types, one for each combination of best responses for each firm. Figure 3 depicts each of these combinations. For example, if $x_f\beta - \delta + \xi_f > 0$ for $f = 1, 2$, i.e., region 9, then entry is a dominant strategy for both firms and $\sigma^L = \sigma^H = (1, 1)$. If $x_1\beta + \xi_1 < 0$ and $x_2\beta - \delta + \xi_2 < 0 < x_2\beta + \xi_2$, i.e., region 2, no-entry is dominant for firm 1, and conditional on $y_1 = 0$ entry is dominant for firm 2, hence $\sigma^L = \sigma^H = (0, 1)$.

It is easy to see that for any of these games Σ_{ISD} is singleton, i.e., $\sigma^H = \sigma^L$, except when $\xi \in$ 5. In that case no firm has a dominant strategy and everything survives ISD, i.e., $\sigma^L = (0, 0)$ and $\sigma^H = (1, 1)$.

3 ISD Bounds and Identified Set

In this section I show that SMSGs produce tractable ISD bounds, and use these to derive an identified set for the parameters of interest.

3.1 ISD Bounds

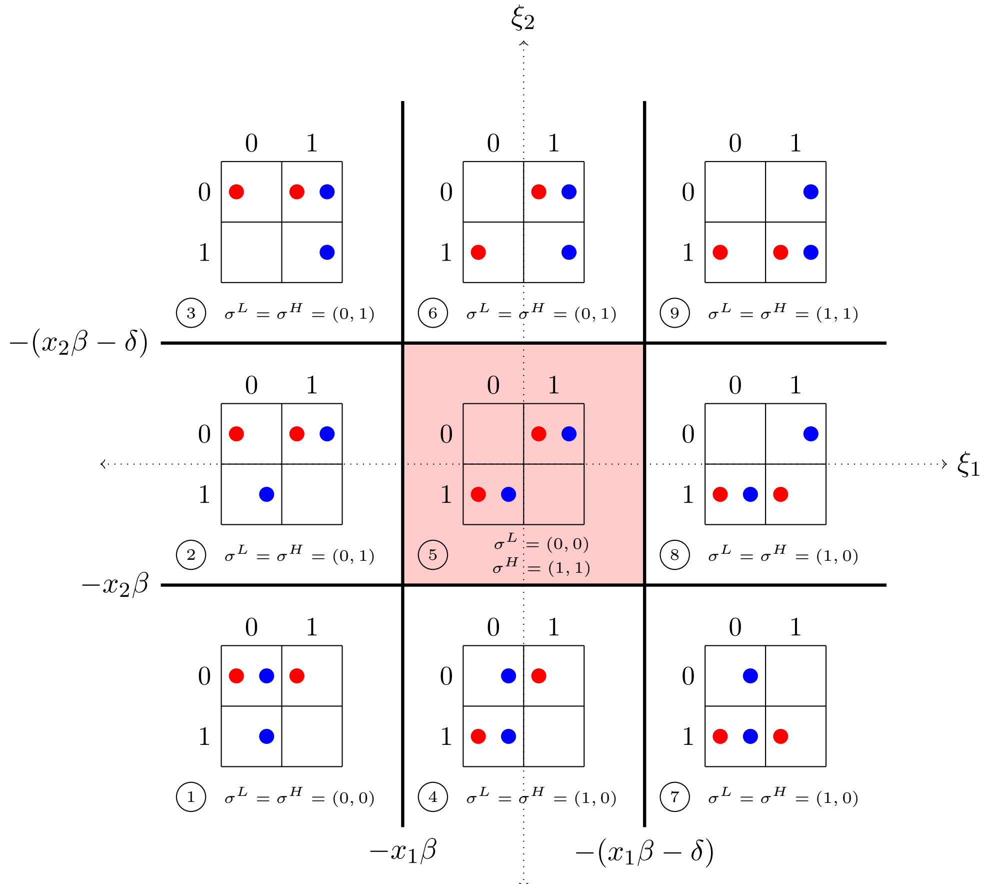
The main assumption behind ISD bounds in SMSGs, Assumption 1 below, simply says is that for all possible values of (x, θ, ξ) the (x, θ, ξ)-game is an SMSG. This assumption implies that the results of Theorem 1 hold for all (x, θ, ξ), and that the best response iterations described in Appendix A.2 apply to all (x, θ, ξ)-games. Importantly, this assumption *does not say* that the set of complements and substitutes of each firm has to be the same for all (x, θ, ξ). This is an important source of flexibility if the researcher does not want to impose the nature of strategic interactions between players, and rather wants this to be revealed by the data (as in Ciliberto and Jäkel (2021)).

Assumption 1 (SMSG Assumption)**.** *The (x, θ, ξ)-game is an SMSG for every $(x, \theta, \xi) \in \mathcal{X} \times \Theta \times \Xi$.*

Most games considered in empirical research satisfy this assumption. Nevertheless, there cases where the assumption fails such as in Fan and Yang (2022) who study a product choice game among breweries in California, or Seim (2006) who studies entry and location choices among video rental stores. In these applications Assumption 1 fails on two accounts: profits are not supermodular, and ID/DD is not guaranteed.

Theorem 2, below, derives ISD bounds in SMSGs. Importantly, the theorem provides bounds on the distribution of *any subset* of the outcome vector y, which allows me to define different identified sets depending on what $\tilde{y} \subseteq y$ one is considering. This is an important source of flexibility for empirical research, as different ISD bounds have different computational burdens and identifying power. I discuss this in more detail on Section 3.3.

Figure 3: Game Matrices with Best Responses for Values of ξ



Note: In each region ①,...,⑨, (ξ_1, ξ_2) generates a different class of games, in the sense that within each region all values of ξ generate the same best responses for both players, and when going from one region to another at least one firm changes its best response. The red dots represent the best response of firm 1. The blue dots represent the best response of firm 2.

Theorem 2 (ISD Bounds in SMSGs). *Say Assumption 1 holds, and let $H(\cdot|x,\theta)$ be the distribution of ξ. Furthermore, let $\tilde{y}$ be a subset of y (e.g., the action of firm f). The following holds:*

$$\underline{P}_{ISD}(\tilde{y}|x,\theta) \leq P^{\rho}(\tilde{y}|x,\theta) \leq \overline{P}_{ISD}(\tilde{y}|x,\theta) \tag{4}$$

where:

$$\begin{aligned}
\underline{P}_{ISD}(\tilde{y}|x,\theta) &\equiv \int_{\Xi}\int_{\mathcal{E}} \mathbb{1}\Big\{\tilde{\sigma}^{L}(\epsilon;x,\theta,\xi) = \tilde{y} = \tilde{\sigma}^{H}(\epsilon;x,\theta,\xi)\Big\} dG(\epsilon|x,\theta,\xi)dH(\xi|x,\theta) \\
P^{\rho}(\tilde{y}|x,\theta) &\equiv \int_{\Xi}\int_{\mathcal{E}} \mathbb{1}\Big\{\tilde{\sigma}^{\rho}(\epsilon;x,\theta,\xi) = \tilde{y}\Big\} dG(\epsilon|x,\theta,\xi)dH(\xi|x,\theta) \\
\overline{P}_{ISD}(\tilde{y}|x,\theta) &\equiv \int_{\Xi}\int_{\mathcal{E}} \mathbb{1}\Big\{\tilde{\sigma}^{L}(\epsilon;x,\theta,\xi) \leq \tilde{y} \leq \tilde{\sigma}^{H}(\epsilon;x,\theta,\xi)\Big\} dG(\epsilon|x,\theta,\xi)dH(\xi|x,\theta)
\end{aligned}$$

and where $\tilde{\sigma}^L$ and $\tilde{\sigma}^H$ are subsets of σ^L and σ^H corresponding to $\tilde{y}$.

Proof. Consider an arbitrary (x,θ,ξ)-game and fix an arbitrary equilibrium selection mechanism ρ. By definition, every equilibrium strategy must survive ISD, hence by Theorem 1.c $\sigma^L \leq \sigma^{\rho} \leq \sigma^H$, which implies $\sigma^L(\epsilon) \leq \sigma^{\rho}(\epsilon) \leq \sigma^H(\epsilon)$ for all ϵ. Now fix an arbitrary subset of y, $\tilde{y}$, and let $\tilde{\sigma}^L$, $\tilde{\sigma}^H$ and $\tilde{\sigma}^{\rho}$ represent the corresponding elements of σ^L, σ^H and σ^{ρ}. Clearly, the following holds: $\tilde{\sigma}^L(\epsilon) \leq \tilde{\sigma}^{\rho}(\epsilon) \leq \tilde{\sigma}^H(\epsilon)$

Say $\tilde{\sigma}^L(\epsilon) = \tilde{y} = \tilde{\sigma}^H(\epsilon)$ for some ϵ, then $\tilde{\sigma}^{\rho}(\epsilon) = \tilde{y}$. This reasoning implies the following inequality (making explicit the dependence on (x,θ,ξ)):

$$\mathbb{1}\Big\{\tilde{\sigma}^{L}(\epsilon;x,\theta,\xi) = \tilde{y} = \tilde{\sigma}^{H}(\epsilon,x,\theta,\xi)\Big\} \leq \mathbb{1}\Big\{\tilde{\sigma}^{\rho}(\epsilon;x,\theta,\xi) = \tilde{y}\Big\}$$

Integrating over ϵ and ξ yields the inequality on the left-hand side of (4).

Similarly, if $\tilde{\sigma}^{\rho}(\epsilon) = \tilde{y}$ for some ϵ, then $\tilde{\sigma}^L(\epsilon) \leq \tilde{y} \leq \tilde{\sigma}^H(\epsilon)$. This reasoning leads to the following inequality (making explicit the dependence on (x,θ,ξ)):

$$\mathbb{1}\Big\{\tilde{y} = \tilde{\sigma}^{\rho}(\epsilon;x,\theta,\xi)\Big\} \leq \mathbb{1}\Big\{\tilde{\sigma}^{L}(\epsilon;x,\theta,\xi) \leq \tilde{y} \leq \tilde{\sigma}^{H}(\epsilon,x,\theta,\xi)\Big\}$$

Integrating over ϵ and ξ yields the inequality on the right-hand side of (4). □

Theorem 2 shows that, regardless of the equilibrium selection mechanism, the probability that the model generates a sub-outcome $\tilde{y} \subseteq y$ is bounded from below by the probability that ISD *mandates* $\tilde{y}$ (i.e., all strategies that survive ISD prescribe $\tilde{y}$), and from above by the probability that ISD *allows* $\tilde{y}$ (i.e., some strategy that survives ISD prescribe $\tilde{y}$). In what follows, I refer to $\overline{P}_{ISD}$ and $\underline{P}_{ISD}$ as the ISD bounds, and use them to construct an identified set for the parameters of interest.

Before going there, however, let us discuss two noteworthy points. First, analogous ISD bounds may be built using i-level rationality, rather than "full-blown" rationality. This is, similar bounds hold for i ISD rounds, as in Aradillas-Lopez and Tamer (2008), Aradillas-Lopez (2010) and Molinari and Rosen (2008). In fact, it is easy to see that the more ISD rounds one uses, the tighter the bounds. Second, and relatedly, ISD bounds apply to any "strategy selection rule" that is consistent with i rounds of ISD. In other words, when building the identified sets, we need not assume that the data is generated by a BNE, but simply by a strategy profile that lies in the i'th level ISD set. This observation is important, as researchers may wish to explore the identifying power of different levels of rationality, or simply may not wish to assume that the data comes from a BNE. I formalize this below.

Remark 1. *Say assumption 1 holds. The i-level rationality bounds are:*

$$\begin{aligned}
\underline{P}^i_{ISD}(y|x,\theta) &= \int_{\Xi}\int_{\mathcal{E}} \mathbb{1}\left\{\sigma^{i,L}(\epsilon;x,\theta,\xi) = y = \sigma^{i,H}(\epsilon;x,\theta,\xi)\right\} dG(\epsilon|x,\theta,\xi)dH(\xi|x,\theta)\\
\overline{P}^i_{ISD}(y|x,\theta) &= \int_{\Xi}\int_{\mathcal{E}} \mathbb{1}\left\{\sigma^{i,L}(\epsilon;x,\theta,\xi) \leq y \leq \sigma^{i,H}(\epsilon;x,\theta,\xi)\right\} dG(\epsilon|x,\theta,\xi)dH(\xi|x,\theta)
\end{aligned}$$

By definition $\sigma^{i,L} \leq \sigma^{i+1,L} \leq \sigma^{i+1,H} \leq \sigma^{i,H}$, hence:

$$\underline{P}^i_{ISD}(y|x,\theta) \leq \underline{P}^{i+1}_{ISD}(y|x,\theta) \leq P^{\rho}(y|x,\theta) \leq \overline{P}^{i+1}_{ISD}(y|x,\theta) \leq \overline{P}^i_{ISD}(y|x,\theta)$$

for all $i = 0, 1, 2, \ldots$.

