Quantitative Economics 16 (2025), 235–266 1759-7331/20250235

An ordinal approach to the empirical analysis of games with monotone best responses

Natalia Lazzati
Department of Economics, University of California at Santa Cruz

John K.-H. Quah
Department of Economics, National University of Singapore

Koji Shirai
School of Economics, Kwansei Gakuin University

We develop a nonparametric and ordinal approach for testing pure strategy Nash equilibrium play in games with monotone best responses, such as those with strategic complements/substitutes. The approach makes minimal assumptions on unobserved heterogeneity, requires no parametric assumptions on payoff functions, and no restriction on equilibrium selection from multiple equilibria. The approach can also be extended in order to make inferences and predictions. Both model-testing and inference can be implemented by a tractable computation procedure based on column generation. To illustrate how our approach works, we include an application to an IO entry game.

Keywords. Revealed preference, monotone comparative statics, single-crossing differences, supermodular games, revealed monotonicity axiom.

JEL classification. C1, C6, C7, D4, L1.

Natalia Lazzati: nlazzati@ucsc.edu
John K.-H. Quah: ecsqkhj@nus.edu.sg
Koji Shirai: kshirai1985@kwansei.ac.jp
For helpful discussions and comments, the authors are grateful to S. Berry, J. Fox, K. Hirano, T. Hoshino, A. Kajii, Y. Kitamura, B. Kline, E. Krasnokutskaya, C. Manski, W. Newey, T. Sekiguchi, J. Stoye, B. Strulovici, S. Takahashi, Y. Takahashi, and especially X. Tang. Various versions of this project have been presented to audiences at the following events and we are grateful for their comments: seminars at University of Arizona, Johns Hopkins, Kyoto, Louvain (CORE), New York University, Rice University, Simon Fraser University, the National University of Singapore, Northwestern University, University of Paris (Dauphine), Queensland, Shanghai University of Finance and Economics, University of Southern California, Singapore Management University, Stanford, UC Davis, UC San Diego, the Canadian Economic Theory Conference (Vancouver, 2017), the Conference on Econometrics for Incomplete Models (CeMMAP and Northwestern, 2018), 13th Greater New York Metropolitan Area Econometrics Colloquium (Princeton University, 2018), and the Econometric Society North American Summer Meeting (UC Davis, 2018). Koji Shirai gratefully acknowledges financial support from the Japan Society for Promotion of Science (KAKENHI 19K00155) and the hospitality of Johns Hopkins University during his visit in the 2019–2020 academic year.

Available at http://qeconomics.org. https://doi.org/10.3982/QE2192

1. Introduction

Economic analysis is often concerned with the effect of an exogenous or strategic variable on an agent's decision: Would a consumer buy more of good A if the price of good B falls? Would a firm follow its rival when the latter raises its price? Is someone more likely to join a demonstration if more people are participating? The theory of monotone comparative statics identifies the single-crossing property (see Milgrom and Shannon (1994)) as a sufficient (and, in a specific sense, necessary) condition for optimal choices to be monotone with respect to opponents' strategies and exogenous variables. The empirically relevant follow-up question is the following: What kind of *observed* choice behavior are necessary and sufficient for the recovery of payoff functions obeying the single-crossing property? The contribution of this paper is to answer this revealed preference question and to show that it forms the basis of an econometric analysis of games with strategic complements.

One obvious and important area of application of our results is to the study of entry games (as in Bresnahan and Reiss (1990), Berry (1992), or Ciliberto and Tamer (2009)) and other games that arise in the empirical IO literature. In these papers, firms' entry decisions are modeled as games of complete information, where each firm's decision on whether or not to enter a given market is a best response to the entry decisions taken by other firms in that market. The payoff functions are assumed to depend on observable variables in a specific parametric form while the unobserved component is additively separable. The unobserved component is heterogenous across markets and belongs to a known class of distributions. Entry decisions by firms across many markets are observed, from which one could then estimate firms' payoff functions. A major issue in this work concerns the effects of strategic interaction and market characteristics in terms of its direction and size: How often does the entry of another firm encourage or deter entry? To what extent does an exogenous variable (such as market size) encourage or deter the entry of other firms?

Our approach has as its starting point a data set of the same type as the papers cited above. With this data set, we can test whether firms are playing pure strategy Nash equilibria (PSNE), subject to single-crossing restrictions on its payoff functions. For example, we can test the hypothesis that a firm's entry into a market is encouraged when the market is large and discouraged when another firm is also entering. Our method works without imposing any parametric assumptions on payoff functions, without assuming that unobserved heterogeneity is additive or that its distribution belongs to a particular family, and without assumptions on equilibrium selection. By specifying a *joint* distribution on the payoff functions, we allow for correlation or other forms of dependence among firms' payoff functions, which is important in many settings (see Chen, Christensen, and Tamer (2018)). To pass our test means that the hypothesis that the data are explained as PSNE by firms with payoff functions satisfying single-crossing restrictions cannot be refuted.

At its most basic, our approach provides a way for researchers to test the general (nonparametric) features of a model, before the implementation of a more restrictive parametric model that could be used for inference and prediction. In some cases, the confirmation of monotone features, which are part of our test could also facilitate estimation procedures.[1] Beyond this, since our test recovers the distributions on firms' payoff functions that satisfy single-crossing restrictions and agree with the observations, the procedure can also be extended for the purposes of inference and prediction (when the data set passes the test).

While we write of recovering "payoff functions," what we are really recovering are a player's preference over different actions, conditional on covariates and the actions of other players; this is as it should be, because in an environment where only PSNE are played, the information recovered from the data *has to be* just ordinal. The specific preference property we test (or when making inferences, assume)—the single-crossing property—is also an ordinal property.

Our econometric approach is similar to that in Kitamura and Stoye (2018) (henceforth KS).[2] This paper tests a random utility model of consumer demand. In the first step, it is assumed that the population distribution of consumer demand at a linear budget set B, which we denote by $\mathrm{P}(\cdot|B)$, is known for a finite collection of budget sets $\mathcal{B}$. Then one could formulate necessary and sufficient conditions under which the stochastic demand system $\mathcal{P}_{\mathrm{KS}} = \{\mathrm{P}(\cdot|B)\}_{B\in\mathcal{B}}$ is generated by a population of utility-maximizing consumers, under the *conditional independence* assumption; this assumption requires the distribution of utility functions (which generates the distribution of demand) to be the same at each budget set $B \in \mathcal{B}$. The characterization of $\mathcal{P}_{\mathrm{KS}}$ in KS is facilitated by the well-known characterization of utility-maximizing demand behavior for a *single* consumer, known as the strong axiom of revealed preference (SARP). The second step in the KS approach is to show how the characterizing conditions on $\mathcal{P}_{\mathrm{KS}}$ could be statistically tested for an actual data set, with empirical frequencies estimated at each budget set $B \in \mathcal{B}$.

The key observation in our paper is that a two-step procedure similar to that implemented in KS could also be used for analyzing specific classes of games. Suppose that there is a large population of groups, with each group playing the same game. We assume that the population distribution over joint action profiles at a given vector of covariate values $\mathbf{x}$, which we denote by $\mathrm{P}(\cdot|\mathbf{x})$, is known for a finite set of covariate val-

[1] This information could be used to build a mapping from specific moments of the data to the identified set of relevant parameters. For instance, in two-player games the sign of the strategic interaction parameters allows us to identify outcomes that could occur *only as* a unique equilibrium; it follows that the probabilities of these outcomes (conditional on various observable variables) do not depend on any equilibrium selection mechanism and can be nicely related to payoff relevant parameters (see Tamer (2003) and Kline and Tamer (2016)). Shape restrictions can also reduce the size of the identified set of relevant parameters (see, e.g., Matzkin (2007)) and allow for the more efficient use of small sample data sets (see, e.g., Beresteanu (2005, 2007)) .

[2] Our approach is also close in spirit, though not in specifics, with the nonparametric random utility models in Tebaldi, Torgovitsky, and Yang (2023), Deb, Kitamura, Quah, and Stoye (2023), Apesteguia, Ballester, and Lu (2017), Hoderlein and Stoye (2014), Manski (2007), McFadden (2005), McFadden and Richter (1991), and Marschak (1960). As far as we know, our paper is the first to exploit this nonparametric approach to study games. Note that Kitamura and Stoye's empirical approach (and hence ours) is based on linear programming, which can be also found in earlier works such as Honoré and Tamer (2006) and Chernozhukov, Fernández-Val, Hahn, and Newey (2013).

ues $\widehat{\mathbf{X}}$.[3] (The set $\widehat{\mathbf{X}}$ takes the place of $\mathcal{B}$ in the KS model.) We then formulate necessary and sufficient conditions under which the set of choice distributions $\mathcal{P} = \{\mathrm{P}(\cdot|\mathbf{x})\}_{\mathbf{x}\in\widehat{\mathbf{X}}}$ is consistent with a population of groups made up of agents having payoff functions that satisfy single-crossing conditions and playing PSNE, under the assumption of conditional independence (which, in this case, means that the distribution of payoff function profiles across groups is the same at different $\mathbf{x} \in \widehat{\mathbf{X}}$). The second step in our approach shows how these conditions on $\mathcal{P}$ could be statistically tested on an actual data set, with empirical frequencies over action profiles at different covariate values; for this second step, we simply follow the statistical procedure in KS. As in KS, the sampling framework requires that for each $\mathbf{x} \in \widehat{\mathbf{X}}$, there are $N_{\mathbf{x}}$ observations of action profiles such that $N_{\mathbf{x}}/N \to \rho_{\mathbf{x}} \in (0, 1)$, where $N = \sum_{\mathbf{x}\in\widehat{\mathbf{X}}} N_{\mathbf{x}} \to \infty$.

Similar to KS, the characterization of $\mathcal{P} = \{\mathrm{P}(\cdot|\mathbf{x})\}_{\mathbf{x}\in\widehat{\mathbf{X}}}$ requires that we find necessary and sufficient conditions under which the joint actions from a *single* group at different covariate values are consistent with our hypothesis of PSNE play and payoff functions satisfying single-crossing conditions (with respect to opponents' actions and covariates). Since, unlike KS, there is no ready-made characterization for this class of games, we need to develop it ourselves. We show that this hypothesis can be characterized by a property we call the revealed monotonicity (RM) axiom. This axiom plays the role of SARP in the KS model.

When the data set passes the test, our approach is in turn useful for making inference and prediction in the spirit of Deb et al. (2023), which deals with a version of the consumer model. For example, we can estimate the fraction of players who are effectively nonstrategic, in the sense that their actions depend only on covariate values and are independent of what other players do. We can also bound the proportion of groups which (at a given covariate vector) has a particular equilibrium profile as a PSNE (along the lines of the analysis in Aradillas-Lopez (2011)); note that this potentially differs from the observed fraction of groups playing that action profile, not just because of sampling variation, but also because a given action profile could be a nonchosen PSNE when there are multiple PSNE.

The procedure in KS is hard to implement when there is a large number of budget sets and Smeulders, Cherchye, and De Rock (2021) propose a column generation method to deal with this difficulty. This method is also applicable in our setting and is useful in easing the computational burden of our test when (e.g.) $\widehat{\mathbf{X}}$ is a big set. In our paper, we develop a new result on column generation that allows for this method to be used, not just for testing but also inference.

The rest of the paper is organized as follows. In Section 2, we provide an outline of how our procedure works in the context of an entry game and contrast it with a parametric approach. Section 3 presents our main results at the population level. We introduce the revealed monotonicity axiom and use it to characterize those distributions over joint actions that are consistent with our hypothesis; properties of the underlying

[3]Variation of feasible sets (as in KS) can be included in our analysis of games (see Lazzati, Quah, and Shirai (2018)), but we have avoided it, in order not to burden the reader with too many model features and also because our empirical application does not have such variation. (See also Carvajal (2004) for a related result.)

TABLE 1. $\mathcal{P} = \{P(\cdot|x_2 = (0,0)), P(\cdot|x_2 = (0,1)), P(\cdot|x_2 = (1,0))\}$.

$x_2 = (0,0)$		Firm 2 N	Firm 2 E
Firm 1	N	3/12	3/12
	E	4/12	2/12

$x_2 = (0,1)$		Firm 2 N	Firm 2 E
Firm 1	N	1/12	5/12
	E	3/12	3/12

$x_2 = (1,0)$		Firm 2 N	Firm 2 E
Firm 1	N	2/12	4/12
	E	2/12	4/12

distribution over payoff function profiles can also be recovered. Section 4 explains how the population-level analysis in Section 3 can be implemented on finite sample data. In this section, we also introduce and extend the column generation method of Smeulders, Cherchye, and De Rock (2021). To illustrate our approach, we carry out an empirical analysis of entry decisions made by airlines; this is found in Section 5. The Supplemental Material (Lazzati, Quah, and Shirai (2024)) contains some additional theoretical/empirical results as well as the omitted details of the statistical procedure.

2. Motivating example

There is a large empirical literature modeling oligopoly entry decisions. We shall use this model to illustrate the basic question we are interested in and the approach we propose to address this question. For simplicity, we treat the case of two firms. Let $y_i \in \{N, E\}$ be the action set of firm i, where E means that the firm enters the market and N that it stays out and let x_i be a real-valued, finite-dimensional vector of exogenous profit shifters (covariates) that affect firm i's profit and are observed by the other firm and the researcher.

We assume that there is a large population of markets, with each market consisting of a Firm 1 and a Firm 2 that make their entry decisions simultaneously. The designation of a player as Firm 1 or Firm 2 is made by the researcher and based on observable characteristics; for example, in Kline and Tamer (2016), one firm is the "Low-Cost Carrier" and the other firm is "Other Airlines" (see Section 5). There is a finite set of realized profit shifters, which we denote by $\widehat{\mathbf{X}}$. For each $(x_1, x_2) \in \widehat{\mathbf{X}}$, we suppose that the population distribution of joint action profiles $P(\cdot|x_1, x_2)$ is known to the researcher. We denote this set of distributions by $\mathcal{P} = \{P(\cdot|(x_1, x_2))\}_{(x_1,x_2)\in\widehat{\mathbf{X}}}$. Table 1 gives an example of $\mathcal{P}$ where there is only variation in x_2 and it takes three possible vector values; for example, the box on the left tells us that $P((E, N)|x_2 = (0, 0)) = 4/12$. We are interested in developing a procedure, which allows us to identify those $\mathcal{P}$ that are compatible with our model of firm entry. Of course, in any empirical analysis these characterizing conditions on $\mathcal{P}$ would have to be statistically tested on an actual data set with sampling variation (as we explain in detail in Section 4). Confining our discussion to $\mathcal{P}$ at this stage allows us to focus on the more distinctive aspects of our analysis.

We now describe the model, which (potentially) generates $\mathcal{P}$. We denote the payoff/profit of Firms 1 and 2 by $\Pi_1(y_1, y_2, x_1)$ and $\Pi_2(y_1, y_2, x_2)$, respectively. We postulate that entry decisions are generated as pure strategy Nash equilibria (PSNE) of an entry game between Firms 1 and 2. We allow for multiple PSNE and impose no restriction on how firms select among these equilibria. There remains unobserved market heterogeneity even after conditioning on (x_1, x_2); this heterogeneity is captured by a joint distribution on (Π_1, Π_2), which in turn leads to a distribution over joint actions $\mathrm{P}(\cdot|x_1, x_2)$. We assume that there is *conditional independence*, in the sense that the distribution over (Π_1, Π_2) does not vary with the realized value of (x_1, x_2).