Finally, let ρ^i be an "i-level rationality strategy selection rule," i.e., a function that selects a strategy from the set of strategies that survives i rounds of ISD. The following holds:

$$\underline{P}^i_{ISD}(y|x,\theta) \leq P^{\rho^i}(y|x,\theta) \leq \overline{P}^i_{ISD}(y|x,\theta)$$

for all $i = 0, 1, 2, \ldots$.

All results below hold for i-level rationality and any i-level rationality strategy selection rule, ρ^i.

3.2 Two Entry-Game Example Continued

Before building the identified set, let us explore how Theorem 2 produces ISD bounds for the case of independent private information introduced in 2.3.1 and the case of complete information introduced in 2.3.2.

3.2.1 Independent Private Information Entry Game Revisited

Consider the independent private information game introduced in 2.3.1, and recall that by Theorem 2 the set of strategies that survive ISD is pinned down by the extreme strategies σ^L and σ^H. Furthermore, recall that any optimal strategy takes the form of a "threshold strategy," i.e., for each firm there is a threshold ϵ^*_f such that $\sigma_f(\epsilon) = \mathbb{1}\{\epsilon_f \geq \epsilon^*_f\}$.

Say we are interested in deriving the ISD bounds of $y_f = 1$. Figure 4 zooms into the strategies of firm f that survive ISD. The values $\epsilon^H_f < \epsilon^L_f$ represent the entry thresholds of σ^H_f, and σ^L_f, respectively. All strategies whose thresholds lie between ϵ^H_f and ϵ^L_f survive ISD, so for any ρ,

$\epsilon_f^H \leq \epsilon_f^\rho \leq \epsilon_f^L$. It is easy to see that $y_f = 1$ *is mandated by ISD*, i.e. $\sigma_1^L(\epsilon_f) = 1 = \sigma_1^L(\epsilon_f)$ when $\epsilon_1 > \epsilon_1^L$. Similarly, $y_f = 1$ *is allowed by ISD*, i.e., $\sigma_f^L(\epsilon_f) \leq 1 \leq \sigma_f^H(\epsilon_f)$, when $\epsilon_f > \epsilon_f^H$.

Finally, assuming that $\xi_f = \xi$ for $f = 1, 2$, is a market-specific binary variable that takes the values ξ_0 and $-\xi_0$ with equal probability, the ISD bounds are:

$$\begin{aligned}\underline{P}_{ISD}(y_f = 1|x,\theta) &= \frac{1}{2}Pr(\epsilon_f > \epsilon_f^L(x,\theta,\xi_0)) + \frac{1}{2}Pr(\epsilon_f > \epsilon_f^L(x,\theta,-\xi_0))\\ \overline{P}_{ISD}(y_f = 1|x,\theta) &= \frac{1}{2}Pr(\epsilon_f > \epsilon_f^H(x,\theta,\xi_0)) + \frac{1}{2}Pr(\epsilon_f > \epsilon_f^H(x,\theta,-\xi_0))\end{aligned}$$

Now say we are interested in deriving ISD bounds for $y = (1, 1)$. Figure 5 depicts the outcomes allowed by ISD for all combinations of ϵ_1 and ϵ_2. The lower bound is given by the values of ϵ such that $\sigma^L(\epsilon) = \sigma^H(\epsilon) = (1, 1)$, which occurs when both firms receive large enough profit shocks, i.e., $\epsilon_f > \epsilon_f^L$. This corresponds to the purple area in the top-right corner. Similarly, the upper bound is given by those values of ϵ for which $\sigma^L(\epsilon) \leq (1, 1) \leq \sigma^H(\epsilon)$. This occurs when both firms receive shocks $\epsilon_f > \epsilon_f^H$, which corresponds to the area inside the brown rectangle.

Assuming as before that ξ_f is a market-specific binary variable, and using independence pf ϵ_1 and ϵ_2, the ISD bounds are (making dependence on (x, θ, ξ) explicit):

$$\begin{aligned}\underline{P}_{ISD}(y = (1,1)|x,\theta) &= \frac{1}{2}\prod_{f=1,2} Pr(\epsilon_f > \epsilon_f^L(x,\theta,\xi_0)) + \frac{1}{2}\prod_{f=1,2} Pr(\epsilon_f > \epsilon_f^L(x,\theta,-\xi_0))\\ \overline{P}_{ISD}(y = (1,1)|x,\theta) &= \frac{1}{2}\prod_{f=1,2} Pr(\epsilon_f > \epsilon_f^H(x,\theta,\xi_0)) + \frac{1}{2}\prod_{f=1,2} Pr(\epsilon_f > \epsilon_f^H(x,\theta,-\xi_0))\end{aligned}$$

It is noteworthy that the ISD bounds on outcomes y are tighter than the ISD bounds on firm actions y_f. Naturally, this implies that ISD bounds on y are more informative about the parameters of interest than ISD bounds on y_f's.

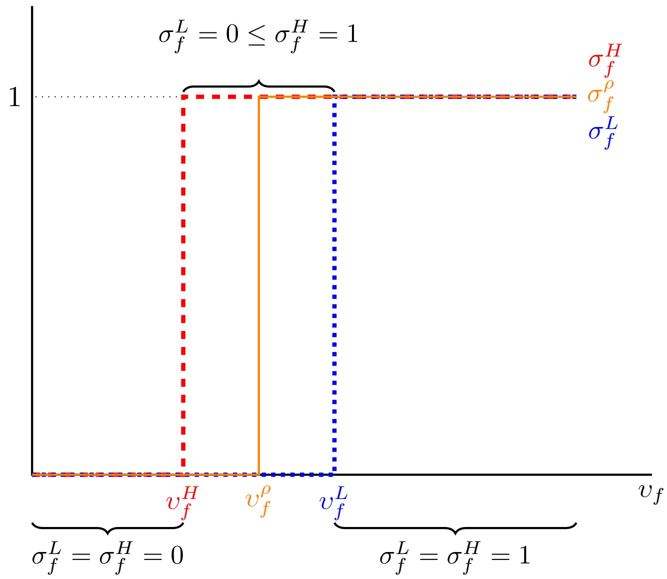
3.2.2 Perfect Information Example

Consider the complete information entry game introduced in 2.3.2. As depicted in Figure 3, every realization of (ξ_1, ξ_2) triggers a different SMSG, hence for each (ξ_1, ξ_2) the set of strategies that survive ISD is pinned down by different extreme strategies σ^H and σ^L. For example, if $\xi_f > -(\beta x_f - \delta)$ for $f = 1, 2$ then entry is dominant for both firms and $\sigma^L = \sigma^H = (1, 1)$. Similarly, if $-\beta x_f < \xi_f < -(\beta x_f - \delta)$ then ISD has no bite and $\sigma^L = (0, 0)$ and $\sigma^H = (1, 1)$.

Say we are interested in the ISD bounds for $y_1 = 1$. The lower ISD bound is given by those (ξ_1, ξ_2) for which $\sigma_1^L = 1 = \sigma_1^H$, which occurs in regions ④, ⑦, ⑧, and ⑨. Similarly, the upper bound is given by those (ξ_1, ξ_2) for which $\sigma_1^L \leq 1 \leq \sigma_1^H$, which occurs in regions ④, ⑤, ⑦, ⑧, and ⑨, Hence:

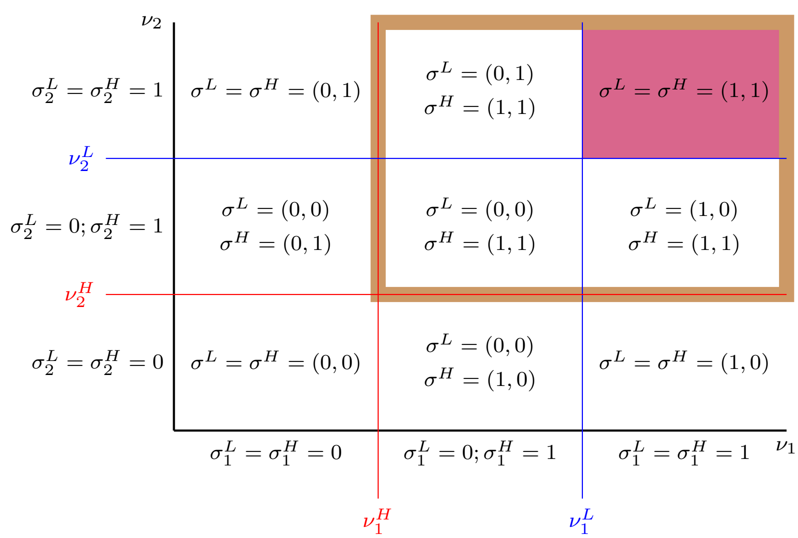
$$\begin{aligned}\underline{P}_{ISD}(y_1 = 1)|x,\theta) &= Pr\big(\xi \in ④ \cup ⑦ \cup ⑧ \cup ⑨|x,\theta\big)\\ \overline{P}_{ISD}(y_1 = 1)|x,\theta) &= Pr\big(\xi \in ④ \cup ⑤ \cup ⑦ \cup ⑧ \cup ⑨|x,\theta\big)\end{aligned}$$

Figure 4: Player Specific Extreme Strategies and Equilibrium Strategy Selected by ρ



Note: Lowest σ_f^L, highest σ_f^H and selected σ_f^ρ, strategies for an entry game of incomplete information. ϵ_f^L, ϵ_f^H and ϵ_f^ρ are the corresponding fixed cost entry thresholds.

Figure 5: Set of Actions allowed by ISD for values of ϵ. Two-Player Case.



Note: Set of actions allowed by ISD in the (ϵ_1, ϵ_2) space. For each value of ϵ_1 (resp. ϵ_2), the horizontal (rest. vertical) axis shows the actions that ISD allows for firm 1, (resp. 2) . Every rectangle, shows the outcomes allowed by the corresponding (ϵ_1, ϵ_2). For example, for outcome $y = (1,1)$ is the only outcome by ISD if and only if $\epsilon \in$ ■. Similarly, outcome $(1,1)$ is allowed by ISD for all $\epsilon \in$ □, Now, say we are interested in computing the ISD bounds of outcome $y = (1,1)$. The lower ISD bound is given by the values of ξ that make $\sigma^L = (1,1) = \sigma^H$, i.e., region ⑨. Similarly the ISD upper bound is given by the values of ξ that make $\sigma^L \leq (1,1) \leq \sigma^H$, i.e., regions ⑤ and ⑨. Hence:

$$\begin{aligned} \underline{P}_{ISD}(y = (1,1)|x,\theta) &= Pr\big(\xi \in ⑨|x,\theta\big) \\ \overline{P}_{ISD}(y = (1,1)|x,\theta) &= Pr\big(\xi \in ⑤ \cup ⑨|x,\theta\big) \end{aligned}$$

As before, ISD bounds on outcomes y are tighter than the ISD bounds on firm actions y_f, so we should expect bounds on y to be more informative about the parameters of interest than bounds on y_f's.

3.3 Identified Set

In this subsection I derive ISD identified set. To this end, consider the following assumption on the data generating process.

Assumption 2 (Data Generating Process (DGP))**.** *There is a real parameter vector and a real equilibrium selection mechanism, θ_0 and ρ_0, respectively. Furthermore, given a private shock vector ϵ, the observed outcome of the (x, θ_0, ξ)-game is:*

$$\sigma_0(\epsilon; x, \xi) = \sigma^{\rho_0}(\epsilon; x, \theta_0, \xi)$$

where $\sigma^{\rho_0}(\epsilon; x, \theta_0, \xi)$ equilibrium strategy chosen by ρ_0 in the (x, θ_0, ξ)-game, evaluated at ϵ.

Assumption 2 says that the model is correctly specified, and that the realization of each game comes from a BNE selected by an arbitrary equilibrium selection mechanism ρ_0. It is worth noting that Assumption 2 is not central to any result below. In particular, as noted in Remark 1, it could be the case that the data does not come from a BNE at all, but that some $\sigma \in \Sigma_{ISD}$ is selected by a "strategy selection mechanism," i.e., a function $\tilde{\rho}$ that chooses an element of Σ_{ISD}. All identified sets I describe below hold under this weaker assumption too.