Lastly, we postulate that the firms' profit functions satisfy single-crossing restrictions (see Milgrom and Shannon (1994)). In this context, it means that Firm 1's entry into the market is encouraged when the profit shifter x_1 takes higher values and is discouraged when Firm 2 chooses to enter. Formally, we require

$$\Pi_1(E, y_2', x_1') > \Pi_1(N, y_2', x_1') \quad \Longrightarrow \quad \Pi_1(E, y_2'', x_1'') > \Pi_1(N, y_2'', x_1'') \tag{1}$$

whenever $x_1'' \geq x_1'$ and either $y_2' = y_2''$ or $y_2' = E$ and $y_2'' = N$. (A similar requirement is imposed on Π_2.) For example, in Ciliberto and Tamer (2009),

$$\Pi_1(y_1, y_2, x_1) = \begin{cases} \alpha_1' x_1 + \delta_1 \mathbf{1}_{y_2} + \varepsilon_1 & \text{if } y_1 = E, \\ 0 & \text{if } y_1 = N, \end{cases} \tag{2}$$

where $\mathbf{1}_E = 1$ and $\mathbf{1}_N = 0$. In this specification, the entry of Firm 2 alters the profit of Firm 1 by δ_1 and unobserved heterogeneity in payoff functions is captured by ε_1, which enters the profit function additively. It is straightforward to check that our single-crossing restrictions are satisfied if $\delta_1 < 0$ and $\alpha_1 > 0$. Note, however, that the converse is not true, that is, there are distributions over payoff functions satisfying (1) that *cannot* be represented in the additive form given by (2), for any distribution on ε_1.

We say that $\mathcal{P}$ is consistent with the single-crossing model, or $\mathcal{SC}$*-rationalizable*, if there is a joint distribution of payoff functions (Π_1, Π_2) that satisfy our single-crossing conditions (1) such that the resulting distribution of PSNE (given some equilibrium selection rule) coincides with $\mathrm{P}(\cdot|x_1, x_2)$ for each $x \in \widehat{\mathbf{X}}$. We would like to answer the following question: What conditions on $\mathcal{P}$ characterize $\mathcal{SC}$-rationalizability? In other words, when presented with $\mathcal{P}$, how could we check if it is $\mathcal{SC}$-rationalizable?

We first observe that our model *does* have structural implications for $\mathcal{P}$. Suppose the observable profit shifters weakly increase entry-by-entry from (x_1', x_2') to (x_1'', x_2'');[4] then, at any particular realization Π_1 of Firm 1's payoff function, if it prefers to enter when the other firm enters at (x_1', x_2'), then the single-crossing condition guarantees that it will continue to prefer entry at (x_1'', x_2''). The same argument applies to Firm 2, and so we conclude that if (E, E) is the Nash equilibrium at (x_1', x_2') for a given realized profit function profile (Π_1, Π_2), then it will be the *unique* Nash equilibrium at (x_1'', x_2'') for this realized profile. Aggregating across all profiles, we establish that

$$\mathrm{P}\big((E, E)|x_1'', x_2''\big) \geq \mathrm{P}\big((E, E)|x_1', x_2'\big),$$

provided conditional independence holds. This inequality constitutes a restriction on $\mathcal{P}$ but it is not the only restriction imposed by our model. We now sketch out the procedure

[4]Formally, (x_1'', x_2'') is weakly higher than (x_1', x_2') in the *product order* (see footnote 8 for its formal definition).

Table 2. Distribution of types rationalizing the choice distributions in table 1.

		$x_2 = (0, 0)$				$x_2 = (0, 1)$				$x_2 = (1, 0)$			
		Action Profiles				Action Profiles				Action Profiles			
Type	Weight	N, N	N, E	E, N	E, E	N, N	N, E	E, N	E, E	N, N	N, E	E, N	E, E
1	1/12			1/12					1/12			1/12	
2	2/12	2/12					2/12			2/12			
3	2/12			2/12				2/12					2/12
4	1/12			1/12				1/12				1/12	
5	1/12	1/12				1/12					1/12		
6	2/12				2/12				2/12				2/12
7	3/12		3/12				3/12				3/12		
Sum	1	3/12	3/12	4/12	2/12	1/12	5/12	3/12	3/12	2/12	4/12	2/12	4/12

for systematically checking whether $\mathcal{P}$ is $\mathcal{SC}$-rationalizable, using $\mathcal{P}$ presented in Table 1 as an example.

Given a particular realization (Π_1, Π_2), the firms will choose an action profile (either (E, E), (E, N), (N, E), or (N, N)) at each realization of x_2, and as x_2 takes different values the action profile of the two firms may change. We shall refer to the map from x_2 to the action profile as a *group type*. Notice that even though firms' profit functions may be heterogenous in infinitely many ways, its manifestation in behavior must be *finite*, since there are only finitely many possible actions and the realized covariates (x_1, x_2) take values in the finite set $\widehat{\mathbf{X}}$.

To be precise, there are in total $4^3 = 64$ group types, but not all are consistent with PSNE play and single-crossing payoff functions. For example, as we have already explained, a group type where (E, E) is played at $x_2 = (0, 0)$ and (N, N) at $x_2 = (0, 1)$ is not compatible with single-crossing. On the other hand, it is quite clear a group type where (N, E) is played at all three values of x_2 can be justified with single-crossing profit functions.

Ascertaining if $\mathcal{P}$ can be rationalized involves a two-step procedure. First, we must identify all *single-crossing group types*, in the sense that the action profile (y_1, y_2) at each value of (x_1, x_2) could be generated as PSNE from payoff functions satisfying (1). This is do-able because we show in Section 3 that these group types are characterized by an easy-to-check condition called the *revealed monotonicity axiom*. Second, we have to check whether there are weights on these group types that could account for the observed distribution of action profiles; this involves solving a system of linear inequalities.

We claim that $\mathcal{P}$ depicted in Table 1 can be rationalized. To understand why, we list in Table 2 seven possible group types. One could check that each of these group types is consistent with the single-crossing property. When these types are represented in the population with the weights indicated in Table 2, they generate the distribution of entry decisions observed in Table 1. (Compare the entries in Table 1 with the last row of Table 2.)

Lastly, we point out that while $\mathcal{P}$ is $\mathcal{SC}$-rationalizable, it is *not* compatible with a model where profit functions have the form (2), so the latter specification does involve a loss of generality. Indeed, with this specification, Firm 2's profit upon entry is

$$\pi_2(E, y_1, x_{21}, x_{22}, \varepsilon_2) = \alpha_{21}x_{21} + \alpha_{22}x_{22} + \delta_{21}\mathbf{1}_{y_1} + \varepsilon_2, \tag{3}$$

where $(\alpha_{21}, \alpha_{22}) > 0$ and $\delta_{21} < 0$.[5] Whether the boost to profits of an increase in x_{21} is greater or smaller than that obtained from the same increase in x_{22} depends on whether α_{21} is bigger or smaller than α_{22} and is *independent of the realization of* ε_2. So, it excludes the case where the realization of ε_2 influences the relative benefit of higher x_{21} versus higher x_{22}. To see why this parametric model cannot explain the choice distributions in Table 1, suppose instead that it does. Then

$$\begin{aligned}&\mathrm{P}\big((E, E)|x_1, (1, 0)\big) - \mathrm{P}\big((E, E)|x_1, (0, 0)\big)\\&\quad= \mu\big(\big\{\varepsilon_1 : \pi_1(E, E, x_1, \varepsilon_1) \geq 0\big\} \times \{\varepsilon_2 : -\delta_{21} \geq \varepsilon_2 \geq -\alpha_{21} - \delta_{21}\}\big),\end{aligned}$$

where μ is the probability measure on the space of $(\varepsilon_1, \varepsilon_2)$; similarly,

$$\begin{aligned}&\mathrm{P}\big((E, E)|x_1, (0, 1)\big) - \mathrm{P}\big((E, E)|x_1, (0, 0)\big)\\&\quad= \mu\big(\big\{\varepsilon_1 : \pi_1(E, E, x_1, \varepsilon_1) \geq 0\big\} \times \{\varepsilon_2 : -\delta_{21} \geq \varepsilon_2 \geq -\alpha_{22} - \delta_{21}\}\big).\end{aligned}$$

Since the former equals 2/12 while the latter equals 1/12, we conclude that $\alpha_{22} < \alpha_{21}$. However,

$$\begin{aligned}\frac{1}{12} &= \mathrm{P}\big((N, N)|x_1, (0, 0)\big) - \mathrm{P}\big((N, N)|x_1, (1, 0)\big)\\&= \mu\big(\big\{\varepsilon_1 : \pi_1(E, N, x_1, \varepsilon_1) \leq 0\big\} \times \{\varepsilon_2 : 0 \geq \varepsilon_2 \geq -\alpha_{21}\}\big)\end{aligned}$$

and

$$\begin{aligned}\frac{2}{12} &= \mathrm{P}\big((N, N)|x_1, (0, 0)\big) - \mathrm{P}\big((N, N)|x_1, (0, 1)\big)\\&= \mu\big(\big\{\varepsilon_1 : \pi_1(E, N, x_1, \varepsilon_1) \leq 0\big\} \times \{\varepsilon_2 : 0 \geq \varepsilon_2 \geq -\alpha_{22}\}\big),\end{aligned}$$

which tells us that $\alpha_{22} > \alpha_{21}$. So, we obtain a contradiction.

In Supplementary Appendix A1, we provide a more elaborate discussion of the contrast between the observable restrictions imposed by a linear parametric model and our (more general) nonparametric model. In particular, using simulations based on an extended version of the above example, we show that the difference between the two models is also picked up at the sample level: the method of Kline and Tamer (2016) (correctly) finds that the data are inconsistent with the linear model, whereas our method (also correctly) finds that the data are consistent with the more general model.

[5] We are grateful to Aureo De Paula for suggesting that we construct an example with this specific feature.

3. $\mathcal{SC}$-Rationalizable distributions

In this section, we consider a population of groups that play pure strategy Nash equilibria (PSNE) within each group. We characterize how the distribution of action profiles in this population will change with covariates when agents have best responses that are monotone with respect to both covariates and the actions of other agents in the group.

3.1 *Games with single-crossing payoff functions*

We assume that there is a population of groups, and for each group, we denote the set of agents by $\mathcal{N} = \{1, 2, \ldots, n\}$. Agent $i \in \mathcal{N}$ chooses an action y_i from an action space Y_i, which we assume is finite and totally ordered. We denote a joint action profile of the group by $\mathbf{y} \in \mathbf{Y} = \times_{i\in\mathcal{N}} Y_i$. For each $i \in \mathcal{N}$, there is an $M(i)$-dimensional covariate $x_i \in X_i = \times_{m=1}^{M(i)} X_{im} \subset \mathbb{R}^{M(i)}$.[6] The profile of $x_i = (x_{i1}, x_{i2}, \ldots, x_{iM(i)})$ across agents is denoted by $\mathbf{x} = (x_i : i \in \mathcal{N})$. The payoff of each agent depends on its action y_i, the actions of the other agents in its group $\mathbf{y}_{-i} = (y_j : j \in \mathcal{N}, j \neq i) \in \mathbf{Y}_{-i} = \times_{j\in\mathcal{N}, j\neq i} Y_j$, and on $x_i \in X_i$. Thus the payoff of agent i is given by a function $\Pi_i : Y_i \times \mathbf{Y}_{-i} \times X_i \to \mathbb{R}$. We use $\mathbf{\Pi} = (\Pi_i : i \in \mathcal{N})$ to indicate a profile of payoff functions.

A pair of payoff functions and covariate profiles $(\mathbf{\Pi}, \mathbf{x})$ induces a game of complete information $\mathrm{G}(\mathbf{\Pi}, \mathbf{x})$. In what follows, we let $\mathbf{X} \subset \times_{i\in\mathcal{N}} X_i$ denote the set of conceivable joint realizations of covariates $\mathbf{x}$. Since some covariates may be shared by multiple agents, $\mathbf{X}$ may not be equal to the direct product of X_i's (as in our empirical application in Section 5). We denote the best response of each player i at $(\mathbf{y}_{-i}, x_i)$ by $\mathrm{BR}_i(\mathbf{y}_{-i}, x_i) = \operatorname{argmax}_{y_i\in Y_i} \Pi_i(y_i, \mathbf{y}_{-i}, x_i)$; throughout this paper, we assume that agents have *strict* preferences over actions, so that $\mathrm{BR}_i(\mathbf{y}_{-i}, x_i)$ has a unique value.[7] The set of pure strategy Nash equilibria (PSNE) of this game is defined as

$$\mathrm{NE}(\mathbf{\Pi}, \mathbf{x}) = \left\{\mathbf{y}^* \in \mathbf{Y} : y_i^* = \mathrm{BR}_i\left(\mathbf{y}_{-i}^*, x_i\right) \text{ for all } i \in \mathcal{N}\right\}.$$

Importantly, even if the best response of every agent is single-valued, there could be multiple PSNE.

We are interested in games where payoff functions obey single-crossing conditions Milgrom and Shannon (1994).

Definition 1. The payoff function Π_i has *single-crossing differences in* $(y_i; (\mathbf{y}_{-i}, x_i))$ if the following condition holds: for every $y_i'' > y_i'$ and $(\mathbf{y}_{-i}'', x_i'') > (\mathbf{y}_{-i}', x_i')$ in the product order,[8]

$$\Pi_i\left(y_i'', \mathbf{y}_{-i}', x_i'\right) > \Pi_i\left(y_i', \mathbf{y}_{-i}', x_i'\right) \quad \Longrightarrow \quad \Pi_i\left(y_i'', \mathbf{y}_{-i}'', x_i''\right) > \Pi_i\left(y_i', \mathbf{y}_{-i}'', x_i''\right). \tag{4}$$

For simplicity, we may refer to such a payoff function as a *single-crossing payoff function*.

[6]We use "$\subset$" to denote the weak set inclusion and use "$\subsetneq$" to emphasize a proper subset relations.

[7]This assumption is needed for the revealed preference analysis to be meaningful, as otherwise, we could justify any behavior by simply claiming that each agent is indifferent among all elements in the action space.

[8]The product order on vectors is defined as follows: for vectors $a = (a_1, a_2, \ldots, a_n)$ and $a' = (a_1', a_2', \ldots, a_n')$, $a \geq a'$ means that $a_p \geq a_p'$ for all $p = 1, 2, \ldots, n$. We write $a > a'$ when $a \geq a'$ and $a \neq a'$; in other words, there is some p such that $a_p > a_p'$.

This condition states that if it is advantageous for agent i to choose a higher action y_i'' over a lower one y_i', then it remains advantageous to do so when other players raise their actions and/or covariates take higher values. The focus of our analysis is the set of payoff profiles

$$\mathcal{SC} = \{\mathbf{\Pi} = (\Pi_i)_{i \in \mathcal{N}} : \Pi_i \text{ is strict and has single-crossing differences in } (y_i; (\mathbf{y}_{-i}, x_i))\}.$$

Note that single-crossing differences is an *ordinal* property since any strictly increasing transformation of a function that obeys single-crossing differences will also obey single-crossing differences. Furthermore, since a player's best responses are pinned down by a player's *preference over actions*, $\text{NE}(\mathbf{\Pi}, \mathbf{x}) = \text{NE}(\widetilde{\mathbf{\Pi}}, \mathbf{x})$ whenever $\widetilde{\mathbf{\Pi}} = (\widetilde{\Pi}_i)_{i \in \mathcal{N}}$ is a strictly increasing transformation of $\mathbf{\Pi} = (\Pi_i)_{i \in \mathcal{N}}$, in the sense that, for every $i \in \mathcal{N}$, $\widetilde{\Pi}_i = f_i(\Pi_i)$ for some strictly increasing function $f_i : \mathbb{R} \to \mathbb{R}$.[9]

The property of single-crossing differences has two key implications, which are central to our study. (See Milgrom and Roberts (1990), Milgrom and Shannon (1994), and Vives (1990).)

Basic Theorem. *If* $\mathbf{\Pi} \in \mathcal{SC}$, *the family of games* $\{\text{G}(\mathbf{\Pi}, \mathbf{x}) : \mathbf{x} \in \mathbf{X}\}$ *has the following properties*:

(i) $\text{BR}_i(\mathbf{y}_{-i}, x_i)$ *is increasing in* $(\mathbf{y}_{-i}, x_i)$ *for each* $i \in \mathcal{N}$ *and*

(ii) $\text{NE}(\mathbf{\Pi}, \mathbf{x})$ *is nonempty.*[10]

This result says that single-crossing differences guarantees that $\text{G}(\mathbf{\Pi}, \mathbf{x})$ is a game of *strategic complements*, in the sense that a player optimally increases his action when other players raise theirs,[11] and that these games have pure strategy Nash equilibria. Furthermore, the best response of each player also increases with the (exogenous) covariate.

Remark 1. It is well known that, simply by reversing the order on the actions of one player, a two-player game of strategic substitutes becomes a game of strategic complements. The applications to strategic entry games in Sections 2 and 4 rely on that fact. However, one needs to proceed with care when applying our modeling framework to games of strategic substitutes with more than two players; while it is straightforward to work out single-crossing-type conditions on payoff functions that guarantee strategic substitutes, these games are not always guaranteed to have PSNE, which conflicts

[9]We say that a function f is increasing if $f(a) \geq f(a')$ whenever $a > a'$, while f is said to be *strictly* increasing if $f(a) > f(a')$ whenever $a > a'$. In this definition, the elements of the domain (and similarly the elements of the range) are ranked according to the product order as defined in footnote 8.

[10]It is also known that $\text{NE}(\mathbf{\Pi}, \mathbf{x})$ has a smallest and a largest element and that they increase with $\mathbf{x}$; however, this property is of limited use in our setting since we make no assumptions on equilibrium selection.

[11]There is also a sense in which single-crossing differences is *necessary* for monotone optimal solutions; see Milgrom and Shannon (1994).

with our hypothesis that the population consists of groups playing PSNE. One important class of games, which does accommodate strategic substitutes and for which the existence of PSNE is also guaranteed, is the class of aggregative games.[12]

3.2 *Rationalizability*

Since the population consists of many groups with heterogenous preferences, even at a given covariate value $\mathbf{x}$, different groups will take different joint actions. This generates a conditional distribution over joint actions $\mathbf{y} \in \mathbf{Y}$, which we denote by $\mathrm{P}(\cdot|\mathbf{x})$. Throughout this section, we assume that $\mathrm{P}(\cdot|\mathbf{x})$ is known for $\mathbf{x}'s$ contained in some *finite* subset $\widehat{\mathbf{X}} \subset \mathbf{X}$. In applications, it may be the case that $\widehat{\mathbf{X}} = \mathbf{X}$ but it is also possible for $\widehat{\mathbf{X}} \subsetneq \mathbf{X}$. The set $\mathbf{X}$ may or may not be finite, but it is important for our results that $\widehat{\mathbf{X}}$ is finite.[13] We wish to consider the conditions under which a set of choice distributions

$$\mathcal{P} = \{\mathrm{P}(\cdot|\mathbf{x}) : \mathbf{x} \in \widehat{\mathbf{X}}\}$$

is consistent with (in other words, generated by) pure strategy Nash equilibrium play in games with single-crossing payoff functions. Note that different choices across groups can arise not only from *heterogeneity in payoff functions* but also from *heterogeneity in equilibrium selection rules* among PSNE (both of which are not directly observed by the researcher).

Conditionally independent random payoff functions. To capture preference heterogeneity, we assume that the profile of payoff functions, $\mathbf{\Pi} = (\Pi_i)_{i\in\mathcal{N}}$, is random and distributed according to $\mathrm{P}_{\mathbf{\Pi}}$. (Notice that we are abusing notation by using $\mathbf{\Pi}$ to denote both the random variable and a particular realization.) By specifying a *joint* distribution on the payoff functions, we allow for correlation or any other type of dependence across the payoffs of the group members. In particular, we can be agnostic about correlations arising from group formation processes in the population (of the type observed in Chen, Christensen, and Tamer (2018), for example). Let $\mathrm{P}_{\mathbf{\Pi}}|_{\mathbf{x}}$ be the distribution of payoff function profiles conditional on the realized values of the covariates $\mathbf{x}$; we assume that $\mathrm{P}_{\mathbf{\Pi}}$ satisfies *conditional independence* in the sense that it does not depend on $\mathbf{x}$, that is, $\mathrm{P}_{\mathbf{\Pi}}|_{\mathbf{x}} = \mathrm{P}_{\mathbf{\Pi}}$ for all $\mathbf{x} \in \widehat{\mathbf{X}}$. The conditional independence assumption states that the distribution of group payoffs remains the same after conditioning on the observable covariates.

[12]In aggregative games, each player's payoff has the form $\Pi_i(y_i, \sum_{j\neq i} y_j, x_i)$, and one can test single-crossing differences either in $(y_i; (\sum_{j\neq i} y_j, x_i))$ (the case of strategic complements) or in $(y_i; (-\sum_{j\neq i} y_j, x_i))$ (the case of strategic substitutes). For these games, our framework *is* applicable because the existence of PSNE in aggregative games is guaranteed (see Dubey, Haimanko, and Zapechelnyuk (2006) and Jensen (2010)).

[13]In Kitamura and Stoye (2018) and Deb et al. (2023), the analog to $\widehat{\mathbf{X}}$ is the set of the price vectors at which the distribution of demand is known, whereas the analog to $\mathbf{X}$ is the set of all strictly positive price vectors; for essentially the same reasons, it is also important in their analyses that the former is finite, and for the statistical procedure, that it remains fixed as the sample size increases. Possible ways of relaxing this condition are discussed briefly in Kitamura and Stoye (2018, Section 8) and their observations are also potentially applicable here.

Loosely speaking, this assumption means that the population of groups A playing the game at some covariate value $\mathbf{x}'$ do not have fundamentally different preferences from the population of groups B who are playing the same game at another covariate value $\mathbf{x}''$: if the covariate imposed on population A is changed from $\mathbf{x}'$ to $\mathbf{x}''$, then the preferences of players over different actions in this population will be the same as those in population B (and vice versa).

Equilibrium selection rule. Given $\mathbf{x}$ and a particular realization $\mathbf{\Pi}$ in $\mathcal{SC}$, the basic theorem tells us that the set of pure strategy Nash equilibria $\text{NE}(\mathbf{\Pi}, \mathbf{x})$ is nonempty, and even though we assume that best replies are single-valued, we cannot rule out the possibility of multiple equilibria. We denote the *equilibrium selection rule* by $\lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x})$; this refers to the fraction of groups in the population with payoff functions $\mathbf{\Pi}$ and covariates $\mathbf{x}$ that select the action profile $\mathbf{y}$. We assume $\lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x}) = 0$ for all $\mathbf{y} \notin \text{NE}(\mathbf{\Pi}, \mathbf{x})$ and $\sum_{\mathbf{y}\in\mathbf{Y}} \lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x}) = 1$.

We are now in position to spell out precisely what it means for a set of distributions $\mathcal{P}$ to be consistent with PSNE in games with single-crossing payoff functions.

Definition 2. A distribution $\text{P}_{\mathbf{\Pi}}$, with support on $\mathcal{SC}$, *rationalizes* the set of choice distributions $\mathcal{P}$ if there is an equilibrium selection mechanism $\lambda(\cdot|\mathbf{\Pi}, \mathbf{x})$ such that

$$\text{P}(\mathbf{y}|\mathbf{x}) = \int \lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x})\, d\text{P}_{\mathbf{\Pi}} \quad \textit{for all } \mathbf{y} \in \mathbf{Y} \textit{ and all } \mathbf{x} \in \widehat{\mathbf{X}}. \tag{5}$$

$\mathcal{P}$ is *single-crossing rationalizable* (or *$\mathcal{SC}$-rationalizable*) if it admits such a distribution $\text{P}_{\mathbf{\Pi}}$; in other words, there is a distribution among payoff function profiles in $\mathcal{SC}$ and an equilibrium selection rule that could account for the observed distribution of joint actions at each $\mathbf{x} \in \widehat{\mathbf{X}}$.

Remark 2. $\mathcal{SC}$-rationalizability requires the domain of each agent's recovered payoff function to be $Y_i \times \mathbf{Y}_{-i} \times X_i$, rather than $Y_i \times \mathbf{Y}_{-i} \times \text{proj}_i\widehat{\mathbf{X}}$, and single-crossing differences must also be satisfied on the entire domain, that is, even for actions that are available but not chosen and covariate values that are not part of $\widehat{\mathbf{X}}$.[14]

Remark 3. We adopt conditional independence in this paper, because it is a tractable and quite prevalent restriction in empirical work. There are ways to weaken or modify this condition. If the modeler has a specific belief about the way that $\mathbf{\Pi}$ depends on the covariate $\mathbf{x}$ other than conditional independence, then it may be possible to replace conditional independence with the new condition and develop a characterization for this modified notion of $\mathcal{SC}$-rationalizability (see Remark 4), along with a statistical procedure to test it. Alternatively, in applications where conditional independence may be suspect, but instrumental variables are available, one may develop control variables for which the distribution of the payoff function profiles (after conditioning on the control variables) is independent of $\mathbf{x}$; with this one could calculate endogeneity-corrected distributions on actions, to which our results are applicable (see Kitamura and Stoye (2018)).

[14]Note that $\text{proj}_i\widehat{\mathbf{X}}$ is the projection of $\widehat{\mathbf{X}}$ to the set of possible values of x_i. That is, letting $\mathbf{x}_{-i}$ be a profile of covariates of agents other than i, $\text{proj}_i\widehat{\mathbf{X}} = \{x_i : (x_i, \mathbf{x}_{-i}) \in \widehat{\mathbf{X}} \text{ for some } \mathbf{x}_{-i}\}$.

3.3 *The revealed monotonicity axiom*

We now explain how $\mathcal{SC}$-rationalizable distributions may be characterized. The characterization has two parts and generalizes the procedure we used in the illustration in Section 2. First, we characterize all group types that could be generated by payoff functions with single-crossing differences. Second, we find weights on these types that could account for the distributions in $\mathcal{P}$.

I: Single-crossing rationalizable group types A *group type* is a function $\mathrm{B}: \widehat{\mathbf{X}} \to \mathbf{Y}$ that associates a profile of actions $\mathbf{y}$ to each covariate $\mathbf{x} \in \widehat{\mathbf{X}}$. For reasons which will be clear later, it is convenient to generalize this notion to correspondences. A *generalized group type* maps elements of $\widehat{\mathbf{X}}$ to nonempty subsets of $\mathbf{Y}$; we denote this correspondence by $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$, with $\mathrm{B}(\mathbf{x}) \subset \mathbf{Y}$ for each $\mathbf{x} \in \widehat{\mathbf{X}}$. We interpret a generalized group type as a set of observations generated by a group of players with fixed preferences, where $\mathrm{B}(\mathbf{x})$ consists of the action profiles that are played at the game with covariate $\mathbf{x}$. We wish to characterize all group types where $\mathrm{B}(\mathbf{x})$ consists of PSNE and players have payoff functions $\mathbf{\Pi}$ in $\mathcal{SC}$.

Definition 3. A generalized group type $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$ is a *single-crossing group type* if there exists a profile of payoff functions $\mathbf{\Pi}$ in $\mathcal{SC}$ such that $\mathrm{B}(\mathbf{x}) \subset \mathrm{NE}(\mathbf{\Pi}, \mathbf{x})$ for all $\mathbf{x} \in \widehat{\mathbf{X}}$.

The next definition provides the key observable feature of single-crossing group types.

Definition 4. A generalized group type $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$ obeys the *revealed monotonicity (RM) axiom*, if for each $\mathbf{x}', \mathbf{x}'' \in \widehat{\mathbf{X}}$,

$$\mathbf{y}' \in \mathrm{B}(\mathbf{x}'), \mathbf{y}'' \in \mathrm{B}(\mathbf{x}''), \text{ and } (\mathbf{y}''_{-i}, x''_i) \geq (\mathbf{y}'_{-i}, x'_i) \Longrightarrow y''_i \geq y'_i \text{ for each } i \in \mathcal{N}. \tag{6}$$

This axiom imposes a monotonicity restriction on B in the sense that it requires player i to take a weakly higher action whenever all other players are choosing higher actions and the covariate values are also higher. Note, however, that it does *not* require that if $\mathbf{x}'' \geq \mathbf{x}'$, $\mathbf{y}'' \in \mathrm{B}(\mathbf{x}'')$, and $\mathbf{y}' \in \mathrm{B}(\mathbf{x}') \Longrightarrow \mathbf{y}'' \geq \mathbf{y}'$. Indeed, the axiom even allows for $\mathbf{y}'' < \mathbf{y}'$, which corresponds to the case where there are two ranked Nash equilibria (at both $\mathbf{x}'$ and $\mathbf{x}''$), with players' jointly playing the lower equilibrium at the higher covariate value. The following theorem states that this axiom fully characterizes single-crossing group types.

Theorem 1. *The correspondence* $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$ *is a single-crossing group type if and only if it satisfies the RM axiom.*

We can think of Theorem 1 as a revealed preference counterpart to the basic theorem. Whereas that theorem tells us that whenever $\mathbf{\Pi} \in \mathcal{SC}$, players have monotone best response functions; this result says that one could rationalize a given group type with some $\Pi \in \mathcal{SC}$, as long as it displays no violations of monotonicity. Theorem 1 gives the econometrician, through the RM axiom, a simple way of checking whether or not a given group type is $\mathcal{SC}$-rationalizable.

It is clear that the RM axiom is *necessary* for $\mathrm{B}(\cdot)$ to be a single-crossing group type. Indeed, suppose $\mathrm{B}(\cdot)$ is of that type and for some $\mathbf{y}'' \in \mathrm{B}(\mathbf{x}'')$ and $\mathbf{y}' \in \mathrm{B}(\mathbf{x}')$, it holds that $(\mathbf{y}''_{-i}, x''_i) \geq (\mathbf{y}'_{-i}, x'_i)$ for some $i \in \mathcal{N}$. Then, for this agent i, there is some single-crossing payoff function Π_i for which $y''_i = \mathrm{BR}_i(\mathbf{y}''_{-i}, x''_i)$ and $y'_i = \mathrm{BR}_i(\mathbf{y}'_{-i}, x'_i)$ and basic theorem (i) immediately implies that $y''_i \geq y'_i$. Thus the more substantial part of this theorem is the claim that the RM axiom is sufficient for a group type to be single-crossing. In the proof (see the Supplemental Appendix), we explicitly construct, for each $i \in \mathcal{N}$, a single-crossing payoff function Π_i that supports y_i as the best response to $(\mathbf{y}_{-i}, x_i)$ for every $\mathbf{y} \in \mathrm{B}(\mathbf{x})$. This payoff function is defined on $Y_i \times \mathbf{Y}_{-i} \times X_i$ (rather than just $Y_i \times \mathbf{Y}_{-i} \times \mathrm{proj}_i\widehat{\mathbf{X}}$) and satisfies single-crossing differences on the entire domain; indeed, it satisfies increasing (and not just single-crossing) differences and is single-peaked in the action y_i.