3.3.1 Probability Identified Set: Discrete Case

Consider a partition of y, i.e., $(\tilde{y}_j)_j = y$. For example, each $\tilde{y}_j$ may represent the action taken by a firm, i.e., for each j, $\tilde{y}_j = y_f$ for some f. Let $P_0(\tilde{y}_j|x) = P^{\rho_0}(\tilde{y}_j|x, \theta_0)$ be the probability of observing $\tilde{y}_j$ according to the DGP. By Theorem 2, expression (4) holds for θ_0 and ρ_0. Hence, any θ that *violates*:

$$\underline{P}_{ISD}(\tilde{y}_j|x,\theta) \leq P_0(\tilde{y}_j|x) \leq \overline{P}_{ISD}(\tilde{y}_j|x,\theta) \tag{5}$$

for some $(x, \tilde{y}_j, j)$ cannot be the real θ. With this intuition, I define the identified set as follows.

Definition 6 (ISD Identified Set). *Let $(\tilde{y}_j)_{j=1,\ldots,J}$ be a partition of y, where J is number of sets in the partition. The identified set, Θ_{ISD}, is the collection of all $\theta \in \Theta$ such that θ satisfies* (5)*, i.e.:*

$$\Theta_{ISD} = \{\theta \in \Theta : \theta \text{ satisfies } (5), \forall x \in \mathcal{X}, \forall \tilde{y}_j \in \tilde{\mathcal{Y}}_j, \forall j \in J\}$$

where $\mathcal{Y}_j$ is the set of values that $\tilde{y}_j$ can take.

The identified set in Definition 6 is really a collection of identified sets that depend on which partition of y is being considered. Two partitions of interest are the trivial partition, i.e., $J = 1$ and $\tilde{y}_1 = y$, and the-firm level partition, i.e., $J = |\mathcal{F}|$ and for each j, $\tilde{y}_j = y_f$ for some f, although many more are possible.

The identification power of the ISD bounds will depend on what partition is being used. Intuitively, coarser partitions place more restrictions on the joint distribution of y, and therefore will lead to tighter ISD bounds (as argued in Section 3.2) and more informative identified sets. Hence, firm-level bounds (i.e., bounds in y_f) are less informative about θ than outcome-level bounds (i.e., bounds on y). The Montecarlo experiments in Section 4 support this argument.

This disadvantage, however, comes with a trade off in terms of computational burden. Typically, $\underline{P}_{ISD}$ and $\overline{P}_{ISD}$ do not have closed form solutions and need to be computed numerically, and the computational burden of doing so will depend on how coarse is the partition of $(\tilde{y}_j)_j$, with coarser partitions being more burdensome. Hence computing bounds on y (the coarsest possible partition) will be costlier than computing bounds on y_f's.

3.3.2 ISD Identified Set: Continuous Case

The identified set from Definition 6 is uninformative when y is continuous. If this is the case, then the collection of (ϵ, ξ)'s such that $\sigma(\epsilon; x, \theta, \xi) = y$ has zero measure, hence:

$$\begin{aligned} \underline{P}_{ISD}(y|x,\theta) &= 0, \forall x, \theta \\ P^{\rho}(y|x,\theta) &= 0, \forall x, \theta \end{aligned}$$

and the bounds proposed in the previous subsection will be trivially satisfied for all θ.

To bypass this issue I propose bounds on the cumulative distribution of y. To see how this works, consider an arbitrary outcome y and note that Theorem 1 implies that $\sigma^L \leq \sigma^\rho \leq \sigma^H$ for any ρ. Hence, the following inequalities hold:

$$\begin{aligned} \mathbb{1}\{\sigma^\rho(\epsilon; x, \theta, \xi) \leq y\} &\leq \mathbb{1}\{\sigma^L(\epsilon; x, \theta, \xi) \leq y\} \\ \mathbb{1}\{\sigma^\rho(\epsilon; x, \theta, \xi) \geq y\} &\leq \mathbb{1}\{\sigma^H(\epsilon; x, \theta, \xi) \geq y\} \end{aligned}$$

In words, any action profile y that is greater than the predicted outcome $\sigma^\rho(\epsilon; \cdot)$ must also be greater than the lower bound $\sigma^L(\epsilon; \cdot)$. Similarly, any action profile that is smaller than $\sigma^\rho(\epsilon; \cdot)$ must also be smaller than the upper bound $\sigma^H(\epsilon; \cdot)$. Integrating over ϵ and ξ, and slightly abusing notation, this implies the following probability-based inequalities:

$$\begin{aligned} P(\sigma^{\rho} \leq y|x,\theta) &\leq P_{ISD}(\sigma^{L} \leq y|x,\theta) \\ P(\sigma^{\rho} \geq y|x,\theta) &\leq P_{ISD}(\sigma^{H} \geq y|x,\theta) \end{aligned}$$

where the probabilities follow from integrating over the distributions $H(\xi|x,\theta)$ and $G(\epsilon|x,\theta,\xi)$.[13] Finally, let

$$\begin{aligned} P(\sigma_0 \leq y|x) &\equiv P(\sigma^{\rho_0} \leq y|x,\theta_0) \\ P(\sigma_0 \geq y|x) &\equiv P(\sigma^{\rho_0} \geq y|x,\theta_0) \end{aligned}$$

be the real probabilities, i.e., the probabilities derived from the real parameter vector and equilibrium selection mechanism (θ_0, ρ_0). Identification rests on the idea that any parameter vector θ that violates:

$$\begin{aligned} P(\sigma_0 \leq y|x) &\leq P_{ISD}(\sigma^{L} \leq y|x,\theta) \\ P(\sigma_0 \geq y|x) &\leq P_{ISD}(\sigma^{H} \geq y|x,\theta) \end{aligned} \tag{6}$$

for some (x, y) cannot have generated the data, i.e., $\theta \neq \theta_0$. The identified set, then, is defined as follows.

Definition 7 (ISD Identified Set - Continuous Case)**.** *The identified set, Θ_{ISD}, is the collection of all $\theta \in \Theta$ such that θ satisfies* (6)*, i.e.:*

$$\Theta_{ISD} = \{\theta \in \Theta : \theta \text{ satisfies } (6), \forall x \in \mathcal{X}, \forall y \in \mathcal{Y}, \} \tag{7}$$

A number of comments regarding this identified set are in order. First, as in Theorem 2 and Definition 6, the ISD bounds and the identified set for continuous variables can be defined for subsets of y rather than y itself. Second, these bounds apply for the discrete case as well as the continuous case. And third, in the case where y_f has discrete elements and continuous elements (say for an entry game followed by a pricing stage), one can use a combination of the bounds of Definition 7 and the ones defined in Section 3.3.1

[13]This is:

$$\begin{aligned} P(\sigma^{\rho} \leq y|x,\theta,\ \rho) &= \int_{\xi}\int_{\epsilon} \mathbb{1}\{\sigma^{\rho}(\epsilon; x,\theta,\xi) \leq y\} dG(\epsilon|x,\theta,\xi) dH(\xi|x,\theta) \\ P(\sigma^{\rho} \geq y|x,\theta,\rho) &= \int_{\xi}\int_{\epsilon} \mathbb{1}\{\sigma^{\rho}(\epsilon; x,\theta,\xi) \geq y\} dG(\epsilon|x,\theta,\xi) dH(\xi|x,\theta) \\ P_{ISD}(\sigma^{L} \leq y|x,\theta) &= \int_{\xi}\int_{\epsilon} \mathbb{1}\{\sigma^{H}(\epsilon; x,\theta,\xi) \leq y\} dG(\epsilon|x,\theta,\xi) dH(\xi|x,\theta) \\ P_{ISD}(\sigma^{H} \geq y|x,\theta) &= \int_{\xi}\int_{\epsilon} \mathbb{1}\{y \leq \sigma^{H}(\epsilon; x,\theta,\xi)\} dG(\epsilon|x,\theta,\xi) dH(\xi|x,\theta) \end{aligned}$$

3.3.3 ISD Identified Set: Outcome Level

Finally, rather than putting bounds on the distribution of outcomes, one can put bounds on outcomes themselves. To see this, note that Theorem 1.a implies that for any ρ:

$$\sigma^L(\epsilon; x, \theta, \xi) \leq \sigma^\rho(\epsilon; x, \theta, \xi) \leq \sigma^H(\epsilon; x, \theta, \xi)$$

Taking expectation over ϵ and ξ:

$$\underbrace{y^L(x,\theta)}_{\equiv E[\sigma^L(\epsilon;x,\theta,\xi)|x,\theta]} \leq \underbrace{y^\rho(x,\theta)}_{\equiv E[\sigma^\rho(\epsilon;x,\theta,\xi)|x,\theta]} \leq \underbrace{y^H(x,\theta)}_{\equiv E[\sigma^H(\epsilon;x,\theta,\xi)|x,\theta]}$$

Finally, letting $y_0(x) = y^{\rho_0}(x, \theta_0)$, we can define the identified set as follows.

Definition 8 (ISD Identified Set for Outcome Bounds)**.** *The identified set based on outcome level bounds is:*

$$\Theta_{ISD} = \{\theta \in \Theta : y^L(x,\theta) \leq y_0(x) \leq y^H(x,\theta), \forall x \in \mathcal{X}\}$$

It is easy to see that outcome level bounds apply for discrete or continuous y. Since computing these bounds requires integrating over y, the identified set that results from outcome level bounds is weakly larger than the ones based on probability bounds. Nevertheless, this disadvantage comes with the benefit that simulating these bounds is less computationally burdensome than any of the aforementioned alternatives.

4 Montecarlo Exercises

Here I provide some Montecarlo exercises to study the performance of ISD bounds. First, in Section 4.1, I focus on a entry game of complete information and compare ISD bounds to previous bounds proposed in the literature. Second, in Section 4.3, I provide Montecarlo experiments for the case of incomplete information with unobserved heterogeneity.

4.1 Complete Information Two-Firm Entry Game

Tamer (2003) and Ciliberto and Tamer (2009) (CT), pioneered the probability bounds approach to set identification for discrete games of complete information. This approach has also been studied by Aradillas-Lopez and Tamer (2008) and Fan and Yang (2022) (FY). In this subsection I study how ISD bounds compare to this earlier literature. To this end, I consider a complete information two-firm entry game, in the spirit of CT, and provide comparisons between CT bounds, FY bounds, and ISD bounds.[14]

[14] The bounds in Aradillas-Lopez and Tamer (2008) are based on rationalizable strategies, which coincides with ISD in two player games. As a result, their bounds and ISD bounds coincide for in this example.

Table 1: CT vs. ISD Bounds

Bound Type	Lower Bound	Upper Bound
CT $(0,1)$	$Pr\left(\xi \in ② \cup ③ \cup ⑥\right)$	$Pr\left(\xi \in ② \cup ③ \cup ⑤ \cup ⑥\right)$
ISD $(0,1)$	$Pr\left(\xi \in ② \cup ③ \cup ⑥\right)$	$Pr\left(\xi \in ② \cup ③ \cup ⑤ \cup ⑥\right)$
CT $(0,0)$	$Pr\left(\xi \in ①\right)$	$Pr\left(\xi \in ①\right)$
ISD $(0,0)$	$Pr\left(\xi \in ①\right)$	$Pr\left(\xi \in ① \cup ⑤\right)$

Consider the two-player complete information entry game described in Sections 2.3 and 3.2. Firm's get profits:

$$\pi_f(y_f, y_{f'}; x_f, \theta, \xi_f) = y_f(\beta x_f - \delta y_{f'} + \xi_f)$$

with $\theta = (\beta, \delta)$, where $\delta \geq 0$ and where $\xi \sim H(\xi)$.

Table 1 shows how CT bounds compare to ISD bounds for this game. CT bounds are built on the idea that if y is the unique equilibrium then $\sigma^\rho = y$, and if $\sigma^\rho = y$ then y must be in *an* equilibrium. In our notation:

$$\begin{array}{ccccc} \mathbb{1}\{\{y\} = \mathcal{B}(x,\theta,\xi)\} & \leq & \mathbb{1}\{y = \sigma^\rho(x,\theta,\xi)\} & \leq & \mathbb{1}\{y \in \mathcal{B}(x,\theta,\xi)\} \\ \underline{P}_{CT}(y|x,\theta) & \leq & P^\rho(y|x,\theta) & \leq & \overline{P}_{CT}(y|x,\theta)) \end{array}$$

where the second line comes from integrating over ξ, and where $\underline{P}_{CT}(y|x,\theta)$ and $\overline{P}_{CT}(y|x,\theta))$ are the integrals of the LHS and the RHS respectively, and they represent the lower and the upper CT probability bounds.