To keep our exposition simple, we assume throughout this paper that each player's action is totally ordered. In fact, all our results remain valid even when each player's action space is multi-dimensional (and hence partially ordered), essentially because the RM axiom and Theorem 1 are both extendable to that context (see Supplementary Appendix Section A2).

II: Finding weights on group types Since the sets $\mathbf{Y}$ and $\widehat{\mathbf{X}}$ are finite, the set of all possible group types is also finite. We denote *the set of single-valued and single-crossing group types by $\mathcal{B}$*. The following result characterizes $\mathcal{SC}$-rationalizable choice distributions $\mathcal{P} = \{\mathrm{P}(\mathbf{y}|\mathbf{x}) : \mathbf{x} \in \widehat{\mathbf{X}}\}$ using the set of single-crossing group types $\mathcal{B}$ or, equivalently (by Theorem 1), those group types that obey the RM axiom.

Theorem 2. *$\mathcal{P}$ is $\mathcal{SC}$-rationalizable if and only if there exists a distribution $\tau = (\tau^{\mathrm{B}})_{\mathrm{B}\in\mathcal{B}}$ on $\mathcal{B}$ such that the following holds*:

$$\mathrm{P}(\mathbf{y}|\mathbf{x}) = \sum_{\{\mathrm{B}\in\mathcal{B}:B(\mathbf{x})=\mathbf{y}\}} \tau^{\mathrm{B}} \quad \textit{for all } \mathbf{y} \in \mathbf{Y} \textit{ and } \mathbf{x} \in \widehat{\mathbf{X}}. \tag{7}$$

The proof of the "only if" part of this theorem is found in the Supplemental Appendix. The "if" part of this theorem should be quite clear since the characterization is *itself* an instance of $\mathcal{SC}$-rationalizability. Indeed, suppose there is a distribution $\tau = (\tau^{\mathrm{B}})_{\mathrm{B}\in\mathcal{B}}$ on $\mathcal{B}$ such that (7) holds. By definition, there is some $\mathbf{\Pi} \in \mathcal{SC}$ that rationalizes B for each $\mathrm{B} \in \mathcal{B}$. By taking strictly increasing transformations if necessary, we can guarantee that distinct group types in $\mathcal{B}$ are rationalized by distinct payoff function profiles in $\mathcal{SC}$. We denote the profile that rationalizes B by $\mathbf{\Pi}^{\mathrm{B}}$. Then $\mathcal{P}$ is $\mathcal{SC}$-rationalizable with a distribution $\mathrm{P}_{\mathbf{\Pi}}$ that assigns probability τ^{B} to $\mathbf{\Pi}^{\mathrm{B}} \in \mathcal{SC}$ and an equilibrium selection rule λ where $\lambda(\mathbf{y}|\mathbf{\Pi}^{\mathrm{B}}, \mathbf{x}) = 1$ if $\mathbf{y} = \mathrm{B}(\mathbf{x})$ and $\lambda(\mathbf{y}|\mathbf{\Pi}^{\mathrm{B}}, \mathbf{x}) = 0$ if $\mathbf{y} \neq \mathrm{B}(\mathbf{x})$; in other words, all groups in the population with payoff profile $\mathbf{\Pi}^{\mathrm{B}}$ will play $\mathrm{B}(\mathbf{x})$ at each $\mathbf{x} \in \widehat{\mathbf{X}}$.

Theorems 1 and 2 together provide us with a way of establishing the $\mathcal{SC}$-rationalizability of $\mathcal{P}$. First, we must identify the single-crossing group types, which by Theorem 1, we can do via the RM axiom. Then Theorem 2 tells us that checking if $\mathcal{P}$ is

$\mathcal{SC}$-rationalizable boils down to finding a positive solution to a set of equations linear in the unknowns τ^{B} for all $\mathrm{B} \in \mathcal{B}$.[15]

Remark 4. It is part of the definition of $\mathcal{SC}$-rationalizability that the distribution of $\boldsymbol{\Pi} = (\Pi_i)_{i \in \mathcal{N}}$ is independent of $\mathbf{x}$. Suppose we drop this condition but still require all payoff functions to consist of single-crossing functions; then it is easy to see that the payoff functions and equilibrium selection rules will induce a distribution over group types in $\mathcal{B}$ at each $\mathbf{x}$, which we may denote by $(\tau_{\mathbf{x}}^{\mathrm{B}})_{\mathrm{B} \in \mathcal{B}}$, such that the following counterpart of (7) holds:

$$\mathrm{P}(\mathbf{y}|\mathbf{x}) = \sum_{\{\mathrm{B} \in \mathcal{B} : \mathrm{B}(\mathbf{x}) = \mathbf{y}\}} \tau_{\mathbf{x}}^{\mathrm{B}} \quad \text{for all } \mathbf{y} \in \mathbf{Y} \text{ and } \mathbf{x} \in \widehat{\mathbf{X}}. \tag{8}$$

This condition is trivially true in the sense that one could always find $(\tau_{\mathbf{x}}^{\mathrm{B}})_{\mathrm{B} \in \mathcal{B}}$ such that it holds. Conditional independence imposes the additional requirement that $\tau_{\mathbf{x}'}^{\mathrm{B}} = \tau_{\mathbf{x}''}^{\mathrm{B}}$ for any $\mathbf{x}', \mathbf{x}'' \in \widehat{\mathbf{X}}$, and this condition in combination with (8) is obviously equivalent to (7). One could imagine situations where the modeler has different views of how the distribution of $\boldsymbol{\Pi}$ (and hence the distribution of the associated group types) varies with $\mathbf{x}$, which may be more permissive than or different from conditional independence; these could be incorporated as further conditions on $\tau_{\mathbf{x}}^{\mathrm{B}}$ that could be tested in combination with (8). Obviously, such a test will remain a linear test if the added conditions are linear in $\tau_{\mathbf{x}}^{\mathrm{B}}$.

3.4 *Recovering properties of a rationalizing distribution* $\mathrm{P}_{\boldsymbol{\Pi}}$

When $\mathcal{P}$ is $\mathcal{SC}$-rationalizable, we are also able to extract information about this rationalization through the properties of $(\tau^{\mathrm{B}})_{\mathrm{B} \in \mathcal{B}}$ that solve (7). In particular, let $\mathcal{SC}^*$ be a subset of single-crossing payoff functions (including all of its strictly increasing transformations) and let

$$\mathcal{B}^* = \left\{\mathrm{B} \in \mathcal{B} : \text{ there is } \boldsymbol{\Pi} \in \mathcal{SC}^* \text{ that rationalizes B}\right\}. \tag{9}$$

By a straightforward adaptation of the proof of Theorem 2 (see the Supplemental Appendix), we can show that

$$\max\left\{\sum_{\mathrm{B} \in \mathcal{B}^*} \tau^{\mathrm{B}} : \left(\tau^{\mathrm{B}}\right)_{\mathrm{B} \in \mathcal{B}} \text{ solves (7)}\right\} = \max\left\{\int_{\boldsymbol{\Pi} \in \mathcal{SC}^*} d\mathrm{P}_{\Pi} : \mathrm{P}_{\Pi} \text{ rationalizes } \mathcal{P}\right\} \tag{10}$$

Notice that the left-hand side of this equation is straightforward to compute when $\mathcal{B}^*$ and $\mathcal{B}$ are known, since it simply involves solving a linear program. Thus we can find the greatest possible weight on a given set of payoff profiles, for any distribution that

[15] In some problems, it may not be computationally feasible to find all the elements of $\mathcal{B}$, but in those cases, one could still test for $\mathcal{SC}$-rationalizability by progressively enlarging the set of single-crossing types (see Section 4.1).

rationalizes $\mathcal{P}$.[16] We give two cases where this exercise is useful, both of which are empirically implemented in Section 5. Other examples can be found in the Supplementary Appendix A3.

Application 1. Bounds on the role of strategic interaction

While our model allows for the possibility that each player reacts strategically to other players in the game, it is conceivable that the conditional choice distributions could be explained more simply, without appealing to strategic effects for one or more players in the game.

To be specific, suppose we wish to check whether it is possible to regard a subgroup $\mathcal{N}'$ of the players as nonstrategic. Let $\mathcal{SC}^*$ be the payoff profiles in $\mathcal{SC}$ such that Π_i does not depend on $\mathbf{y}_{-i}$ for every $i \in \mathcal{N}'$ and let $\mathcal{B}^*$ be its corresponding set of group types (as defined by (9)). The types in $\mathcal{B}^*$ can be characterized by a stricter version of the RM axiom: a group type is in $\mathcal{B}^*$ if and only if it obeys the RM axiom and, for each $i \in \mathcal{N}'$, we require that $\mathbf{y}'' \in \mathrm{B}(\mathbf{x}''), \mathbf{y}' \in \mathrm{B}(\mathbf{x}'),$ and $x''_i \geq x'_i \Longrightarrow y''_i \geq y'_i$. With this characterization, we can construct $\mathcal{B}^*$. If we find that

$$\max\left\{\sum_{\mathrm{B}\in\mathcal{B}^*} \tau^{\mathrm{B}} : \left(\tau^{\mathrm{B}}\right)_{\mathrm{B}\in\mathcal{B}} \text{ solves (7)}\right\} = 1,$$

we conclude (by (10)) that $\mathcal{P}$ can be $\mathcal{SC}$-rationalized without requiring the players in $\mathcal{N}'$ to be strategic; on the other hand, if the upper bound is strictly below 1, then we *must* incorporate strategic interactions among these players to $\mathcal{SC}$-rationalize $\mathcal{P}$.

Application 2. Probability bounds for Nash equilibrium profiles

Given a strategy profile $\bar{\mathbf{y}}$ and covariate $\bar{\mathbf{x}}$, we pose the following question: Among all the possible $\mathcal{SC}$-rationalizations of $\mathcal{P}$, what is the greatest fraction of groups, which could have $\bar{\mathbf{y}}$ as a pure strategy Nash equilibrium at $\bar{\mathbf{x}}$? Here, $\bar{\mathbf{x}} \in \mathbf{X}$ may or may not be an element of $\widehat{\mathbf{X}}$, and when $\bar{\mathbf{x}} \notin \widehat{\mathbf{X}}$, the answer to this question provides information on how the game would be played at an hitherto unobserved covariate value. However, the question is interesting even when $\bar{\mathbf{x}} \in \widehat{\mathbf{X}}$.

To see why, notice that there is a distinction between $\mathrm{P}(\bar{\mathbf{y}}|\bar{\mathbf{x}})$, the observed fraction of groups in the population that play $\bar{\mathbf{y}}$ at $\bar{\mathbf{x}}$, and the fraction of groups for which $\bar{\mathbf{y}}$ is a Nash equilibrium. The former is typically smaller than the latter because groups might have multiple Nash equilibria. Thus some groups who play strategy profiles other than $\bar{\mathbf{y}}$ *may also have* $\bar{\mathbf{y}}$ *as a Nash equilibrium*.[17] The distinction between $\mathrm{P}(\bar{\mathbf{y}}|\bar{\mathbf{x}})$ and the greatest possible weight on those groups, which have $\bar{\mathbf{y}}$ as a Nash equilibrium at $\mathbf{x} = \bar{\mathbf{x}}$ is relevant,

[16] To obtain $\min\{\int_{\Pi\in\mathcal{SC}^*} d\mathrm{P}_\Pi : \mathrm{P}_\Pi \text{ rationalizes } \mathcal{P}\}$, we use the similarly easy-to-prove identity

$$\min\left\{\sum_{\mathrm{B}\in\mathcal{B}_0} \tau^{\mathrm{B}} : \left(\tau^{\mathrm{B}}\right)_{\mathrm{B}\in\mathcal{B}} \text{ solves (7)}\right\} = \min\left\{\int_{\Pi\in\mathcal{SC}^*} d\mathrm{P}_\Pi : \mathrm{P}_\Pi \text{ rationalizes } \mathcal{P}\right\},$$

where $\mathcal{B}_0 = \{\mathrm{B} \in \mathcal{B} : \mathrm{B} \text{ can only be rationalized by } \Pi \in \mathcal{SC}^*\}$.

[17] In our empirical application of an entry game with two firms, if (E, E) or (N, N) is played by a pair of firms, then it has to be their unique equilibrium, but any pair that plays (E, N) may also have (N, E) as another (albeit unselected) equilibrium. Thus if $\mathrm{P}(E, N|\bar{\mathbf{x}})$ and $\mathrm{P}(N, E|\bar{\mathbf{x}})$ are the observed probabilities of action profiles (E, N) and (N, E), respectively, then the probability that (E, N) (similarly, (N, E)) is a Nash equilibrium profile at $\mathbf{x} = \bar{\mathbf{x}}$ is no greater than $\mathrm{P}(E, N|\bar{\mathbf{x}}) + \mathrm{P}(N, E|\bar{\mathbf{x}})$.

because if the gap is small, then we are sure that changing the equilibrium selection scheme *cannot* significantly increase the frequency with which $\bar{\mathbf{y}}$ is played. This means (e.g.) that a policymaker who wants $\bar{\mathbf{y}}$ to be played more often must alter payoffs in some way and it is not possible to simply convince players to coordinate on a different equilibrium. An earlier analysis of questions of this type can be found in Aradillas-Lopez (2011), which focuses on a different class of games.

To answer our question, let $\mathcal{SC}^* = \{\Pi \in \mathcal{SC} : \bar{\mathbf{y}} \in \mathrm{NE}(\mathbf{\Pi}, \bar{\mathbf{x}})\}$ and let $\mathcal{B}^*$ be its corresponding set of group types. We can check whether B belongs to $\mathcal{B}^*$ by using the RM axiom. Indeed $\mathrm{B} \in \mathcal{B}^*$ if and only if the (possibly) multivalued group type $\overline{B}$ defined as follows obeys the RM-axiom: $\overline{\mathrm{B}}(\bar{\mathbf{x}}) = \{\mathrm{B}(\bar{\mathbf{x}}), \bar{\mathbf{y}}\}$ and $\overline{\mathrm{B}}(\mathbf{x}) = \mathrm{B}(\mathbf{x})$ for every $\mathbf{x} \in \widehat{\mathbf{X}} \setminus \{\bar{\mathbf{x}}\}$. The proportion of the population which has $\bar{\mathbf{y}}$ as a PSNE cannot exceed $\max\{\sum_{\mathrm{B}\in\mathcal{B}^*} \tau^{\mathrm{B}} : (\tau^{\mathrm{B}})_{\mathrm{B}\in\mathcal{B}} \text{ solves } (7)\}$ and can equal this number.[18]

4. The statistical procedure

This section outlines the statistical procedure that implements the results in the previous section, which are based on population distributions. The test of $\mathcal{SC}$-rationalizability is explained in Section 4.1 and relies on the statistical hypothesis testing proposed by Kitamura and Stoye (2018). The efficient implementation of this test when $\mathcal{B}$ is large (and cannot be fully listed) uses the column generation approach proposed in Smeulders, Cherchye, and De Rock (2021). Section 4.2 outlines the procedure (in essence provided by Deb et al. (2023)) to obtain confidence intervals for the weights on certain group types; the efficient implementation of this procedure requires a nontrivial extension of the column generation method in Smeulders, Cherchye, and De Rock (2021) and we provide this in Proposition 3.

4.1 *Statistical hypothesis testing*

We begin with a matrix reformulation of the characterization given in Theorem 2. Each generalized group type $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$ can be represented as a vector $\mathbf{b} = (b_{\mathbf{y},\mathbf{x}})_{\mathbf{Y}\times\widehat{\mathbf{X}}}$ such that $b_{\mathbf{y},\mathbf{x}} = 1$ if $\mathbf{y} \in \mathrm{B}(\mathbf{x})$ and $b_{\mathbf{y},\mathbf{x}} = 0$ otherwise. Conversely, for any $\mathbf{b} \in \{0, 1\}^{|\mathbf{Y}\times\widehat{\mathbf{X}}|}$ corresponds to a generalized group type, with a vector $\mathbf{b} \in \{0, 1\}^{|\mathbf{Y}\times\widehat{\mathbf{X}}|}$ representing a single-valued group type if and only if $\sum_{\mathbf{y}\in\mathbf{Y}} b_{\mathbf{y},\mathbf{x}} = 1$ at every $\mathbf{x} \in \widehat{\mathbf{X}}$. Similarly, since $\mathcal{P}$ consists of $|\widehat{\mathbf{X}}|$ distributions on $\mathbf{Y}$, it can be captured by the column vector $\mathbf{p} \in [0, 1]^{|\mathbf{Y}\times\widehat{\mathbf{X}}|}$, where the $(\mathbf{y}, \mathbf{x})$-th entry of $\mathbf{p}$ is $\mathrm{P}(\mathbf{y}|\mathbf{x})$ (and hence, $\sum_{\mathbf{y}\in\mathbf{Y}} p_{\mathbf{y},\mathbf{x}} = 1$ for each $\mathbf{x} \in \widehat{\mathbf{X}}$).

In what follows, we shall abuse notation and use $\mathcal{B}$ to denote both the set of group types obeying the RM axiom and also the vectors corresponding to those types. We denote by $\mathbf{B}$ the matrix where each column represents a group type in $\mathcal{B}$. Theorem 2 states that $\mathcal{P}$ is $\mathcal{SC}$-rationalizable if and only if there is $\tau \in \Delta^{\mathcal{B}}$, the set of distributions on $\mathcal{B}$, that

[18] Our analysis here gives the most optimistic estimate on the possibility of switching the equilibrium action to $\bar{\mathbf{y}}$, in the sense that it assumes that every group type, which *can* be rationalized by an element in $\mathcal{SC}^*$, actually does have a payoff function profile in $\mathcal{SC}^*$. We could also find the most conservative estimate of the proportion of the population that could switch to $\bar{\mathbf{y}}$ by changing equilibrium selection rules; this is explained in Supplementary Appendix A6.2.

solves $\mathbf{B}\tau = \mathbf{p}$. ($\Delta^{\mathcal{B}}$ could be thought of as elements of the standard $(|\mathcal{B}| - 1)$-simplex.) We would like to test if the data is consistent with the $\mathcal{SC}$-rationalizability of $\mathcal{P}$. Equivalently, letting $\mathbb{P}^{\text{SC}} = \{\mathbf{B}\tau : \tau \in \Delta^{\mathcal{B}}\}$ (i.e., the set of $\mathcal{SC}$-rationalizable distributions in vector form), our null hypothesis is

$$\min_{\eta \in \mathbb{P}^{\text{SC}}} (\mathbf{p} - \eta) \cdot (\mathbf{p} - \eta) = 0. \tag{11}$$

The data set consists of $N_{\mathbf{x}}$ observations of the action profiles at each realization of $\mathbf{x} \in \widehat{\mathbf{X}}$. We assume that $N_{\mathbf{x}}/N \to \rho_{\mathbf{x}} \in (0, 1)$ at each $\mathbf{x} \in \widehat{\mathbf{X}}$, as $N = \sum_{\mathbf{x} \in \widehat{\mathbf{X}}} N_{\mathbf{x}} \to \infty$. We denote the empirical distribution over action profiles by

$$\mathcal{Q} = \{\mathrm{Q}(\cdot|\mathbf{x}) : \mathbf{x} \in \widehat{\mathbf{X}}\},$$

and we estimate $\mathcal{P}$ by this sample analog. As in the case with $\mathcal{P}$, we can represent $\mathcal{Q}$ by a column vector $\mathbf{q} \in [\mathbf{0}, \mathbf{1}]^{|\mathbf{Y} \times \mathbf{X}|}$ where the $(\mathbf{y}, \mathbf{x})$-th entry is equal to $\mathrm{Q}(\mathbf{y}|\mathbf{x})$.

The testing procedure by Kitamura and Stoye (2018) depends on the simple, but important observation that $\mathbf{B}\tau = \mathbf{p}$ holds for some $\tau \in \Delta^{\mathcal{B}}$, if and only if $\mathbf{B}\tau = \mathbf{p}$ holds for some $\tau \geq 0$ (Theorem 3.1 in their paper). Thus, by letting $\mathcal{A} = \{\mathbf{B}\tau : \tau \geq 0\}$, the null hypothesis is equivalent to whether $\mathbf{p}$ lives in this convex cone, that is,

$$\min_{\eta \in \mathcal{A}} (\mathbf{p} - \eta) \cdot (\mathbf{p} - \eta) = 0. \tag{12}$$

Given this, following Kitamura and Stoye (2018), we adopt the test statistic

$$J_N := \min_{\eta \in \mathcal{A}} N(\mathbf{q} - \eta) \cdot (\mathbf{q} - \eta) = \min_{\tau \in \mathbb{R}_+^{|\mathcal{B}|}} N(\mathbf{q} - \mathbf{B}\tau) \cdot (\mathbf{q} - \mathbf{B}\tau). \tag{13}$$

Calculating the critical value. Note that we cannot simply adopt a solution to the problem (13) as the bootstrap estimator for the empirical choice distribution, due to the possible discontinuity of the limiting distribution of J_N. Addressing this issue involves introducing a tuning parameter and considering the corresponding *tightened* problem. We follow the procedure by Smeulders, Cherchye, and De Rock (2021), which is a modification of the one in Kitamura and Stoye (2018).

Choose $\mathcal{B}' \subset \mathcal{B}$ so that it contains a basis of the space spanned by $\mathcal{B}$, and define $\mathcal{T}_{\kappa_N} = \{\tau \in \mathbb{R}_+^{|\mathcal{B}|} : \tau^{\mathbf{b}} \geq \kappa_N / |\mathcal{B}'| \text{ for all } \mathbf{b} \in \mathcal{B}'\}$, with κ_N being selected so that $\kappa_N \downarrow 0$ and $\sqrt{N}\kappa_N \uparrow \infty$ as $N \to \infty$. (See Supplementary Appendix A5, for the procedure for constructing $\mathcal{B}'$ and a detailed justification of our procedure.[19]) Letting $\mathcal{A}_{\kappa_N} = \{\mathbf{B}\tau : \tau \in \mathcal{T}_{\kappa_N}^{\mathcal{B}'}\}$, we adopt

$$\eta^* = \operatorname*{argmin}_{\eta \in \mathcal{A}_{\kappa_N}} N(\mathbf{q} - \eta) \cdot (\mathbf{q} - \eta) = \operatorname*{argmin}_{\tau \in \mathcal{T}_{\kappa_N}^{\mathcal{B}'}} N(\mathbf{q} - \mathbf{B}\tau) \cdot (\mathbf{q} - \mathbf{B}\tau). \tag{14}$$

[19] In the original formulation by Kitamura and Stoye (2018), positive weights are required on all elements in $\mathcal{B}$, which is inconvenient when applying the column generation procedure described later in this subsection. The modification of that approach by Smeulders, Cherchye, and De Rock (2021) (which we are using here) requires positive weights only for the types in $\mathcal{B}'$.

as the bootstrap estimator of the empirical choice frequency. Compared to the problem (13), the feasible set in the minimization problem is tightened by the tuning parameter, with positive weights required for the elements in $\mathcal{B}'$. We then generate a bootstrap sample $\mathbf{q}^{(r)}$ (for $r = 1, 2, \dots, R$) using standard nonparametric bootstrap resampling from η^* and recenter this sample by setting $\widehat{\mathbf{q}}^{(r)} := (\mathbf{q}^{(r)} - \mathbf{q}) + \eta^*$. With $\widehat{\mathbf{q}}^{(r)}$, we can calculate the bootstrap test statistic

$$J_N^{(r)} := \min_{\eta \in \mathcal{A}_{\kappa_N}} N\big(\widehat{\mathbf{q}}^{(r)} - \eta\big) \cdot \big(\widehat{\mathbf{q}}^{(r)} - \eta\big) = \min_{\tau \in \mathcal{T}_{\kappa_N}^{\mathcal{B}'}} N\big(\widehat{\mathbf{q}}^{(r)} - \mathbf{B}\tau\big) \cdot \big(\widehat{\mathbf{q}}^{(r)} - \mathbf{B}\tau\big), \tag{15}$$

and the empirical distribution of $J_N^{(r)}$ allows us to obtain the p-value $p = \#\{J_N^{(r)} > J_N\}/R$. The null hypothesis that $\mathbf{q}$ is a sample from some $\mathbf{p} \in \mathcal{A}$ (equivalently, $\mathbf{p} \in \mathbb{P}^{\mathrm{SC}}$) is not rejected if the p-value is greater than the critical value.

Column Generation. A major hurdle in implementing the above test is that the computation of J_N and $J_N^{(r)}$ involves $\mathcal{B}$, which is often too large to be listed in its entirety. We cope with this problem by applying the *column generation* procedure in Smeulders, Cherchye, and De Rock (2021). This procedure involves first testing a more stringent version of the model corresponding to a strict subset $\mathcal{B}_0$ of $\mathcal{B}$, which is completely known. For instance, we may choose the "starter" set $\mathcal{B}_0$ to be the set of *constant types*, in which every player takes the same action regardless of opponents' actions and covariates; these group types obviously obey the RM axiom. Then the set $\mathcal{B}_0$ is progressively enlarged by including more group types from $\mathcal{B}$, up to the point where further additions will not improve the model's ability to explain the data.

To be precise, let $\mathbf{B}_0$ be the matrix where the columns are elements of $\mathcal{B}_0$. We can calculate

$$J_{N,0} := \min_{\tau \in \mathbb{R}_+^{|\mathcal{B}_0|}} N(\mathbf{q} - \mathbf{B}_0\tau) \cdot (\mathbf{q} - \mathbf{B}_0\tau). \tag{16}$$

Obviously, $J_{N,0} \geq J_N$ and we could check if it is possible to decrease $J_{N,0}$ by including some $\mathbf{b} \in \mathcal{B}$. We say that $\mathbf{b} \in \mathcal{B}$ *improves* $\mathcal{B}_0$ if, when $\mathbf{b}$ is included in $\mathcal{B}_0$, the new value of $J_{N,0}$ is strictly lower. The following result, which follows from the convex projection theorem, provides a necessary and sufficient condition for $\mathcal{B}_0$ to be improvable.

Proposition 1. *A set of group types $\mathcal{B}_0$ is improved by some $\mathbf{b} \in \mathcal{B}$, if and only if*

$$\max_{\mathbf{b} \in \mathcal{B}} (\mathbf{q} - \eta_0) \cdot (\mathbf{b} - \eta_0) > 0, \tag{17}$$

where $\eta_0 = \mathbf{B}_0\tau_0$ and $\tau_0 = \operatorname{argmin}_{\tau \in \mathbb{R}_+^{|\mathcal{B}_0|}} (\mathbf{q} - \mathbf{B}_0\tau) \cdot (\mathbf{q} - \mathbf{B}_0\tau)$.

To solve problem (17) without fully enumerating $\mathcal{B}$, we must find a computationally efficient way to characterize $\mathcal{B}$. Conveniently for us, the RM axiom—and hence the set $\mathcal{B}$—can be characterized as solutions to an integer linear programming problem.

Proposition 2. *We can construct a matrix C and a column vector θ, both with nonnegative integer entries, such that for any $\mathbf{b} \in \{0, 1\}^{|\mathbf{Y} \times \widehat{\mathbf{X}}|}$, we have $\mathbf{b} \in \mathcal{B}$ if and only if $C\mathbf{b} \leq \theta$.*

The formulae for C and θ are found in our proof of this proposition in the Supplemetnal Appendix. Combining this result with Proposition 1, $\mathcal{B}_0$ is improved by some $\mathbf{b} \in \mathcal{B}$, if and only if

$$\max(\mathbf{q} - \eta_0) \cdot (\mathbf{b} - \eta_0), \quad \text{subject to } \mathbf{b} \in \{0, 1\}^{|\mathbf{Y} \times \widehat{\mathbf{X}}|} \text{ and } C\mathbf{b} \leq \theta, \tag{18}$$

is strictly positive. If it is, we add this $\mathbf{b}$ to $\mathcal{B}_0$ and then repeat the process. In other words, we recalculate $J_{N,0}$ and η_0 based on the new $\mathcal{B}_0$, and try to find another element in $\mathcal{B}$ that improves $\mathcal{B}_0$ by checking if (18) has a strictly positive solution. Since $\mathcal{B}$ is finite, this algorithm must terminate, and at the end we can be sure we have found $\mathcal{B}_0$ such that $J_{N,0} = J_N$.

The column generation procedure described above can be also applied to the computation of $J_N^{(r)}$ defined by (15). Since the constraint in problem (15) requires positive weights on $\mathcal{B}'$, this set needs to be contained in the initial choice of $\mathcal{B}_0$.[20]

Summary. Below is a step-by-step summary of the test procedure.[21]

I Obtain the test statistic J_N defined in (13) as follows:

(i) Based on $\mathcal{B}_0$, solve the minimization problem (16) to get $J_{N,0}$ and η_0.

(ii) Check the value of (18). IF it is strictly positive, then update $\mathcal{B}_0$ by adding a solution of (18) and go to (i). ELSE, set $J_N = J_{N,0}$ and STOP.

II Obtain the tightened estimator η^* in (14) as follows:

(i) Obtain $\mathcal{B}' \subset \mathcal{B}$ using the procedure explained in Supplementary Appendix A5.2.

(ii) Set $\mathcal{B}'$ as $\mathcal{B}_0$, and run the procedure in Step I replacing $\mathbb{R}_+^{|\mathcal{B}_0|}$ in problem (16) with $\mathcal{T}_{\kappa_N}^{\mathcal{B}_0} = \{\tau \in \mathbb{R}_+^{|\mathcal{B}_0|} : \tau^{\mathbf{b}} \geq \kappa_N / |\mathcal{B}'| \text{ for } \mathbf{b} \in \mathcal{B}'\}$. When it stops, the resulting η_0 is η^*.