Consider outcome $y = (0,1)$. The CT lower bound for $y = (0,1)$ is the probability that $(0,1)$ is a unique equilibrium which, from Figure 3, occurs when $\xi \in ② \cup ③ \cup ⑥$, whereas the upper bound corresponds to the probability that $(0,1)$ is *an* equilibrium which occurs when $\xi \in ② \cup ③ \cup ⑤ \cup ⑥$. Coincidentally, region $② \cup ③ \cup ⑥$ is also where $\sigma^L = (0,1) = \sigma^H$, and region $② \cup ③ \cup ⑤ \cup ⑥$ is where $\sigma^L \leq (0,1) \leq \sigma^H$, so CT and ISD bounds coincide.

CT bounds are built on a stronger concept than ISD bounds (all Nash equilibria survive ISD, but not everything that survives ISD is a Nash equilibrium), hence CT bounds are (weakly) tighter than their ISD counterparts. This is the case for outcome $(0,0)$. Ignoring mixed strategies, there is no value of ξ for which $(0,0)$ is one of many equilibrium outcomes, so the CT upper and lower bounds coincide and correspond to the probability that $\xi \in ①$. For ISD bounds, in contrast, $\sigma^L = (0,0) = \sigma^H$ in region ① and $\sigma^L \leq (0,0) \leq \sigma^H$ in regions ① and ⑤, so the ISD upper bound for outcome $(0,0)$ is larger than the CT bound.

Table 2 shows how the bounds proposed by FY compare to the firm-level ISD bounds (i.e.,

Table 2: FY vs. ISD Bounds

Bound Type	Lower Bound	Upper Bound
FY $y_1 = 0$	$Pr\left(\xi \in ① \cup ② \cup ③\right)$	$Pr\left(\xi \in ① \cup ② \cup ③ \cup ④ \cup ⑤ \cup ⑥\right)$
Firm ISD $y_1 = 0$	$Pr\left(\xi \in ① \cup ② \cup ③ \cup ⑥\right)$	$Pr\left(\xi \in ① \cup ② \cup ③ \cup ⑤ \cup ⑥\right)$

bounds on y_f). FY bounds are built on the idea that if y_f is dominant then it must be part of a Nash equilibrium and therefore $\sigma_f^\rho = y_f$. Similarly, if y_f is played in a Nash equilibrium, i.e., $\sigma_f^\rho = y_f$, then y_f cannot be dominated. Somewhat abusing notation:

$$\begin{array}{rcccl} \mathbb{1}\{\{y_f\} = \tilde{\mathcal{Y}}_{f,ISD}(x,\theta,\xi)\} & \leq & \mathbb{1}\{y_f = \sigma_f^\rho(x,\theta,\xi)\} & \leq & \mathbb{1}\{y_f \in \tilde{\mathcal{Y}}_{f,ISD}(x,\theta,\xi)\} \\ \underline{P}_{FY}(y_f|x,\theta) & \leq & P^\rho(y_f|x,\theta) & \leq & \overline{P}_{FY}(y_f|x,\theta) \end{array}$$

where $\tilde{\mathcal{Y}}_{f,ISD}$ represents the set of strategies that are not dominated.[15] As before, the second line comes from integrating over ξ, and $\underline{P}_{FY}(y_f|x,\theta)$ and $\overline{P}_{FY}(y_f|x,\theta))$ are the lower and the upper FY probability bounds.

For firm 1, $y_1 = 0$ is dominant if and only if $\xi_1 < -\beta_1 x_1$, i.e., $\xi \in ① \cup ② \cup ③$, and it is not dominated if $\xi_1 < -(\beta_1 x_1 - \delta_1)$, i.e., $\xi \in ① \cup ② \cup ③ \cup ④ \cup ⑤ \cup ⑥$. The corresponding ISD bounds come from values of ξ for which $\sigma_1^L = 0 = \sigma_1^H$, which occurs when $\xi \in ① \cup ② \cup ③ \cup ⑥$, and the values of ξ for which $\sigma_f^L \leq 0 \leq \sigma_f^H$, which occurs when $\xi \in ① \cup ② \cup ③ \cup ⑤ \cup ⑥$. As one would expect, since ISD bounds are constructed going through multiple rounds ISD, they are (weakly) tighter than FY bounds which only goes though one round.

4.2 Complete Information Montecarlo

Consider the entry game described in the previous subsection, except that there are $|\mathcal{F}| \geq 2$ firms who receive payoffs, and assume that $\xi_f \sim N(0,1)$ for all f

For $|\mathcal{F}| = 2, 3$, $\beta_0 = 1$, $\delta_0 = 1/(|\mathcal{F}| - 1)$, and $\xi_f \sim N(0,1)$, I simulate $MC = 100$ samples of $M = 1000$ independent markets. For each game, x_f can take the values in $\{-2, 0, 2\}$ with equal probability, and for each realization of $x = (x_f)_f$, I compute CT bounds, outcome-level ISD bounds, firm-level ISD bounds, and FY bounds.

For inference, I use the critical values for vector inequality hypotheses of Chernozhukov et al. (2019). Letting,$(l_k(x))_k$ be a collection of K non-negative functions, I compute unconditional empirical moment inequalities. In particular, for each type of bound, $B \in$ {CT,FY,ISD outcome, ISD firm},

[15] Note that if $\tilde{\mathcal{Y}}_{f,ISD}$ is singleton, then it contains a dominant strategy.

and each for observation (X_m, Y_m), I compute the empirical moment functions:[16]

$$\begin{aligned}\overline{\psi}_B(y|Y_m, X_m, \theta) &\equiv \overline{P}_B(y|X_m, \theta) - \mathbb{1}\{y = Y_m\}\\ \underline{\psi}_B(y|Y_m, X_m, \theta) &\equiv \mathbb{1}\{y = Y_m\} - \underline{P}_B(y|X_m, \theta)\end{aligned}$$

Let $M(x) = \sum_m \mathbb{1}\{X_m = x\}$ represent the number of markets where (y, x) was observed. Furthermore, somewhat abusing notation, let $\mu \in \{\underline{\mu}, \overline{\mu}\}$, sd $= \{\underline{\text{sd}}, \overline{\text{sd}}\}$, and $t = \{\underline{t}, \overline{t}\}$ and define:

$$\begin{aligned}\mu_B(\theta; y, x) &= \frac{\sum_m \mathbb{1}\{X_m = x\}\psi_B(y|X_m, Y_m, \theta)}{M(x)}\\ \text{sd}_B(\theta; y, x) &= \left(\frac{1}{M(x)}\sum_m \mathbb{1}\{X_m = x\}\left(\psi_B(y|X_m, Y_m, \theta) - \mu_B(\theta; y, x)\right)^2\right)^{1/2}\\ t_B(\theta; y, x) &= M(x)\frac{\mu_B(\theta; y, x)}{\text{sd}_B(\theta; y, x)}\end{aligned}$$

In a nutshell, CCK propose critical values for a null hypothesis of the form $\max\{v\} \leq 0$, for a normalized vector v,[17] and find a critical value for the statistic $\max\{v\}$.

Applied to our setting, the null hypothesis corresponds to:

$$\max_{x,y}\left\{\max\{\underline{t}_B(\theta; y, x), \overline{t}_B(\theta; y, x)\}\right\} \leq 0$$

and the corresponding critical value at significance of α is:

$$CCK(\alpha) = \frac{\Phi^{-1}(1 - \alpha/2|\mathcal{X}||\mathcal{Y}|)}{\sqrt{1 - \Phi^{-1}(1 - \alpha/2|\mathcal{X}||\mathcal{Y}|)/M}}$$

where Φ^{-1} is the standard normal quantile function.

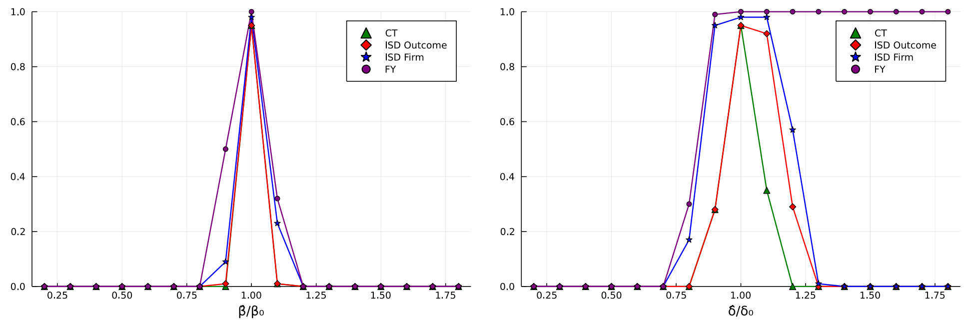
Let $\hat{\beta}$ and $\hat{\delta}$ be guesses for β_0 and δ_0 respectively. Keeping $\hat{\delta} = \delta_0$, and for each $\hat{\beta} \in \{\beta_0\frac{1}{5}, \beta_0\frac{2}{5}, \ldots, \beta_0\frac{8}{5}, \beta_0\frac{9}{5}\}$, I compute the CT bounds, ISD outcome bounds, ISD firm bounds, and FY bounds for each game in each sample. For each sample, I compute the outcome of the test described above, and from this I get the share of samples for which the null $(\hat{\beta}, \delta_0) \in \Theta_{ISD}$ gets rejected. I conduct an analogue exercise for $\hat{\delta}$ keeping $\hat{\beta} = \beta_0$. The results can be found in Figures 6 and 7.

As expected, CT bounds provide the smallest identified set, followed by ISD outcome bounds, ISD firm bounds, and FY bounds. The identified set for bothe parameters is larger with 3 firms rather than 2. This is to be expected, since the equilibrium multiplicity problem is exacerbated as the number if firms increases. It is noteworthy that FY bounds are not able to reject $(\beta_0, \hat{\delta}) \in \Theta_{ISD}$

[16] The moment functions I present here are based on outcomes, for the firm-level ISD bounds and FY bounds similar moment conditions apply.

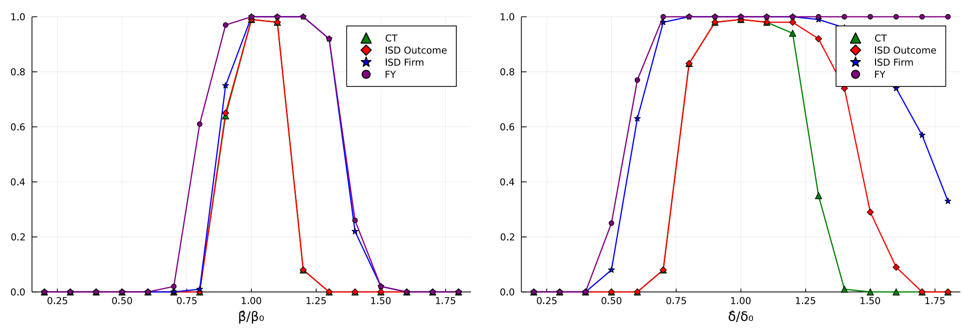
[17] This is, each element of v has mean zero and a standard deviation of one.

Figure 6: Monte Carlo Probability of $\hat{\theta} \in \Theta_{ISD}$, for $|\mathcal{F}| = 2$ under Perfect Information



Note: Left panel shows the results for β. The right panel shows the results for δ. y-axis is $Pr(\hat{\theta} \in \Theta_{ISD})$, for $\hat{\theta} \in \{\hat{\beta}, \hat{\delta}\}$, for each type of Identified Set. x-axis is $\hat{\beta}/\beta_0$ (left) and $\hat{\delta}/\delta_0$ (right).