III Obtain the bootstrap test statistics $J_N^{(r)}$ defined in (15) for $r = 1, 2, \ldots, R$:

(i) Obtain the recentered bootstrap sample $\widehat{\mathbf{q}}^{(r)} = (\mathbf{q}^{(r)} - \mathbf{q}) + \eta^*$.

(ii) Set $\mathcal{B}'$ as $\mathcal{B}_0$, and run the procedure in Step I replacing $\mathbf{q}$ and $\mathbb{R}_+^{|\mathcal{B}_0|}$ in problem (16) with $\widehat{\mathbf{q}}^{(r)}$ and $\mathcal{T}_{\kappa_N}^{\mathcal{B}_0}$, respectively. When it stops, adopt the resulting $J_{N,0}$ as $J_N^{(r)}$.

IV Lastly, calculate the p-value $p = \#\{J_N^{(r)} > J_N\}/R$.

[20] Note that, even if $\mathcal{B}_0$ contains $\mathcal{B}'$, and $\mathcal{B}'$ contains a linear basis of $\mathcal{B}$, the *conical hull* of $\mathcal{B}_0$ need not coincide with the conical hull of $\mathcal{B}$ (though the linear hull of $\mathcal{B}_0$ of course coincides with the linear hull of $\mathcal{B}$). So, it is still possible for $\mathcal{B}_0$ to be improvable.

[21] The Supplementary Appendix (Section A5.2) provides a couple of shortcuts that improves the computation time.

4.2 *Inference on types*

Suppose we have a data set that is consistent with $\mathcal{SC}$-rationalizability (in the sense that the null hypothesis (12) is not rejected) and would now like to form a confidence interval on $\sum_{\mathbf{b}\in\mathcal{B}^*}\tau^{\mathbf{b}}$, the total weight of a subset of single-crossing group types $\mathcal{B}^*$ (see Section 3.4). To do this, we follow the procedure in Deb et al. (2023). The problem of determining whether a given weight of $\mathcal{B}^*$ falls within the confidence interval can be determined by testing a suitably modified version of the null hypothesis (12), with $\mathbb{P}^{\text{SC}}$ replaced by a different set of distributions. To be specific, suppose we would like to find the upper bound of the confidence interval. For each $\beta \in (0, 1)$, we let

$$\mathbb{P}^{SC}(\beta; \mathcal{B}^*) = \left\{\mathbf{B}\tau : \tau \in \Delta^{\mathcal{B}} \text{ and } \sum_{\mathbf{b}\in\mathcal{B}^*} \tau^{\mathbf{b}} \geq \beta\right\},$$

and test the null hypothesis

$$\min_{\eta\in\mathbb{P}^{\text{SC}}(\beta;\mathcal{B}^*)} (\mathbf{p}-\eta)\cdot(\mathbf{p}-\eta) = 0 \tag{19}$$

at some significance level $\bar{p}$. We then use binary search to obtain the maximal value of β under which the null hypothesis is not rejected; the resulting maximal value of β corresponds to the supremum of the $100(1-\bar{p})\%$ confidence interval of β.

For a given β, the test statistic is[22]

$$\begin{aligned} J_N(\beta) &:= \min_{\eta\in\mathbb{P}^{\text{SC}}(\beta;\mathcal{B}^*)} N(\mathbf{q}-\eta)\cdot(\mathbf{q}-\eta) \\ &= \min_{\tau\in\Delta^{\mathcal{B}}} N(\mathbf{q}-\mathbf{B}\tau)\cdot(\mathbf{q}-\mathbf{B}\tau) \quad \text{subject to} \sum_{\mathbf{b}\in\mathcal{B}^*}\tau^{\mathbf{b}} \geq \beta. \end{aligned} \tag{20}$$

As in the preceding subsection, it may not be possible to fully enumerate $\mathcal{B}$ or $\mathcal{B}^*$, and so a version of the column generation procedure outlined there is needed. This in turn requires an extension of Proposition 1, which we now explain.

Let $\mathcal{B}_0 \subset \mathcal{B}$ be such that $\mathcal{B}_0 \cap \mathcal{B}^* \neq \emptyset$, and let us calculate

$$J_{N,0}(\beta) = \min_{\tau\in\Delta^{\mathcal{B}_0}} N(\mathbf{q}-\mathbf{B}_0\tau)\cdot(\mathbf{q}-\mathbf{B}_0\tau) \quad \text{s.t.} \sum_{\mathbf{b}\in(\mathcal{B}_0\cap\mathcal{B}^*)} \tau^{\mathbf{b}} \geq \beta, \tag{21}$$

where $\Delta^{\mathcal{B}_0}$ is the standard $(|\mathcal{B}_0|-1)$-simplex. We say that $\mathcal{B}_0$ is *improvable given problem (20)*, if $J_{N,0}(\beta) > J_N(\beta)$. The following proposition is the counterpart of Proposition 1 and provides a necessary and sufficient condition for a given $\mathcal{B}_0$ to be improvable.

Proposition 3. *If the set $\mathcal{B}_0 \subset \mathcal{B}$ is improvable given problem* (20), *then there is a pair of types $\{\mathbf{b}^*, \mathbf{b}\}$, with $\mathbf{b}^* \in \mathcal{B}^*$ and $\mathbf{b} \in \mathcal{B}$ such that*

$$(\mathbf{q}-\eta_0)\cdot\big(\beta\mathbf{b}^* + (1-\beta)\mathbf{b} - \eta_0\big) > 0, \tag{22}$$

[22] As pointed out in Deb et al. (2023), unlike the case of the preceding subsection (see (13)), τ must be chosen from the simplex $\Delta^{\mathcal{B}}$, rather than the nonnegative orthant.

where $\eta_0 = \mathbf{B}_0\tau_0$ *with* τ_0 *being the distribution that achieves* $J_{N,0}(\beta)$. *Conversely, suppose there is* $\mathbf{b}^* \in \mathcal{B}^*$ *and* $\mathbf{b} \in \mathcal{B}$ *such that* (22) *holds*; *then* $\{\mathbf{b}^*, \mathbf{b}\}$ *improves* $\mathcal{B}_0$ *given problem* (20).

We already know (from Proposition 2) that we can construct a matrix C and a column vector θ so that $\mathbf{b} \in \mathcal{B}$ if and only if $C\mathbf{b} \le \theta$. Suppose that, in addition, we can construct a matrix C^* and a column vector θ^* with integer entries so that, for any $\mathbf{b} \in \{0, 1\}^{|\mathbf{Y}\times\widehat{\mathbf{X}}|}$, we have $\mathbf{b}^* \in \mathcal{B}^* \iff C^*\mathbf{b}^* \le \theta^*$. Then a pair $\{\mathbf{b}^*, \mathbf{b}\}$ obeying (22) exists, if and only if the problem

$$\max(\mathbf{q} - \eta_0) \cdot \big(\beta\mathbf{b}^* + (1-\beta)\mathbf{b} - \eta_0\big) \tag{23}$$
$$\text{s.t. } \mathbf{b}, \mathbf{b}^* \in \{0, 1\}^{|\mathbf{Y}\times\widehat{\mathbf{X}}|} \text{ and } \begin{pmatrix} C^* & O \\ O & C \end{pmatrix} \begin{pmatrix} \mathbf{b}^* \\ \mathbf{b} \end{pmatrix} \le \begin{pmatrix} \theta^* \\ \theta \end{pmatrix}$$

has a positive optimal value. Note that every $\mathcal{B}^*$ in our empirical application has a matrix characterization like the one described above (see Supplementary Appendix A4 for the specific construction). If there is a pair $\{\mathbf{b}, \mathbf{b}^*\}$ that improves $\mathcal{B}_0$, then we update $\mathcal{B}_0$ by including the pair in $\mathcal{B}_0$ and recalculate $J_{N,0}(\beta)$. Since $\mathcal{B}$ is finite, this process terminates and we obtain $J_{N,0}(\beta) = J_N(\beta)$.

To find the valid critical value, we need a suitable tightening that imposes strictly positive weights on a certain subset of group types. The tightening here must depend on β and its formulation is rather involved, so we postpone this discussion to Supplementary Appendix A5.3. That said, once we have constructed a suitably tightened subset of $\Delta^{\mathcal{B}}$ by some tuning parameter κ_N, the rest of the procedure is similar to the one outlined in the preceding subsection. Denoting this subset by $\Delta^{\mathcal{B}}_{\kappa_N}(\beta; \mathcal{B}^*)$ and letting

$$\mathbb{P}^{\mathrm{SC}}_{\kappa_N}(\beta; \mathcal{B}^*) = \Big\{\mathbf{B}\tau : \tau \in \Delta^{\mathcal{B}}_{\kappa_N}(\beta; \mathcal{B}^*) \text{ and } \sum_{\mathbf{b}\in\mathcal{B}^*} \tau^{\mathbf{b}} \ge \beta\Big\}, \tag{24}$$

the bootstrap estimator and the recentered bootstrap samples can be obtained as in (14)–(15), after replacing the sets $\mathcal{A}_{\kappa_N}$ and $\mathcal{T}_{\kappa_N}$ by $\mathbb{P}^{\mathrm{SC}}_{\kappa_N}(\beta; \mathcal{B}^*)$ and $\Delta^{\mathcal{B}}_{\kappa_N}(\beta; \mathcal{B}^*)$, respectively. We can also implement here the column generation procedure we used before. The details of the procedure, including a step-by-step summary, are provided in Supplementary Appendix A5.3.

5. Empirical illustration

We apply our results in the preceding sections to an entry game using a data set taken from Kline and Tamer (2016). The data set contains the entry decisions of airlines in 7882 markets, where a market is defined as a trip between two airports irrespective of intermediate stops. Airline firms are divided into two categories: LCC (low cost carriers) and OA (other airlines).[23] In Kline and Tamer's analysis (and in ours), the two categories

[23]The data were collected from the second quarter of the 2010 Airline Origin and Destination Survey (DB1B). The low cost carriers are AirTran, Allegiant Air, Frontier, JetBlue, Midwest Air, Southwest, Spirit, Sun Country, USA3000, and Virgin America. A firm that is not a low cost carrier is, by definition, an "other airline".

are treated as two firms. Thus, in each market, the two firms, LCC and OA, can either both enter a market, both stay out, or one could enter with the other staying out.

This data set also contains information on two covariates: market presence (MP) and market size (MS). Market presence is a market- and airline-specific variable. For each airline and for each airport, one counts the number of markets that the airline serves from that airport and divide it by the total number of markets served from that airport by any airline; the market presence variable for a given market and airline is the average of these ratios at endpoints of that market/trip. The construction and inclusion of this covariate is not novel and follows Berry (1992). Since the airlines are aggregated into two firms, the market presence variable is also aggregated: the market presence for LCC (resp., OA) is the maximum among the actual airlines in the LCC category (resp., OA category). The second covariate, market size, is a market-specific variable (shared by all airlines in that market) and is defined as the population at endpoints of the corresponding trip.

Furthermore, Kline and Tamer (2016) discretize these variables, where each of them takes value 1 if the variable is higher than its median value and 0 otherwise. Thus, in our data set, there are three binary covariates, MP_{LCC}, MP_{OA}, and MS, and markets are partitioned into eight groups according to realizations of them. Formally, $\mathbf{X} = \{0, 1\}^3$, and in this case, it also holds that $\widehat{\mathbf{X}} = \mathbf{X}$. Note that MS simultaneously influences the payoffs of both LCC and OA, and hence the covariates affecting LCC's payoff can be written as $x_{LCC} = (\text{MP}_{LCC}, \text{MS})$, and similarly, $x_{OA} = (\text{MP}_{OA}, \text{MS})$.

Observations in the data set can be used to calculate the empirical choice distributions that we include in Table 3. It consists of eight blocks, with the markets in each block sharing the same covariates. For example, there are 1271 markets with $(\text{MP}_{LCC}, \text{MP}_{OA}, \text{MS}) = (0, 0, 0)$, of which around 30% are not served by either airline and about 68% are served only by airlines in the OA category (an action profile is written as $(y_{LCC}, y_{OA}) \in \{E, N\} \times \{E, N\}$). The entries in Table 3 seem "reasonable," in the sense that it appears as though a firm's entry is encouraged whenever its market presence is large or the market size is large, and it is deterred by the entry of the other firm. For example, going from $(0, 0, 0)$ to $(1, 0, 0)$ (so the market presence of LCC has increased), both $\text{Q}(N, N)$ and $\text{Q}(N, E)$ fall, while $\text{Q}(E, N)$ and $\text{Q}(E, E)$ both increase.

Testing $\mathcal{SC}$-rationalizability. Our hypothesis is that, in each market, two firms (LCC and OA) are playing a pure strategy Nash equilibrium in a game of strategic substitutes with monotone effects from covariates. The payoff function of LCC, say, $\Pi_{LCC}(y_{LCC}, y_{OA}, \text{MP}_{LCC}, \text{MS})$, is required to obey single-crossing differences in $(y_{LCC}; (-y_{OA}, \text{MP}_{LCC}, \text{MS}))$, and similarly, the payoff function of OA, $\Pi_{OA}(y_{OA}, y_{LCC}, \text{MP}_{OA}, \text{MS})$, is required to obey single-crossing differences in $(y_{OA}; (-y_{LCC}, \text{MP}_{OA}, \text{MS}))$. This ensures that a firm's entry is discouraged by the opponent's entry and enhanced by an increase in own covariates. The data set is supposed to arise from a population of those firms, with unobserved heterogeneity generating a distribution of realizations of payoff functions $\mathbf{\Pi} = (\Pi_{LCC}, \Pi_{OA})$, which we denote by $\text{P}_{\mathbf{\Pi}}$, and an equilibrium selection rule.

Employing the statistical test in Section 4.1, we find a p-value of 0.138, and hence, the hypothesis that the empirical choice frequencies are explained by our modeling restric-

Table 3. Empirical distribution across each realization of covariates.

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (0,0,0)$ 1271 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.304	0.682	0.006	0.009

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (0,1,0)$ 763 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.190	0.785	0.003	0.022

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (1,0,0)$ 1125 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.194	0.367	0.253	0.186

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (1,1,0)$ 782 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.122	0.542	0.050	0.286

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (0,0,1)$ 869 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.159	0.823	0.001	0.017

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (0,1,1)$ 1039 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.078	0.889	0.000	0.033

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (1,0,1)$ 677 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.106	0.326	0.306	0.261

$(\mathrm{MP}_{LCC}, \mathrm{MP}_{OA}, \mathrm{MS}) = (1,1,1)$ 1356 markets			
$\mathrm{Q}(N,N)$	$\mathrm{Q}(N,E)$	$\mathrm{Q}(E,N)$	$\mathrm{Q}(E,E)$
0.055	0.501	0.021	0.423

tions cannot be refuted at 5% (or 10%) significance level. We choose the tuning parameter $\kappa_N = 10^{-3}\sqrt{\log \underline{N}_{\mathbf{x}}/\underline{N}_{\mathbf{x}}}$, where $\underline{N}_{\mathbf{x}} = \min_{\mathbf{x}\in\widehat{\mathbf{X}}} N_{\mathbf{x}}$, and the number of bootstrap samples as $R = 2000$.[24] Note that having a p-value strictly less than 1 means that J_N defined in (13) is strictly positive, that is, there is a strictly positive distance between our empirical distribution and $\mathbb{P}^{\mathrm{SC}}$, the set of (exactly) $\mathcal{SC}$-rationalizable distributions. Using our R code on a desktop computer with Apple M1 processor and 16 GB RAM, the p-value was calculated in less than 3 minutes.