Figure 7: Monte Carlo Probability of $\hat{\theta} \in \Theta_{ISD}$, for $|\mathcal{F}| = 3$ under Perfect Information



Note: Left panel shows the results for β. The right panel shows the results for δ. y-axis is $Pr(\hat{\theta} \in \Theta_{ISD})$, for $\hat{\theta} \in \{\hat{\beta}, \hat{\delta}\}$, for each type of Identified Set. x-axis is $\hat{\beta}/\beta_0$ (left) and $\hat{\delta}/\delta_0$ (right).

even for $\hat{\delta} = \frac{9}{5}\delta_0$. To see why this is the case note that:

$$\begin{aligned} \underline{P}_{FY}(y_f = 0|x, \theta) &= \Phi(-\beta x_f) \\ \overline{P}_{FY}(y_f = 0|x, \theta) &= \Phi(-(\beta x_f - \delta)) \end{aligned}$$

For $y_f = 0$, the lower FY bound does not depend on δ, and therefore it is uninformative about its value. The upper bound of $y_f = 0$, in contrast, is increasing in δ hence the condition $P_0(y_f = 0|x) \leq \overline{P}_{FY}(y_f = 0|x, \theta)$ is never violated by a δ that is "too large." An analogue result holds for $y_f = 1$.

I should note that this draw back is specific to the application at hand. In the product choice model studied by FY the competitive effects (i.e., the equivalent to δ) come from a Bertrand competition stage which is estimated separately, and the probability bounds are used to estimate product/firm entry costs. In that setting, FY bounds provide informative upper and lower bounds for all the parameters of interest.

4.3 Incomplete Information Entry Game with Unobserved Heterogeneity

Consider a game as described above, only now each player receives a private profitability shock, ϵ_f, which follows a standard normal distribution (i.e., private shocks are uncorrelated). Firm f's payoff is:

$$\pi_f(y_f, y_{-f}, \epsilon_f; x_f, \theta, \xi) = y_f \left(\beta x_f - \delta \sum_{f' \neq f} y_{f'} + \omega \xi + \epsilon_f \right)$$

where ξ_f both and ϵ_f follow a $\sim N(0,1)$. Parameter $\omega \geq 0$ controls the amount of unobserved heterogeneity. In particular, when $\omega = 0$ the role of unobserved heterogeneity vanishes. The parameters of interest is $\theta = (\beta, \delta, \omega)$.

It is easy to see that given any σ_{-f}, f's optimal strategy induces an entry probability:

$$p_f(\sigma_{-f}) = \Phi \left(- \left(\beta x_f - \delta \sum_{f' \neq f} Pr(\sigma_{f'} = 1) + \omega \xi \right) \right)$$

Conditional on (x, θ, ξ), we can compute the extreme strategies σ^L and σ^H, which imply extreme entry probabilities for firm f of $p_f^L(x, \theta, \xi) = p_f(\sigma_{-f}^H | x, \theta, \xi)$, and $p_f^H(x, \theta, \xi) = p_f(\sigma_{-f}^L | x, \theta, \xi)$, with $p_f^L < p_f^H$. With this we can compute ISD bounds of outcome y as (omitting dependence p^L and p^H on (x, θ, ξ) for brevity):

$$\underline{P}_{ISD}(y|x, \theta, \xi) = \prod_f (p_f^L)^{y_f} (1 - p_f^H)^{1 - y_f}$$
$$\overline{P}_{ISD}(y|x, \theta, \xi) = \prod_f (p_f^H)^{y_f} (1 - p_f^L)^{1 - y_f}$$

Finally, integrating over ξ:

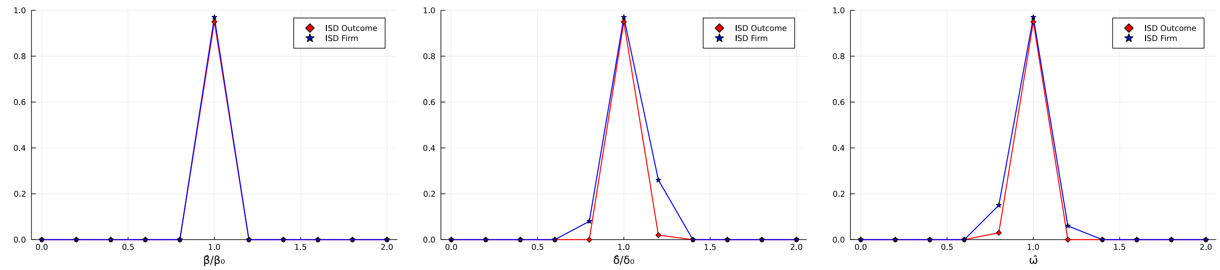
$$\underline{P}_{ISD}(y|x, \theta) = \int_{-\infty}^{\infty} \underline{P}_{ISD}(y|x, \theta, \xi) \phi(\xi) d\xi$$
$$\overline{P}_{ISD}(y|x, \theta) = \int_{-\infty}^{\infty} \overline{P}_{ISD}(y|x, \theta, \xi) \phi(\xi) d\xi$$

where ϕ represents the standard normal density.

For $|\mathcal{F}| = 2, 3$, and $\theta_0 = (\beta_0, \delta_0, \omega_0) = (1, \frac{1}{|\mathcal{F}|-1}, 1)$, I simulate $MC = 100$ samples of $M = 1000$ independent markets. For each market, x_f's take values in $\{-2, 0, 2\}$ with equal probability, and for each realization of $x = (x_f)_f$ I compute ISD bounds. Inference follows exactly as before.

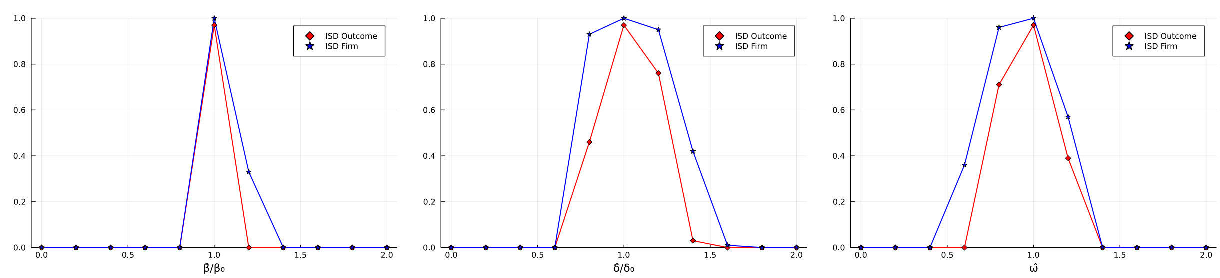
The results of this exercise can be found in Figure 8 for the two-firm case, and Figure 9 for the three-firm case. In either case, simulations show that ISD bounds do a remarkably good job at pinning down the parameters of interest. It is particularly interesting to see that the estimates are able to reject the absence of unobserved heterogeneity, i.e., $\omega = 0$.

Figure 8: $Pr(\hat{\theta} \in \Theta_{ISD})$, for $|\mathcal{F}| = 2$ under Incomplete Information and Unobserved Heterogeneity



Note: The left panel shows the results for β, the middle panel shows the result for δ, and the right panel shows the results for ω. In the left panel the y-axis shows $Pr((\hat{\beta}, \delta_0, \omega_0) \in \Theta_{ISD})$, for the firm-level and the outcome-level ISD set. The mid and right panels show analogue quantities for $\hat{\delta}$ and $\hat{\omega}$, respectively.

Figure 9: $Pr(\hat{\theta} \in \Theta_{ISD})$, for $|\mathcal{F}| = 3$ under Incomplete Information and Unobserved Heterogeneity



Note: The left panel shows the results for β, the middle panel shows the result for δ, and the right panel shows the results for ω. In the left panel the y-axis shows $Pr((\hat{\beta}, \delta_0, \omega_0) \in \Theta_{ISD})$, for the firm-level and the outcome-level ISD set. The mid and right panels show analogue quantities for $\hat{\delta}$ and $\hat{\omega}$, respectively.

5 Airline Application: Data and Empirical Model

The rest of the paper uses ISD bounds to estimate an esntry game in the airline industry as in Ciliberto and Tamer (2009). In this section I summarize the data and present the model I will use for estimation. Empirical results are reserved for Section 6.

5.1 Data

My main data source is the Origin and Destination Survey (DB1B) collected by the Bureau of Transportation Statistics (BTS). The data consists of a sample of 10% of all trips taken within the U.S. in a given quarter/year. For each trip it contains the price of the ticket as well as the origin and destination airports, and all layover airports. The DB1B is a well known data source in the airline literature, and has been used (e.g., Berry (1992), Ciliberto and Tamer (2009), Aguirregabiria and Ho (2012)).

I use the DB1B data set for the first quarter of 2005, and supplement it with information on airport locations (city) from the BTS, with county level income from the Bureau of Economic Analysis,[18] as well as county level population data from the Census Bureau.

I keep the airports located at the 70 top MSAs in terms of population, which yields a total of 72 airports. Table 3 presents the list of the top 20 airports ranked by the population of their corresponding MSAs, while Table 4 shows some airport level summary statistics.

A market corresponds to a non-directional airport pair regardless of the number of stops. With 72 airports, this would imply $2556(= 72 \cdots 71/2)$ markets, however, I drop airport pairs that lie in the same MSA (e.g. JFK and La Guardia), leaving a total of 2541 markets.

In terms of carriers, I keep America Airlines (AA), Delta (DL), United Airlines (UA), US Airways (US), and Southwest (WN). Table 5 presents summary statistics, across airports, of the number of non-stop destinations served by each carrier.

I follow Ciliberto and Tamer (2009) in defining controls. In particular, for each of these market I compute six market specific variables. *Market size* equal to the geometric mean between the population at each endpoint. *Per-capita income* and *income growth* which correspond to the average of these variables across the two end point MSAs. *Distance* and *Distance center* which correspond to the linear distance between airports and the average of the distance of each end point airport and the population weighted centroid of the U.S., which corresponds to Crawford County, Missouri. This last variable is meant to account for the fact that, due to geography, airports near the coasts or the borders have fewer closer airports than airports near the center of the country. Finally, to measure substitutability between airports, I compute *Close airport* which corresponds to the average over endpoints of the distance between the end point airport and the closest airport (including airports in the same MSA).

For each firm-market I compute two variable. First, following the insight of Berry (1992), I

[18] CAINC4: Personal Income and Employment by Major Component by County. See, https://apps.bea.gov/regional/downloadzip.cfm

Table 3: Top 20 Cities by MSA Population

Airports	City	State	Population	Per-capita Inc.	Inc. Growth
JFK, LGA	New York	NY	18.6	45.6	2.4%
LAX, LGB, SNA	Los Angeles	CA	12.6	42.0	4.2%
MDW, ORD	Chicago	IL	9.3	39.5	2.5%
PHL	Philadelphia	PA	5.8	47.2	3.6%
DAL	Dallas	TX	5.8	33.5	2.3%
HOU, IAH	Houston	TX	5.3	33.5	3.5%
FLL, MIA	Miami	FL	5.3	42.0	4.0%
DCA, IAD	Washington	DC	5.2	52.9	4.0%
ATL	Atlanta	GA	4.8	30.9	2.6%
BOS	Boston	MA	4.5	48.2	2.9%
DTW	Detroit	MI	4.4	35.7	1.2%
OAK, SFO	San Francisco	CA	4.2	61.3	2.4%
ONT	Riverside	CA	3.7	28.5	3.8%
PHX	Phoenix	AZ	3.7	29.9	4.8%
SEA	Seattle	WA	3.2	39.9	2.6%
MSP	Minneapolis	MN	3.1	36.5	3.2%
SAN, CLD	San Diego	CA	3.0	41.5	4.2%
STL	St. Louis	MO	2.8	32.8	3.4%
BWI	Baltimore	MD	2.6	42.9	3.9%
PIE, TPA	Tampa	FL	2.6	32.4	3.6%

Note: Top 20 cities in terms of MSA population, and their airports. Population is measured in millions of people. Per-capita Inc. corresponds to 2005 per-capita income in the MSA in thousands of dollars. Income growth measures the annualized growth in per-capita income between 2000 and 2005.

Table 4: Airport Summary Statistics

	Population	Per-capita Inc.	Inc. Growth	Distance Center
Mean	3.4	36.3	3.3	1367
S.D.	3.8	7.8	1.0	729
Min.	0.7	20.5	0.2	112
p25	1.2	30.4	2.6	774
p50	2.0	34.3	3.4	1235
p75	4.3	41.6	4.0	1787
Max.	18.6	61.3	5.8	2717

Note: Airport summary statistics. Population is measured in millions of people. Per-capita income corresponds to 2005 per-capita income income in thousands of dollards. Incom growth measures annualized growth in per-capita income between 2000 and 2005. Dist. Center measures the distance between the airports and the US population centroid.

Table 5: Number of Connections per Carrier (Across Airports)

Carrier	Min.	P25	Mean	Median	P75	Max.	S.D.
DL	0	12.8	24.8	24.0	38.25	64	15.2
UA	0	7.0	21.0	22.0	32.25	60	15.1
WN	0	0.0	17.6	21.5	34.0	44	15.5
AA	0	6.0	16.4	13.5	26.0	60	13.1
US	0	1.0	15.0	12.0	25.0	51	13.7

Note: Number of connections summary statistics for each carrier. A connection is a non-directional flight between two airports, regardless of the number of stops.

compute *Airport presence* as the ratio between the number of markets served by a particular carrier in a particular airport, and the total number of markets served from said airport. This variable captures is meant to capture the benefits of the hub-and-spoke network that many airlines have. These benefits range from cost reductions on the supply side due to the economies of scale and scope that arise from concentrating activities in a particular airport, and demand side benefits that arise from flying to/from well connected airports. Finally, to measure the opportunity cost of entering a market I compute the difference between the non-stop distance between the end points and the distance between them while stopping on the carriers' closest hub. I divide this quantity by direct distance and average it across the end points. This variable, which I refer to as *Cost*, is meant to capture the opportunity cost of serving a market. Table 6 below shows summary statistics for each of these variables.

Table 6: Carrier Level Summary Statistics

Carrier	Active		Airport Presence		Cost	
AA	0.67	(0.46)	0.30	(0.14)	2.0	(11.0)
DL	0.74	(0.43)	0.46	(0.16)	2.0	(12.2)
UA	0.67	(0.47)	0.39	(0.16)	1.4	(7.40)
US	0.42	(0.49)	0.27	(0.16)	2.3	(13.7)
WN	0.37	(0.48)	0.33	(0.19)	1.8	(10.0)

Note: Carrier level summary statistics, standard deviations in parenthesis. Active is the share of markets where the carrier is active.

5.2 Airline Entry Model

The model I estimate is a generalization of the entry game I presented above. There is a set of carriers, $\mathcal{F}$, that simultaneously decide whether to enter a market m, $y_{fm} = 1$, or not, $y_{fm} = 0$. A market is a directionless airport pair, regardless of whether they are connected by a direct or a one-stop flight. Carrier f in market m gets a profit of:

$$\pi_{fm} = y_{fm} \left(s'_m \beta_s + x'_{fm} \beta_f - \sum_{g \neq f} \delta^f_g y_g + \omega \xi_{fm} + \epsilon_{fm} \right) \tag{8}$$

where, s_m is a vector of market specific characteristics common to all firms (e.g., *Market size*, *Per-capita income*) and x_{fm} is a vector of firm-market specific characteristics (e.g., *Airport presence*, *Cost*). As before $\epsilon_{fm} \sim N(0,1)$ is a privately observed profit shock, and $\xi_{fm} \sim N(0,1)$ is a common knowledge profit shock which is unobserved by the econometrician. Finally, vector of parameters of interest is: $\theta = (\beta_s, (\beta_f)_{\forall f}, (\delta_g^f)_{\forall f,g:f\neq g}, \omega)$, where δ_g^f represents the competitive effect that carrier g's presence has on f's profit.

5.3 Estimation

For estimation I follow closely the approach outlined in the Montecarlo exercises in Section 4. Data consists of a entry decisions, Y_m, firm/market specific observables, X_m and S_m for a collection of $m = 1, \ldots, M$ markets. For each market m I compute ISD bounds, i.e., $\underline{P}_{ISD}(y_m|S_m, X_m, \theta)$ and $\overline{P}_{ISD}(y_m|S_m, X_m, \theta)$, as well as the following empirical moment functions:

$$\begin{aligned}
\underline{\psi}_k(y|Y_m, S_m, X_m, \theta) &= \left(1\{y_m = Y_m\} - \underline{P}_{ISD}(y|S_m, X_m, \theta)\right) l_k(S_m, X_m) \\
\overline{\psi}_k(y|Y_m, S_m, X_m, \theta) &= \left(\overline{P}_{ISD}(y|S_M, X_M, \theta) - 1\{y = Y_m\}\right) l_k(S_m, X_m)
\end{aligned}$$

for some set of non-negative functions $l_k()$, with $k = 1, \ldots, K$. Finally, for each outcome y and each function k, I compute:

$$\begin{aligned}
\mu_k(\theta; y) &= M^{-1} \sum_m \psi_k(y|X_m, Y_m, \theta) \\
\sigma_k(\theta; y) &= \left(M^{-1} \sum_m \left(\psi_k(y|X_m, Y_m, \theta) - \mu(\theta; y)\right)^2 \right)^{1/2} \\
t(\theta) &= \max_{k,y} \left\{ \max \left\{ \frac{\overline{\mu}_k(\theta; y)}{\overline{\sigma}_k(\theta; y)}, \frac{\underline{\mu}_k(\theta; y)}{\underline{\sigma}_k(\theta; y)} \right\} \right\}
\end{aligned}$$

and compare it to the critical value provided in Section 4.

In reporting the results, I use the test proposed by Chernozhukov et al. (2019) to characterize the identified set, $\Theta_C(\alpha)$, where α represents the significance. For each parameter I compute the minimum and maximum values of the parameter consistent with $\hat{\theta}$ lying in the confidence set. In other words, letting $\Theta_C(\alpha) = \{\theta \in \Theta_I : t(\theta) \leq CCK(\alpha)\}$, and for a given parameter, say δ_g^f, I report:

$$\left[\text{argmin } \left\{ \delta_g^f \text{ s.t. } \theta \in \Theta_C(\alpha) \right\}, \text{ argmax} \left\{ \delta_g^f \text{ s.t. } \theta \in \Theta_C(\alpha) \right\} \right]$$

where, obviously, δ_g^f is an element of θ.

6 Empirical Results [TBA]

7 Closing Remarks

In this paper, I provided probability bounds on (the distribution of) outcomes of games, and show that they pin down an identified set for the parameters of interests. The bounds are based on an ISD argument (ISD bounds), and are robust to multiple equilibria both in pure and mixed strategies. Furthermore, as opposed to previous bounds proposed in the literature, ISD bounds can accommodate games of discrete or continuous strategies of any dimensionality, and allow for any informational structure regarding the players' private shocks (e.g., complete information, independent private information, priviledged information), and they are informative about the underlying informational structure. i.e., different informational structures will produce different bounds.

To maximize the bite of ISD bounds I introduce the Strategically Monotonic Supermodular Games, i.e., games where payoffs are supermodular on own actions, and exhibit either increasing differences or decreasing differences between own and competitors' actions. I argue that for these games ISD is informative, in that it rules out large swaths of the strategy set, and useful, in that the bounds are easy to compute.

In Montecarlo simulations, I show that ISD bounds are informative about the parameters of interest. Furhtermore, I show that the bounds are able to inform about the relative degree of private information vs. unobserved heterogeneity in the underlying DGP.

Finally, I provide an application to the airline industry. [TBA].

References

Ackerberg, D. A. and G. Gowrisankaran (2006). Quantifying equilibrium network externalities in the ach banking industry. *The RAND Journal of Economics 37*(3), 738–761.

Aguirregabiria, V. and C.-Y. Ho (2012, May). A dynamic oligopoly game of the us airline industry: Estimation and policy experiments. *Journal of Econometrics 168*(1), 156–173.

Aguirregabiria, V. and P. Mira (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica 70*(4), 1519–1543.

Aguirregabiria, V. and P. Mira (2019). Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Quantitative Economics 10*(4), 1659–1701.

Aradillas-Lopez, A. (2010). Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics 157*(2), 409–431.

Aradillas-Lopez, A. (2011). Nonparametric probability bounds for nash equilibrium actions in a simultaneous discrete game. *Quantitative Economics 2*(2), 135–171.

Aradillas-López, A. (2020). The econometrics of static games. *Annual Review of Economics 12*(1), 135–165.

Aradillas-López, A. and A. M. Rosen (2022). Inference in ordered response games with complete information. *Journal of Econometrics 226*(2), 451–476.

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

Atal, J. P., J. I. Cuesta, and M. Sæthre (2022). Quality regulation and competition: Evidence from pharmaceutical markets.

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

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

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

Bresnahan, T. F. and P. C. Reiss (1991). Entry and competition in concentrated markets. *Journal of Political Economy 99*(5), 977–1009.

Chernozhukov, V., D. Chetverikov, and K. Kato (2019). Inference on causal and structural parameters using many moment inequalities. *The Review of Economic Studies 86*(5), 1867–1900.

Ciliberto, F. and I. C. Jäkel (2021). Superstar exporters: An empirical investigation of strategic interactions in danish export markets. *Journal of International Economics 129*, 103405.

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

de Paula, A. and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica 80*(1), 143–172.

Draganska, M., M. Mazzeo, and K. Seim (2009, Jun). Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *QME 7*(2), 105–146.

Ellickson, P. B., S. Houghton, and C. Timmins (2013). Estimating network economies in retail chains: a revealed preference approach. *The RAND Journal of Economics 44*(2), 169–193.

Fan, Y. and C. Yang (2020, May). Competition, product proliferation, and welfare: A study of the us smartphone market. *American Economic Journal: Microeconomics 12*(2), 99–134.

Fan, Y. and C. Yang (2022, June). Estimating discrete games with many firms and many decisions: An application to merger and product variety. Working Paper 30146, National Bureau of Economic Research.

Hotz, V. J. and R. A. Miller (1993, July). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

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

Li, S. Y., J. Mazur, Y. Park, J. W. Roberts, A. Sweeting, and J. Zhang (2018, January). Endogenous and selective service choices after airline mergers. Working Paper 24214, National Bureau of Economic Research.

Magnolfi, L. and C. Roncoroni (2020). Estimation of discrete games with weak assumptions on information. The warwick economics research paper series (twerps), University of Warwick, Department of Economics.

Mazzeo, M. J. (2002). Product choice and oligopoly market structure. *The RAND Journal of Economics 33*(2), 221–242.

Milgrom, P. and J. Roberts (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica 58*(6), 1255–1277.

Molinari, F. and A. M. Rosen (2008). [the identification power of equilibrium in simple games]: Comment. *Journal of Business & Economic Statistics 26*(3), 297–302.

Otsu, T. and M. Pesendorfer (2022, 04). Equilibrium Multiplicity in Dynamic Games: Testing and Estimation. *The Econometrics Journal*. utac006.

Otsu, T., M. Pesendorfer, and Y. Takahashi (2016). Pooling data across markets in dynamic markov games. *Quantitative Economics 7*(2), 523–559.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2015). Moment inequalities and their application. *Econometrica 83*(1), 315–334.

Reguant, M. (2016). Bounding equilibria in counterfactual analysis. *WP*.

Seim, K. (2006, September). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

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

Uetake, K. and Y. Watanabe (2013, 12). Estimating supermodular games using rationalizable strategies. *Advances in Econometrics 31*, 233–247.

Uetake, K. and Y. Watanabe (2020). Entry by merger: Estimates from a two-sided matching model with externalities. *SSRN*.

Van Zandt, T. and X. Vives (2007). Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory 134* (1), 339–360.

Wollmann, T. G. (2018, June). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review 108* (6), 1364–1406.

A Proof of Theorem 1 and Best Response Iteration

A.1 Proof of Theorem 1

Here I prove Theorem 2. To this end, it is useful to show lemma 1, below, which states that conditions the define an SMSG (Definition 4), hold for the game written with interim

Lemma 1 (Interim SMSG). *Let the (x, θ, ξ)-game be an SMSG. Then (omitting dependence on (x, θ, ξ) for brevity):*

1.a. ***Complete Lattice Strategy Set:*** *The strategy set Σ_f, together with the partial order "$\geq$" is a complete and compact lattice for all $f \in \mathcal{F}$, where $\sigma_f \geq \sigma'_f \Leftrightarrow \sigma_f(\epsilon_f) \geq \sigma'_f(\epsilon_f)$ for all $\epsilon_f \in \mathcal{E}_f$.*

1.b. ***Order Upper-Semi Continuity:*** *The interim profit function, Π_f, is order upper semi-continuous. This is, for any totally ordered set $C \subset \mathcal{Y}_f$:*

$$\limsup_{y_f \in C, y_f \downarrow \inf(C)} \Pi_f(y_f, \sigma_{-f}, \epsilon_f) \leq \Pi_f(\inf(C), \sigma_{-f}, \epsilon_f)$$
$$\limsup_{y_f \in C, y_f \uparrow \sup(C)} \Pi_f(y_f, \sigma_{-f}, \epsilon_f) \leq \Pi_f(\sup(C), \sigma_{-f}, \epsilon_f)$$

for all $\sigma_{-f} \in \Sigma_{-f}$, all $f \in \mathcal{F}$.