In this setting, the set $\mathbf{X} = \widehat{\mathbf{X}}$ has exactly eight elements, and hence, the number of possible group types is $4^8 \approx 65{,}000$. In this small environment, it is in fact not difficult to check the RM axiom for each of these group types. Doing that, we find that only 482 types satisfy single-crossing (equivalently, satisfy the RM axiom). This gives a sense of the "empirical bite" of our test: the data set has to be explained by using a very small fraction (less than 1%) of all possible group types.

Significance of strategic interactions. Having established that the data set is (statistically) $\mathcal{SC}$-rationalizable, we can now go on to explore its properties. In particular, we can assess the extent to which strategic interactions play a role in explaining the data, in the sense discussed in Section 3.4, by considering the subclasses of single-crossing group types that correspond to: (i) the LCC firm having a payoff function that is independent of the actions of OA; (ii) the OA firm having a payoff function that is independent of the actions of LCC; and (iii) both firms having payoff functions that are independent of the other firm's action. Applying the procedure explained in Section 4.2, we find that the greatest possible weights on these three subclasses of single-crossing group types are (i) 0.923, (ii) 0.790, and (iii) 0.789 (within 5% significance level). Since these weights are all strictly less than 1, we conclude that any $\mathcal{SC}$-rationalization of the data *requires* strategic

[24] Recall that $N_{\mathbf{x}}$ is the number of observations with covariates $\mathbf{x}$. We follow Kitamura and Stoye (2018) in having κ_N proportional to $\sqrt{\log \underline{N}_{\mathbf{x}}/\underline{N}_{\mathbf{x}}}$.

TABLE 4. Probability bounds for equilibrium action profiles.

$(\text{MP}_{LCC}, \text{MP}_{OA}, \text{MS})$	$(0,0,0)$		$(0,1,0)$		$(1,0,0)$		$(1,1,0)$	
Action profile	(N,E)	(E,N)	(N,E)	(E,N)	(N,E)	(E,N)	(N,E)	(E,N)
$\max \Pr[\bar{\mathbf{y}} \in \mathbf{NE}(\boldsymbol{\Pi}, \bar{\mathbf{x}})]$	0.699	0.544	0.815	0.503	0.503	0.644	0.558	0.555
Observed Prob.	0.682	0.006	0.785	0.003	0.367	0.253	0.542	0.050
$(\text{MP}_{LCC}, \text{MP}_{OA}, \text{MS})$	$(0,0,1)$		$(0,1,1)$		$(1,0,1)$		$(1,1,1)$	
Action profile	(N,E)	(E,N)	(N,E)	(E,N)	(N,E)	(E,N)	(N,E)	(E,N)
$\max \Pr[\bar{\mathbf{y}} \in \mathbf{NE}(\boldsymbol{\Pi}, \bar{\mathbf{x}})]$	0.841	0.616	0.913	0.496	0.485	0.661	0.523	0.497
Observed Prob.	0.832	0.001	0.910	0.000	0.326	0.306	0.501	0.021

behavior for both LCC and OA firms. The computation time for each case was about 27 minutes.

Probability bounds for equilibrium actions. Under our behavioral hypothesis, the action profiles (N, N) and (E, E) can only be played as the *unique* equilibrium at any realization of $\mathbf{x} = (\text{MP}_{LCC}, \text{MP}_{OA}, \text{MS})$. On the other hand, when (N, E) is played, it is possible that (E, N) is also a Nash equilibrium of the game. For this reason, the probability that (E, N) is a Nash equilibrium of the game can be strictly higher than the observed frequency with which this profile is played, even after accounting for sampling variability (the same goes for (N, E)). Applying the argument in Sections 3.4 and 4.2, we can recover the greatest possible weight on group types in the population which have (E, N) as a Nash equilibrium at a given covariate value (and similarly for (N, E)). These are reported as $\max \Pr[\bar{\mathbf{y}} \in \mathbf{NE}(\boldsymbol{\Pi}, \bar{\mathbf{x}})]$ in Table 4.[25]

For example, at $(\text{MP}_{LCC}, \text{MP}_{OA}, \text{MS}) = (1, 0, 0)$, the greatest possible weight on those group types that may have (N, E) as a Nash equilibrium of the game is 0.503: this includes types which are already playing (N, E) (with observed frequency 0.367) as well as types that are playing (E, N) but may have (N, E) as an alternative Nash equilibrium.[26] Thus, even if we allow for equilibrium selection rules to change, and (N, E) is chosen whenever it is a PSNE, the frequency with which (N, E) is played at $(1, 0, 0)$ will not exceed 0.503. Notice that, in general, $\max \Pr[\bar{\mathbf{y}} \in \text{NE}(\Pi, \bar{\mathbf{x}})]$ is closer to the observed frequency in the case where $\bar{\mathbf{y}} = (N, E)$, while the same gap in the case of $\bar{\mathbf{y}} = (E, N)$ is considerably bigger. For $\bar{\mathbf{y}} = (N, E)$ (and similarly for $\bar{\mathbf{y}} = (E, N)$), the calculation of $\max \Pr[\bar{\mathbf{y}} \in \text{NE}(\Pi, \bar{\mathbf{x}})]$ for all $\mathbf{x} \in \widehat{\mathbf{X}}$ took around 38 minutes.[27]

Finer division of covariates. The tests that we have done so far do not really put the column generation method through its paces: the total number of possible group types

[25] To be precise, $\max \Pr[\bar{\mathbf{y}} \in \mathbf{NE}(\boldsymbol{\Pi}, \bar{\mathbf{x}})]$ is the upper limit of the confidence interval on $\sum_{\mathrm{B} \in \mathcal{B}^*} \tau^{\mathrm{B}}$ where $\mathcal{B}^*$ is defined as in Application 2 of Section 3.4.

[26] But it is *not* the case that every single-crossing group type with (E, N) as a Nash equilibrium at $(1, 0, 0)$ must also have (N, E) as a Nash equilibrium at $(1, 0, 0)$. For example, if the group type chooses (E, N) at $(1, 0, 0)$ and (E, E) at $(1, 1, 0)$, then (N, E) cannot be a Nash equilibrium at $(1, 0, 0)$. On the other hand, there are single-crossing group types that choose (E, N) at $(1, 0, 0)$ and (N, E) at $(1, 1, 0)$; in these cases, (N, E) may be a Nash equilibrium at $(1, 0, 0)$. The latter types are the ones included in the estimated weight, while the former types are excluded.

[27] See Supplementary Appendix A6.2 for an estimate of $\min \Pr[\bar{\mathbf{y}} \in \text{NE}(\boldsymbol{\Pi}, \bar{\mathbf{x}})]$.

($4^8 = 65{,}536$) is just about small enough to be completely listed; one could then find all the SC-rationalizable group types using the RM axiom (of which there are 482) and avoid using column generation altogether.

To check the performance of the column generation method in a "larger" model, we repeat our analysis with a finer division of the covariates. (A fuller discussion is found in Supplementary Appendix A6.1.) Instead of aggregating covariates into binary variables, we let each of MP_{LCC}, MP_{OA}, and MS take four possible values using quantiles: each variable takes value $k-1$, if it is in the kth quartile. In this way, all markets are partitioned into $4^3 = 64$ covariate values and there is a distribution of entry decisions at each of them. In this environment, the total number of possible group types is enormous (4^{64}) and the same is true of the number of $\mathcal{SC}$-rationalizable group types.[28] While a direct approach is no longer feasible, the column generation method still works, with the test of $\mathcal{SC}$-rationalizability finishing in around 3 minutes, including the bootstrap procedure. In this case, we find that the null hypothesis is rejected with p-value equal to 0.[29]

The conflicting results cast doubts on the robustness of the model to explain choices of airline firms. In Supplementary Appendix 6.1, we implement the tests for even finer discretizations. Naturally, the number of markets at each discrete value of the covariate falls as the discretization becomes finer and so we only use those covariate values that contain a certain minimal number of markets; in other words, we have a case where $\widehat{\mathbf{X}}$ is a strict subset of $\mathbf{X}$. Broadly speaking (see Supplementary Appendix A6.1 for details), we find that the hypothesis is supported if the market presence variables remain binary, even with finer discretizations of market size. Finer discretizations of the market presence variables lead to rejection of the hypothesis, even when the market size variable remains binary.

The precise reasons for this failure are unclear and require a more careful analysis. In terms of the model's explicit assumptions, the failure could be attributed to a failure of the single-crossing property on payoff functions, the failure of firms to play PSNE in each market, or the failure of the conditional independence assumption (especially with finer discretization of the market presence covariate). There could also be problems having to do with the model's basic structure, such as the particular way in which market presence is calculated or the modeling of interaction as a *two*-action, *two*-agent game (which ignores the possible presence of multiple carriers within the LCC or OA category in each market, the scale of their operations if they enter, or the fact that carriers may operate and interact in multiple markets and have more complicated payoff functions).[30]

It is worth noting that the uneven performance of the model with finer discretizations of the covariates is also detectable when we implement the procedure of Kline and

[28] It is straightforward to see that any group type where either (E, N) or (N, E) is played at a covariate obeys the RM axiom. Hence, there are at least 2^{64} ($\approx 3.1 \times 10^{19}$) types obeying the RM axiom.

[29] Since the validity of our test still hinges on a sampling framework in which $N_{\mathbf{x}} \to \infty$ as $N \to \infty$, for each $\mathbf{x} \in \widehat{\mathbf{X}}$ (see Section 4.1), the covariates cannot be dividedly too finely. (See Supplementary Appendix Section A6.1 for details.)

[30] The last point is related to the literature on multimarket contact; see, for example, Evans and Kessides (1994).

Tamer (2016), which assumes conditional independence but has a parametric specification (see Supplementary Appendix A6.1). Specifically, while their approach is mainly developed for inference of partially identified parameters, they point out that it could also be used for "specification testing." In their procedure, the data is repeatedly sampled and, for each sample, one could derive the set of model parameters. Kline and Tamer point out that the frequency with which the identified set (of parameters) is nonempty could be thought of as a form of specification testing. Following this idea, we implement their inference approach for different levels of discretization of the covariates. We find that, with binary covariates, the frequency of nonempty sets of the estimated parameters is 1, but this frequency goes to 0 with finer discretizations.

Appendix

Proof of Theorem 1. We have already shown that the RM axiom is necessary for a group type to be single-crossing. It remains for us to prove the converse. We show that if a generalized group type $\mathrm{B} : \widehat{\mathbf{X}} \rightrightarrows \mathbf{Y}$ obeys the RM axiom, then we can always find a profile of payoff functions $\mathbf{\Pi} = (\Pi_i)_{i \in \mathcal{N}}$ obeying the requirements in Definition 3 so that each Π_i has increasing differences and is single-peaked. We say that $\Pi_i(y_i, \mathbf{y}_{-i}, x_i)$ has *increasing differences* in $(y_i; \mathbf{y}_{-i}, x_i)$, if

$$\Pi_i\big(y_i'', \mathbf{y}_{-i}'', x_i''\big) - \Pi_i\big(y_i', \mathbf{y}_{-i}'', x_i''\big) \geq \Pi_i\big(y_i'', \mathbf{y}_{-i}', x_i'\big) - \Pi_i\big(y_i', \mathbf{y}_{-i}', x_i'\big)$$

for every $y_i'' > y_i'$ and $(\mathbf{y}_{-i}'', x_i'') > (\mathbf{y}_{-i}', x_i')$. (This property obviously implies single-crossing differences.) The single-peaked property means that, for each $(\mathbf{y}_{-i}, \mathbf{x})$, there is $\bar{y}_i$ such that $\Pi_i(\bar{y}_i, \mathbf{y}_{-i}, \mathbf{x}) > \Pi_i(y_i, \mathbf{y}_{-i}, \mathbf{x})$ for all $y_i \neq \bar{y}_i$, with $\Pi_i(y_i, \mathbf{y}_{-i}, \mathbf{x})$ being strictly increasing in y_i for $y_i \leq \bar{y}_i$ and strictly decreasing in y_i for $\bar{y}_i \geq y_i$.

We shall show that, for each $i \in \mathcal{N} = \{1, 2, \dots, n\}$, there is a single-peaked payoff function $\Pi_i : Y_i \times \mathbf{Y}_{-i} \times X_i \to \mathbb{R}$ such that (a) Π_i has increasing differences in $(y_i; \mathbf{y}_{-i}, x_i)$; (b) $\mathbf{y} \in \mathrm{B}(\mathbf{x}) \Longrightarrow \mathbf{y} \in \mathrm{NE}(\mathbf{\Pi}, \mathbf{x})$; and (c) for each $(\mathbf{y}_{-i}, x_i)$, $\mathrm{BR}_i(\mathbf{y}_{-i}, x_i)$ is a singleton (even when $x_i \notin \mathrm{proj}_i\widehat{\mathbf{X}}$).

For each $i \in \mathcal{N}$, let $\mathbf{Z}_i = \mathbf{Y}_{-i} \times X_i$. Since $\widehat{\mathbf{X}}$ and $\mathbf{Y}$ are finite sets, the graph of $\mathrm{B}(\mathbf{x})$, which is $\mathcal{G}(\mathrm{B}) := \{(\mathbf{y}, \mathbf{x}) : \mathbf{y} \in \mathrm{B}(\mathbf{x}) \text{ for some } \mathbf{x} \in \widehat{\mathbf{X}}\}$, is also a finite set. Hence, it can be written as $\mathcal{G}(\mathrm{B}) = \{(\mathbf{y}^t, \mathbf{x}^t) : \mathbf{y}^t \in \mathrm{B}(\mathbf{x}^t) \text{ for } t \in \mathcal{T}\}$ where $\mathcal{T} = \{1, 2, \dots, T\}$ is a finite index set. Letting $\mathbf{z}_i^t := (\mathbf{y}_{-i}^t, x_i^t)$ for $t \in \mathcal{T}$, each $(\mathbf{y}^t, \mathbf{x}^t) \in \mathcal{G}(\mathrm{B})$ can be written as $(y_i^t, \mathbf{z}_i^t)$ for every $i \in \mathcal{N}$.

To obtain a payoff function Π_i defined on $Y_i \times \mathbf{Z}_i (= Y_i \times \mathbf{Y}_{-i} \times X_i)$, we begin with a family of single-peaked functions, $f_i : Y_i \times \mathcal{T} \to \mathbb{R}$ satisfying the following properties: (i) $f_i(y_i^t, t) > f_i(a, t)$ for all $a \neq y_i^t$ and $a \in Y_i$; and (ii) if $y_i^s = y_i^t$, then $f_i(\cdot, s) = f_i(\cdot, t)$ and if $y_i^s > y_i^t$, then $f_i(a'', s) - f_i(a', s) > f_i(a'', t) - f(a', t)$ for all $a'' > a'$ in Y_i. This can be obtained, for example, by letting $f_i(a, t) = -(a - y_i^t)^2$. Then we define $\Pi_i : Y_i \times \mathbf{Z}_i \to \mathbb{R}$ as follows. For each $\mathbf{z} \in \mathbf{Z}_i$, let $T(\mathbf{z}) = \{t \in \mathcal{T} : \mathbf{z}^t \geq \mathbf{z}\} \cup \{\hat{t}\}$, where $\hat{t}$ is any index that satisfies $y_i^{\hat{t}} \geq y_i^t$ for all $t \in \mathcal{T}$. Since it contains $\hat{t}$ at least, $T(\mathbf{z})$ is nonempty. Choose $\tilde{t}(\mathbf{z}) \in T(\mathbf{z})$ such that $y_i^{\tilde{t}(\mathbf{z})} \leq y_i^t$ for all $t \in T(\mathbf{z})$, and define $\Pi_i(\cdot, \mathbf{z}) = f_i(\cdot, \tilde{t}(\mathbf{z}))$. Although there may be more than one candidate for $\tilde{t}(\mathbf{z})$, by property (ii) of f_i, the value of Π_i is not affected by the choice.