1.c. ***Supermodularity:*** *The interim profit function Π_f is supermodular in y_f for all σ_{-f}.*

1.d. ***Strategic Monotonicity:*** *For all $\epsilon_f \in \mathcal{E}_f$, and all $f, f' \in \mathcal{F}$, if $f' \in C(f)$ then Π_f has ID in $(y_f, \sigma_{f'})$, and if or $f' \in S(f)$, then Π_f has DD in $(y_f, \sigma_{f'})$.*

Proof. Fix an arbitrary SMSG. I begin by showing that Σ_f, together with the partial order $\leq$, where $\sigma_f \leq \sigma'_f \Leftrightarrow \sigma_f(\epsilon_f) \leq \sigma'_f(\epsilon_f)$, for all ϵ_f,[19] conform a complete lattice.

Take two strategies σ_f and σ'_f. By definition, for all ϵ_f, $\sigma_f(\epsilon_f), \sigma'_f(\epsilon_f) \in \mathcal{Y}_f$, hence $\sup\{\sigma_f(\epsilon_f), \sigma'_f(\epsilon_f)\} \in \mathcal{Y}_f$ and $\inf\{\sigma_f(\epsilon_f), \sigma'_f(\epsilon_f)\} \in \mathcal{Y}$, for all ϵ, which implies $\sup\{\sigma_f, \sigma'_f\}, \inf\{\sigma_f, \sigma'_f\} \in \Sigma_f$.

This shows that Σ_f is a lattice. The argument for completeness is analogous. Consider a collection of strategies $\tilde{\Sigma}_f \subseteq \Sigma_f$, and let $\tilde{\mathcal{Y}}_f(\epsilon_f) = \{y_f \in \mathcal{Y}_f : \sigma_f(\epsilon_f) = y_f \text{ for some } \sigma_f \in \tilde{\Sigma}_f\}$. Since $\tilde{\mathcal{Y}}_f(\epsilon_f) \subseteq \mathcal{Y}_f$, and $\mathcal{Y}_f$ is a complete lattice, then $\sup\{\tilde{\mathcal{Y}}_f(\epsilon_f)\}, \inf\{\tilde{\mathcal{Y}}_f(\epsilon_f)\} \in \mathcal{Y}_f$ for all ϵ_f, which implies $\sup\{\tilde{\Sigma}_f\}, \inf\{\tilde{\Sigma}_f\} \in \Sigma_f$.

To see that Π_f is order upper semi-continuous simply fix a strategy for f's competitors σ_{-f}. By order upper semi0continuity of π_f, for any ϵ and any totally ordered set $C \subset \mathcal{Y}_f$:

$$\limsup_{y_f \in C, y_f \downarrow \inf(C)} \pi_f(y_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \leq \pi_f(\inf(C), \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f})$$
$$\limsup_{y_f \in C, y_f \uparrow \sup(C)} \pi_f(y_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \leq \pi_f(\sup(C), \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f})$$

[19] I slightly abuse notation by using "$\leq$" to denote the standard vector inequality and the partial order in Σ.

Integrating over ϵ_{-f}.

$$\limsup_{y_f \in C, y_f \downarrow \inf(C)} \Pi_f(y_f, \sigma_{-f}, \epsilon_f) \leq \Pi_f(\inf(C), \sigma_{-f}, \epsilon_f)$$
$$\limsup_{y_f \in C, y_f \uparrow \sup(C)} \Pi_f(y_f, \sigma_{-f}, \epsilon_f) \leq \Pi_f(\sup(C), \sigma_{-f}, \epsilon_f)$$

as desired.

To see that Π_f is supermodular consider any two actions y_f and y'_f, and fix an arbitrary strategy for f's competitors σ_{-f}. By supermodularity of π_f, for any $\epsilon = (\epsilon_f, \epsilon_{-f})$:

$$\begin{aligned} & \pi_f(\sup\{y_f, y'_f\}, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) + \pi_f(\inf\{y_f, y'_f\}, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \\ \geq \; & \pi_f(y_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) + \pi_f(y'_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \end{aligned}$$

which, integrating over ϵ_{-f}, yields:

$$\Pi_f(\sup\{y_f, y'_f\}, \sigma_{-f}, \epsilon_f) + \Pi_f(\inf\{y_f, y'_f\}, \sigma_{-f}, \epsilon_f) \geq \Pi_f(y_f, \sigma_{-f}, \epsilon_f) + \Pi_f(y'_f, \sigma_{-f}, \epsilon_f)$$

Finally, I show that if π_f has ID in (y_f, y_{-f}), then Π_f has ID in (y_f, σ_{-f}) (the proof for the DD case is analogous). Fix actions $y'_f \geq y_f$ and a pair strategy for f's competitors, $\sigma'_{-f} \geq \sigma_{-f}$. By ID of π_f, for any ϵ_f:

$$\begin{aligned} & \pi_f(y'_f, \sigma'_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) - \pi_f(y'_f, \sigma'_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \geq \\ & \quad \pi_f(y'_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) - \pi_f(y'_f, \sigma_{-f}(\epsilon_{-f}), \epsilon_f, \epsilon_{-f}) \end{aligned}$$

Integrating over ϵ_{-f},

$$\Pi_f(y'_f, \sigma'_{-f}, \epsilon_f) - \Pi_f(y_f, \sigma'_{-f}, \epsilon_f) = \Pi_f(y'_f, \sigma_{-f}, \epsilon_f) - \Pi_f(y_f, \sigma_{-f}, \epsilon_f)$$

as desired. $\square$

Having shown Lemma 1, we are in a position to show Theorem 1, which I restate below.

Theorem 1. *Let the (x, θ, ξ)-game be an SMSG, and let Σ^i_{ISD} denote the set of strategies that survive i ISD rounds. Furthermore, let $\sigma \leq \sigma'$ if and only if $\sigma(\epsilon) \leq \sigma'(\epsilon)$ for all ϵ. The following holds (omitting dependence on (x, θ, ξ) for brevity):*

1.a For all $i = 0, 1, 2, \ldots$, there exists $\sigma^{i,L}, \sigma^{i,H} \in \Sigma$ such that $\sigma^{i,L} \leq \sigma^{i,H}$, and such that the set of strategies that survive i rounds of ISD is:

$$\Sigma^i_{ISD} = \{\sigma \in \Sigma : \sigma^{i,L} \leq \sigma \leq \sigma^{i,H}\}$$

1.b Both $\sigma^{i,L}$ and $\sigma^{i,H}$ result from a sequence of best response iterations.

1.c As $i \to \infty$, $(\sigma^{i,L}, \sigma^{i,L}) \to (\sigma^L, \sigma^H)$, with $\sigma^L \leq \sigma^H$.

Proof. I start by generalizing Lemma 1 of Milgrom and Roberts (1990) to the case of Strategic Monotonicity. Consider an SMSG and let $\tilde{\Sigma}(s^L, s^H) = \{\sigma \in \Sigma : s^L \leq \sigma \leq s^H\}$ for some pair of strategy profiles $s^L \leq s^H$ in Σ. Let $\lambda_f^L(\sigma_{-f})$ and $\lambda_f^H(\sigma_{-f})$ be f's lowest and highest best responses to σ_{-f} in $\tilde{\Sigma}_f(s_f^L, s_f^H)$.[20] Furthermore, let $\lambda_f^L(\epsilon_f; \sigma_{-f})$ and $\lambda_f^H(\epsilon_f; \sigma_{-f})$ be these strategies evaluated at ϵ_f. Finally let $\sigma_{-f}^B = (s_{C(f)}^H, s_{S(f)}^L)$ be the "best case scenario" for firm f, i.e., the case where f's complements are playing their highest possible strategy and f's substitutes are playing their lowest possible strategy. I argue that any $\sigma_f \in \tilde{\Sigma}_f(s_f^L, s_f^H)$, such that $\sigma_f \not\leq \lambda_f^H(\sigma_{-f}^B)$ is strictly dominated (in $\tilde{\Sigma}(s_f^L, s_f^H)$) by $\inf\{\sigma_f, \lambda_f^H(\sigma_{-f}^B)\}$.

Fix $\sigma_f \not\leq \lambda_f^H(\sigma_{-f}^B)$. By definition there is at least one ϵ_f such that $\sigma_f(\epsilon_f) \not\leq \lambda_f^H(\epsilon_f; \sigma_{-f}^B)$, so for such ϵ_f, $\sigma_f(\epsilon_f) \geq \inf\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; \sigma_{-f}^B)\}$. Then, for any $\sigma_{-f} = (\sigma_{C(f)}, \sigma_{S(f)}) \in \tilde{\Sigma}_{-f}(s^L, s^H)$:

$$\begin{array}{rcl}
& \Pi_f\Big(\sigma_f(\epsilon_f), \big(\sigma_{C(f)}, \sigma_{S(f)}\big), \epsilon_f\Big) & - \quad \Pi_f\Big(\inf\Big\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; \sigma_{-f}^B)\Big\}, \big(\sigma_{C(f)}, \sigma_{S(f)}\big), \epsilon_f\Big) \\
< & \Pi_f\Big(\sigma_f(\epsilon_f), \big(s_{C(f)}^H, \sigma_{S(f)}\big), \epsilon_f\Big) & - \quad \Pi_f\Big(\inf\Big\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; \sigma_{-f}^B)\Big\}, \big(s_{C(f)}^H, \sigma_{S(f)}\big), \epsilon_f\Big) \\
< & \Pi_f\Big(\sigma_f(\epsilon_f), \big(s_{C(f)}^H, s_{S(f)}^L\big), \epsilon_f\Big) & - \quad \Pi_f\Big(\inf\Big\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; \sigma_{-f}^B)\Big\}, \big(s_{C(f)}^H, s_{S(f)}^L\big), \epsilon_f\Big) \\
\leq & \Pi_f\Big(\sup\big\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; \sigma_{-f}^B)\big\}, \sigma_{-f}^B, \epsilon_f\Big) & - \quad \Pi_f\Big(\lambda_f^H(\epsilon_f; \sigma_{-f}^B), \sigma_{-f}^B, \epsilon_f\Big) \\
\leq & 0 &
\end{array}$$

where the first inequality uses the fact Π_f has ID in $(y_f, \sigma_{C(d)})$, and the second inequality comes from the fact that Π_f has DD in $(y_f, \sigma_{S(f)})$. The third comes from supermodularity of Π_f and from substituting $\sigma_{-f}^B = (s_{C(f)}^H, s_{S(f)}^L)$, while the fourth inequality follows from fact that $\lambda_f^L(\epsilon_f; \sigma_{-f}^B)$ is a maximizes Π_f given σ_{-f}^B and ϵ_f. It follows that:

$$\Pi_f(\sigma_f(\epsilon_f), \sigma_{-f}, \epsilon_f) < \Pi_f(\inf\{\sigma_f(\epsilon_f), \lambda_f^H(\epsilon_f; s_{-f}^B)\}, \sigma_{-f}, \epsilon_f)$$

for all $\sigma_{-f} \in \tilde{\Sigma}_{-f}(s_{-f}^L, s_{-f}^H)$.

Letting $\sigma_{-f}^W = (s_{C(f)}^L, s_{S(f)}^H)$ be the "worst case scenario," for f, an analogue argument shows that $\sigma_f \not\geq \lambda_f^L(\sigma_{-f}^W)$ is strictly dominated by $\sup\{\sigma_f, \lambda_f^L(\sigma_{-f}^W)\}$.

From these two results, it follows that every strategy in $\tilde{\Sigma}_f(s^L, s^H) \setminus \tilde{\Sigma}_f(\lambda^L(\sigma^W), \lambda^H(\sigma^B))$ is strictly dominated by a strategy in $\Sigma_f(\lambda^L(\sigma^W), \lambda^H(\sigma^B))$, and can be safely discarded.

This concludes the generalization of Lemma 1 from Milgrom and Roberts (1990) to the case of strategic monotonicity. With this result at hand, we are in a position to conduct ISD on the original

[20] By assumption 1 and 1, these are guaranteed to exist.

game. To this end, consider the following sequence.

$$
\begin{aligned}
&\textbf{Set up:} && \\
\Sigma_f^0 &= \Sigma \\
\mathcal{Y}_f^0(\epsilon_f) &= \mathcal{Y}_f \\
\sigma_f^{H,0} &= \{\sup\{\mathcal{Y}_f\} : \epsilon_f \in \mathcal{E}_f\} \\
\sigma_f^{L,0} &= \{\inf\{\mathcal{Y}_f\} : \epsilon_f \in \mathcal{E}_f\} \\
\Sigma_f^i &= \{\sigma_f \in \Sigma_f : \sigma_f^{L,i} \leq \sigma_f \leq \sigma_f^{H,i}\} \\
\Sigma^i &= \bigtimes_{f \in F} \Sigma_f^i \\
\mathcal{Y}_f^i(\epsilon_f) &= \{y_f \in \mathcal{Y}_f : \sigma_f^{L,i}(\epsilon_f) \leq y_f \leq \sigma_f^{H,i}(\epsilon_f)\} \\
&\textbf{Best/Worst:} && \\
\sigma_{-f}^{B,i} &= \left(\sigma_{C(f)}^{H,i}, \sigma_{S(f)}^{L,i}\right) \\
\sigma_{-f}^{W,i} &= \left(\sigma_{C(f)}^{L,i}, \sigma_{S(f)}^{L,i}\right) \\
&\textbf{Update:} && \\
\sigma_f^{H,i}(\epsilon_f) &= \sup\left\{\underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{B,i-1}, \epsilon_f)\right\} \\
\sigma_f^{L,i}(\epsilon_f) &= \inf\left\{\underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{W,i-1}, \epsilon_f)\right\} \\
\sigma_f^{H,i} &= \left\{\sigma_f^{L,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f\right\} \\
\sigma_f^{L,i} &= \left\{\sigma_f^{H,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f\right\}
\end{aligned}
$$

Consider the i'th game of the sequence described above and note that $\Sigma^i = \tilde{\Sigma}(\sigma^{L,i}, \sigma^{H,i})$. By the result above, any strategy in $\Sigma^i \setminus \Sigma^{i+1}$, is strictly dominated and can be safely discarded. Hence, by induction, each step in the sequence corresponds to an ISD step. This proves parts 1.a and 1.b.

Part 1.c follows from the fact that $\sigma^{L,i}$ is increasing, and $\sigma^{H,i}$ is decreasing, in i. □

A.2 Applying ISD

Here I show how to apply ISD to an SMSG to find the extreme strategy profiles (σ^L, σ^H). In particular, I outline the sequence of best response iterations that result in ISD steps for three cases of interest: the pure ID case, where all players are complements, i.e., $C(f) = \mathcal{F} \setminus \{f\}$ for all f; the pure DD case, where all players are substitutes, i.e., $S(f) = \mathcal{F} \setminus \{f\}$ for all f; and the general case. The pure ID case encompasses coordination games, while the pure DD case encompasses games with strategic substitution (like the entry game in the example).

A.2.1 Pure ID Case

In this case, the best response iteration that converges to (σ^L, σ^H) follows directly from Milgrom and Roberts (1990) and Van Zandt and Vives (2007). The details of the sequence are outlined in (9).

To get an intuition, consider the case where y_f is uni-variate, and start from f's "best case scenario,"[21] i.e., $\sigma_{-f}^{H,0}(\epsilon_{-f}) = \sup\{\mathcal{Y}_{-f}\}$ for all ϵ_{-f}. By ID, f's best response to $\sigma_{-f}^{H,0}$, i.e., $\sigma_f^{H,1}$, is the largest strategy that f can optimally choose, and it strictly dominates all $\sigma_f > \sigma_f^{H,1}$. Since this holds for all f, all strategy profiles $\sigma > \sigma^{H,1}$ are eliminated by $\sigma^{H,1}$. Iterating over this procedure yields the largest strategy profile not eliminated by ISD, σ^H. An analogous sequence, starting from $\sigma^{L,0}$, yields σ^L.

ISD sequence for the ID case.

$$
\begin{array}{rcl}
\multicolumn{3}{l}{\textbf{Set-up}} \\
\hline
\sigma_f^{L,0}(\epsilon_f) & = & \inf\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\sigma_f^{H,0}(\epsilon_f) & = & \sup\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\Sigma_f^i & = & \{\sigma_f \in \Sigma_f : \sigma_f^{L,i} \leq \sigma_f \leq \sigma_f^{H,i}\} \\
\Sigma^i & = & \times_{f \in F} \Sigma_f^i \\
\mathcal{Y}_f^i(\epsilon_f) & = & \left\{ y_f \in \mathcal{Y}_f : \sigma_f^{L,i}(\epsilon_f) \leq y_f \leq \sigma_f^{H,i}(\epsilon_f) \right\} \\
\\
\multicolumn{3}{l}{\textbf{ISD Step}} \\
\hline
\sigma_f^{L,i}(\epsilon_f) & = & \inf \left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{L,i-1}, \epsilon_f) \right\} \\
\sigma_f^{H,i}(\epsilon_f) & = & \sup \left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{H,i-1}, \epsilon_f) \right\} \\
\sigma_f^{L,i} & = & \left\{ \sigma_f^{L,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f \right\} \\
\sigma_f^{H,i} & = & \left\{ \sigma_f^{H,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f \right\}
\end{array}
\tag{9}
$$

A.2.2 Pure DD Games

The intuition for the pure DD case is similar. Consider the case of uni-variate y_f for all f, and start from f's "best case scenario," i.e., $\sigma_{-f}^{L,0}(\epsilon_{-f}) = \inf\{\mathcal{Y}_{-f}\}$ for all ϵ_{-f}, and its "worst case scenario," i.e., $\sigma_{-f}^{H,0}(\epsilon_{-f}) = \sup\{\mathcal{Y}_{-f}\}$ for all ϵ_{-f}. In the best case scenario, DD implies that $\sigma_f^{H,1}$ is the largest strategy that player f could plausibly choose, hence any $\sigma_f > \sigma_f^{H,1}$ is dominated by $\sigma_f^{H,1}$. Since

[21] Here I am using the terms "best case scenario" ("worst case scenario") loosely to mean "the strategy choice by f's competitors that maximizes (minimizes) f's strategy choice." ID does not imply increasingness of π_f with respect to y_{-f} (nor does DD imply decreasingness of π_f with respect to y_{-f}), so these terms should not be taken to mean "the σ_{-f} that maximizes (minimizes) f's profits."

this is true for all f, we can discard all $\sigma > \sigma^{H,1}$. Similarly, in the worst case scenario, $\sigma_f^{L,1}$ is the smallest best response that f could plausibly choose, hence any $\sigma_f < \sigma_f^{L,1}$ is strictly dominated by $\sigma_f^{L,1}$. Since this is true for all f we can safely discard all $\sigma < \sigma^{L,1}$. Putting these two arguments together, we build a new game with strategy set $\Sigma^1 = \{\sigma \in \Sigma : \sigma^{L,1} \leq \sigma \leq \sigma^{H,1}\}$. Finally, applying this argument iteratively, yields the extreme strategy profiles σ^L and σ^H.

ISD sequence for the DD case.

$$
\begin{array}{rcl}
\multicolumn{3}{l}{\textbf{Set-up}} \\
\hline
\sigma_f^{L,0}(\epsilon) & = & \inf\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\sigma_f^{H,0}(\epsilon) & = & \sup\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\Sigma_f^i & = & \{\sigma_f \in \Sigma_f : \sigma_f^{L,i} \leq \sigma_f \leq \sigma_f^{H,i}\} \\
\Sigma^i & = & \underset{f \in F}{\times} \Sigma_f^i \\
\mathcal{Y}_f^i(\epsilon_f) & = & \left\{ y_f \in \mathcal{Y}_f : \sigma_f^{L,i}(\epsilon_f) \leq y_f \leq \sigma_f^{H,i}(\epsilon_f) \right\} \\
\\
\multicolumn{3}{l}{\textbf{ISD Step}} \\
\hline
\sigma_f^{H,i}(\epsilon_f) & = & \sup \left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{L,i-1}, \epsilon_f) \right\} \\
\sigma_f^{L,i}(\epsilon_f) & = & \inf \left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{H,i-1}, \epsilon_f) \right\} \\
\sigma_f^{L,i} & = & \left\{ \sigma_f^{L,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f \right\} \\
\sigma_f^{H,i} & = & \left\{ \sigma_f^{H,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f \right\}
\end{array}
\tag{10}
$$

A.2.3 General Case

The sequence specified in (11) converges to (σ^L, σ^H) for the general case. The intuition is similar to the previous cases, with the complication that the "best case scenario" and the "wort case scenario" for firm f involve slightly more intricate strategies for its competitors.

As before, to get an intuition consider the case of uni-variate y_f, and let $\sigma_{-f}^{L,0}(\epsilon_{-f}) = \inf\{\mathcal{Y}_{-f}\}$ and $\sigma_{-f}^{H,0}(\epsilon_{-f}) = \sup\{\mathcal{Y}_{-f}\}$ for all ϵ_{-f}. The "best case scenario" for firm f is that all its complements (substitutes) play their highest (lowest) strategy, this is: $\sigma_{-f}^{B,0} = (\sigma_{C(f)}^{H,0}, \sigma_{S(f)}^{L,0})$. By ID/DD, f's best response to this strategy is the largest strategy that f could play, $\sigma_f^{H,1}$, so that any $\sigma_f > \sigma_f^{H,1}$ is dominated by $\sigma_f^{H,1}$. Because this argument applies to all firms, any $\sigma > \sigma^{H,1}$ is discarded.

Similarly, the "worst case scenario," for firm f is for its complements (substitutes) to play their lowest (highest) strategies. This is: $\sigma^{W,0} = (\sigma_{C(f)}^{L,0}, \sigma_{S(f)}^{H,0})$. By ID/DD firm f's best response, $\sigma_f^{L,1}$, is the lowest strategy that it can plausibly play, so any $\sigma_f < \sigma_f^{L,1}$ is dominated by $\sigma_f^{L,1}$. Since this argument applies to all f, any $\sigma \leq \sigma^{L,1}$ is discarded.

Putting these step together, we can build a new game with strategies $\Sigma^1 = \left\{\sigma : \sigma^{L,1} \leq \sigma \leq \sigma^{H,1}\right\}$. Iterating over this procedure we get (σ^L, σ^H).

ISD sequence for the general case.

$$
\begin{array}{rcl}
\multicolumn{3}{l}{\textbf{Set-up}} \\
\hline
\sigma_f^{H,0}(\epsilon_f) & = & \sup\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\sigma_f^{L,0}(\epsilon_f) & = & \inf\{\mathcal{Y}_f\}, \forall \epsilon_f \in \mathcal{E}_f \\
\mathcal{Y}_f^0(\epsilon_f) & = & \mathcal{Y}_f \\
\Sigma_f^i & = & \{\sigma_f \in \Sigma_f : \sigma_f^{L,i} \leq \sigma_f \leq \sigma_f^{H,i}\} \\
\Sigma^i & = & \times_{f \in F} \Sigma_f^i \\
\mathcal{Y}_f^i(\epsilon_f) & = & \{y_f \in \mathcal{Y}_f : \sigma_f^{L,i}(\epsilon_f) \leq y_f \leq \sigma_f^{H,i}(\epsilon_f)\} \\
 & & \\
\multicolumn{3}{l}{\textbf{Best/Worst Case Scenario for } f} \\
\hline
\sigma_{-f}^{B,i} & = & \left(\sigma_{C(f)}^{H,i}, \sigma_{S(f)}^{L,i}\right) \\
\sigma_{-f}^{W,i} & = & \left(\sigma_{C(f)}^{L,i}, \sigma_{S(f)}^{L,i}\right) \\
 & & \\
\multicolumn{3}{l}{\textbf{ISD Step}} \\
\hline
\sigma_f^{H,i}(\epsilon_f) & = & \sup\left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{B,i-1}, \epsilon_f) \right\} \\
\sigma_f^{L,i}(\epsilon_f) & = & \inf\left\{ \underset{y_f \in \mathcal{Y}_f^{i-1}(\epsilon_f)}{\operatorname{argmax}} \quad \Pi_f(y_f, \sigma_{-f}^{W,i-1}, \epsilon_f) \right\} \\
\sigma_f^{H,i} & = & \left\{\sigma_f^{L,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f\right\} \\
\sigma_f^{L,i} & = & \left\{\sigma_f^{H,i}(\epsilon_f) : \epsilon_f \in \mathcal{E}_f\right\}
\end{array}
\tag{11}
$$