We claim that $\Pi_i(\cdot, \mathbf{z})$ defined above obeys properties (a)–(c). For property (a), suppose that at $\mathbf{z}'$, we have $\Pi_i(\cdot, \mathbf{z}') = f_i(\cdot, t')$ and for $\mathbf{z}''$, we have $\Pi_i(\cdot, \mathbf{z}'') = f_i(\cdot, t'')$. If $\mathbf{z}'' > \mathbf{z}'$, then $T(\mathbf{z}'') \subseteq T(\mathbf{z}')$, and so $y_i^{t''} \geq y_i^{t'}$. By property (ii) of f_i, we obtain

$$\Pi_i(a'', \mathbf{z}'') - \Pi_i(a', \mathbf{z}'') \geq \Pi_i(a'', \mathbf{z}') - \Pi_i(a', \mathbf{z}') \quad \text{for all } a'' > a'.$$

Thus Π_i satisfies *increasing-differences*, which means it satisfies single-crossing differences. For property (b), notice that, at any $\mathbf{z}^s$, we have $s \in T(\mathbf{z}^s)$ and, by the RM axiom, $y_i^t \geq y_i^s$ for any $t \in T(\mathbf{z}^s)$. It follows that $\Pi_i(\cdot, \mathbf{z}^s) = f_i(\cdot, s)$, and so $\operatorname{argmax}_{a \in Y_i} \Pi_i(a, \mathbf{z}^s) = y_i^s$. Lastly, property (c) flows from the single-peakedness of each $f_i(\cdot, t)$. □

Proof of the "only if" part of Theorem 2. Suppose $\mathcal{P}$ is $\mathcal{SC}$-rationalizable with distribution $\mathrm{P}_{\mathbf{\Pi}}$ and the equilibrium selection rule λ. Let

$$p(\mathrm{B}, \mathbf{\Pi}) = \times_{\mathbf{x} \in \widehat{\mathbf{X}}} \lambda\big(\mathrm{B}(\mathbf{x}) | \mathbf{\Pi}, \mathbf{x}\big) \tag{25}$$

and let $\tau^{\mathrm{B}} = \int p(\mathrm{B}, \mathbf{\Pi})\, d\mathrm{P}_{\mathbf{\Pi}}$. If B is not a single-crossing type, then $p(\mathrm{B}, \mathbf{\Pi}) = 0$ for all $\mathbf{\Pi} \in \mathcal{SC}$. Therefore, for $\mathbf{\Pi} \in \mathcal{SC}$,

$$\sum_{\mathrm{B} \in \mathcal{B}} p(\mathrm{B}, \mathbf{\Pi}) = 1, \tag{26}$$

which guarantees that $\sum_{\mathrm{B} \in \mathcal{B}} \tau^{\mathrm{B}} = \int_{\mathcal{SC}} d\mathrm{P}_{\mathbf{\Pi}} = 1$ (since the support of $\mathrm{P}_{\mathbf{\Pi}}$ lies in $\mathcal{SC}$). Furthermore, it follows from (25) that $\lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x}) = \sum_{\{\mathrm{B} \in \mathcal{B}: \mathrm{B}(\mathbf{x}) = \mathbf{y}\}} p(\mathrm{B}, \mathbf{\Pi})$, and thus

$$\mathrm{P}(\mathbf{y}|\mathbf{x}) = \int \lambda(\mathbf{y}|\mathbf{\Pi}, \mathbf{x})\, d\mathrm{P}_{\mathbf{\Pi}} = \sum_{\{\mathrm{B} \in \mathcal{B}: \mathrm{B}(\mathbf{x}) = \mathbf{y}\}} \int p(\mathrm{B}, \mathbf{\Pi})\, d\mathrm{P}_{\mathbf{\Pi}} = \sum_{\{\mathrm{B} \in \mathcal{B}: \mathrm{B}(\mathbf{x}) = \mathbf{y}\}} \tau^{\mathrm{B}}.$$ □

Proof of equation (10). If $\mathcal{P}$ is $\mathcal{SC}$-rationalizable and there is $(\tau^{\mathrm{B}})_{\mathrm{B} \in \mathcal{B}}$ that solves (7), then there is P_{Π} that rationalizes $\mathcal{P}$ such that $\int_{\Pi \in \mathcal{SC}^*} d\mathrm{P}_{\Pi} = \sum_{\mathrm{B} \in \mathcal{B}^*} \tau^{\mathrm{B}}$; this is clear from the proof of the "if" part of Theorem 2. Conversely, for any rationalization P_{Π} of $\mathcal{P}$, we claim there is $(\tau^{\mathrm{B}})_{\mathrm{B} \in \mathcal{B}}$ that solves (7) such that $\int_{\Pi \in \mathcal{SC}^*} d\mathrm{P}_{\Pi} \leq \sum_{\mathrm{B} \in \mathcal{B}^*} \tau^{\mathrm{B}}$. Indeed, construct $(\tau^{B})_{\mathrm{B} \in \mathcal{B}}$ in the same way as in the proof of Theorem 2. Notice that if $\Pi \in \mathcal{SC}^*$, then $p(\mathrm{B}, \Pi) = 0$ for any $\mathrm{B} \notin \mathcal{B}^*$ and so it follows from (26) that $\sum_{\mathrm{B} \in \mathcal{B}^*} p(\mathrm{B}, \mathbf{\Pi}) = 1$ since $p(\mathrm{B}, \Pi) = 0$ for any $\mathrm{B} \notin \mathcal{B}^*$. This gives us (10). □

The proof of Proposition 1 requires the following well-known result from convex analysis.

Lemma 1. *Let V be a closed convex set in $\mathbb{R}^n$ and let $\mathbf{r} \in \mathbb{R}^n \setminus V$. Then there is a unique $\mathbf{v}^* \in V$ such that $\|\mathbf{r} - \mathbf{v}^*\| = \min_{\mathbf{v} \in V} \|\mathbf{r} - \mathbf{v}\|$. The point $\mathbf{v}^*$ is the unique point in V with the property that $(\mathbf{r} - \mathbf{v}^*) \cdot (\mathbf{v} - \mathbf{v}^*) \leq 0$ for all $\mathbf{v} \in V$.*

Proof of Proposition 1. If for all $\mathbf{b} \in \mathcal{B}$, we have $(\mathbf{p} - \eta_0) \cdot (\mathbf{b} - \eta_0) \leq 0$, then $(\mathbf{p} - \eta_0) \cdot (\mathbf{v} - \eta_0) \leq 0$ for all $\mathbf{v}$ in the conical hull of $\mathcal{B}$. This implies, by Lemma 1 that the distance between $\mathbf{p}$ and any $\mathbf{v}$ in the conical hull of $\mathcal{B}$ is again minimized at η_0, which means that $\mathcal{B}_0$ is not improvable. Conversely, if there is $\widehat{\mathbf{b}}$ in $\mathcal{B}$ such that (17) holds, then appealing to Lemma 1 again, we know that η_0 does *not* minimize the distance between $\mathbf{p}$ and the conical hull of $\mathcal{B}_0 \cup \{\widehat{\mathbf{b}}\}$ and $\widehat{\mathbf{b}}$ improves $\mathcal{B}_0$. □

Proof of Proposition 2. For each $(\mathbf{y}, \mathbf{x}) \in \mathbf{Y} \times \widehat{\mathbf{X}}$, define $\mathcal{R}(\mathbf{y}, \mathbf{x}) \subset \mathbf{Y} \times \widehat{\mathbf{X}}$ such that

$$\mathcal{R}(\mathbf{y}, \mathbf{x}) = \{(\mathbf{y}', \mathbf{x}') : \mathbf{y} = \mathrm{B}(\mathbf{x}) \Longrightarrow \mathbf{y}' \neq \mathrm{B}(\mathbf{x}') \text{ for all } \mathrm{B} \in \mathcal{B}\}.$$

Recalling the definition of the RM axiom (6), $(\mathbf{y}', \mathbf{x}') \in \mathcal{R}(\mathbf{y}, \mathbf{x})$ holds if there exists some $i \in \mathcal{N}$ such that $y_i' < (>) y_i$ and $(y_{-i}', x_i') \geq (\leq) (y_{-i}, x_i)$. Impose any linear ranking on the elements of $\mathbf{Y} \times \widehat{\mathbf{X}}$; we define $C = (c_{(\mathbf{y},\mathbf{x}),(\mathbf{y}',\mathbf{x}')})_{\mathbf{Y}\times\widehat{\mathbf{X}},\mathbf{Y}\times\widehat{\mathbf{X}}}$ to be a $|\mathbf{Y} \times \widehat{\mathbf{X}}| \times |\mathbf{Y} \times \widehat{\mathbf{X}}|$ matrix where $c_{(\mathbf{y},\mathbf{x}),(\mathbf{y},\mathbf{x})} = |\widehat{\mathbf{X}}|$ and, if $(\mathbf{y}, \mathbf{x}) \neq (\mathbf{y}', \mathbf{x}')$, then $c_{(\mathbf{y},\mathbf{x}),(\mathbf{y}',\mathbf{x}')} = \mathbf{1}[(\mathbf{y}', \mathbf{x}') \in \mathcal{R}(\mathbf{y}, \mathbf{x})]$. By setting $\theta = (|\widehat{\mathbf{X}}|, |\widehat{\mathbf{X}}|, \dots, |\widehat{\mathbf{X}}|)$ (a column vector of length $|\mathbf{Y} \times \widehat{\mathbf{X}}|$), we claim that a single-valued group type $\mathbf{b}$ (thought of as a column vector) obeys RM axiom if and only if $C\mathbf{b} \leq \theta$. Indeed, since $\mathbf{b}$ is single-valued, we have $\sum_{\mathbf{y}\in\mathbf{Y}} \mathbf{b}_{(\mathbf{y},\mathbf{x})} = 1$ for all $\mathbf{x} \in \widehat{\mathbf{X}}$, which guarantees that $(C\mathbf{b})_{(\mathbf{y},\mathbf{x})} \leq |\widehat{\mathbf{X}}|$ if $\mathbf{b}_{(\mathbf{y},\mathbf{x})} = 0$. Note that $(C\mathbf{b})_{(\mathbf{y},\mathbf{x})} \geq c_{(\mathbf{y},\mathbf{x}),(\mathbf{y},\mathbf{x})} = |\widehat{\mathbf{X}}|$ if $\mathbf{b}_{(\mathbf{y},\mathbf{x})} = 1$. If $\mathbf{b}$ satisfies the RM axiom, then $(C\mathbf{b})_{(\mathbf{y},\mathbf{x})} = |\widehat{\mathbf{X}}|$ for all $(\mathbf{y}, \mathbf{x})$ with $\mathbf{b}_{(\mathbf{y},\mathbf{x})} = 1$; if $\mathbf{b}$ violates the RM axiom, then there is $(\tilde{\mathbf{y}}, \tilde{\mathbf{x}})$ with $\mathbf{b}_{(\tilde{\mathbf{y}},\tilde{\mathbf{x}})} = 1$ such that $(C\mathbf{b})_{(\tilde{\mathbf{y}},\tilde{\mathbf{x}})} > |\widehat{\mathbf{X}}|$. □

We make crucial use of the following result in our proof of Proposition 3.

Lemma 2. *Suppose that $\mathcal{B}' \subset \mathcal{B}$ where $\mathcal{B}' \cap \mathcal{B}^*$ is nonempty. Let $V(\mathcal{B}')$ be the set such that $\mathbf{v} \in V(\mathcal{B}')$ if $\mathbf{v} = \mathbf{B}'\tau$ and $\sum_{\mathbf{b}\in\mathcal{B}'\cap\mathcal{B}^*} \tau^{\mathbf{b}} \geq \beta$, where $\mathbf{B}'$ is a matrix representation of $\mathcal{B}'$. Then $V(\mathcal{B}')$ is the convex hull of vectors of the form $\beta\mathbf{b}^* + (1-\beta)\mathbf{b}$, where $\mathbf{b}^* \in \mathcal{B}' \cap \mathcal{B}^*$ and $\mathbf{b} \in \mathcal{B}'$.*

Proof. Clearly, the convex hull of those vectors is contained in $V(\mathcal{B}')$, so we need only show the other inclusion. Note that any $\mathbf{v} \in V(\mathcal{B}')$ can be written as $\beta(\sum_{l=1}^{\bar{l}} t_l \mathbf{b}_l^*) + (1 - \beta)(\sum_{k=1}^{\bar{k}} s_k \mathbf{b}_k)$ where $t_l, s_k \geq 0$, $\sum_{l=1}^{\bar{l}} t_l = \sum_{k=1}^{\bar{k}} s_k = 1$, $\mathbf{b}_l^* \in \mathcal{B}^*$, and $\mathbf{b}_k \in \mathcal{B}'$. By breaking up the convex sums into smaller parts if necessary, we can, with no loss of generality, assume that $t_l = s_k$ and $\bar{l} = \bar{k}$. Then

$$\mathbf{v} = \beta\left(\sum_{l=1}^{\bar{l}} t_l \mathbf{b}_l^*\right) + (1-\beta)\left(\sum_{l=1}^{\bar{l}} t_l \mathbf{b}_l\right) = \sum_{l=1}^{\bar{l}} t_l\left[\beta\mathbf{b}_l^* + (1-\beta)\mathbf{b}_l\right],$$

which establishes our claim. □

Proof of Proposition 3. Note that $J_{N,0}(\beta)$ is the distance between $\mathbf{q}$ and $V(\mathcal{B}_0)$ and this distance is achieved at $\eta_0 \in V(\mathcal{B}_0)$. If, for all $\beta\mathbf{b}^* + (1-\beta)\mathbf{b}$ where $\mathbf{b}^* \in \mathcal{B}^*$ and $\mathbf{b} \in \mathcal{B}$, we have

$$(\mathbf{q} - \eta_0) \cdot (\beta\mathbf{b}^* + (1-\beta)\mathbf{b} - \eta_0) \leq 0,$$

then $(\mathbf{q} - \eta_0) \cdot (\mathbf{v} - \eta_0) \leq 0$ for all $\mathbf{v} \in V(\mathcal{B})$, by Lemma 2. This in turn means, by Lemma 1, that $\mathcal{B}_0$ is not improvable given problem (20). Conversely, suppose that there is a pair of group types $\{\widehat{\mathbf{b}}^*, \widehat{\mathbf{b}}\}$, with $\widehat{\mathbf{b}}^* \in \mathcal{B}^*$ and $\widehat{\mathbf{b}} \in \mathcal{B}$, such that (22) holds. Then, by Lemma 1, η_0 does *not* minimize the distance between $\mathbf{q}$ and the convex hull of $V(\mathcal{B}_0)$ and $\beta\widehat{\mathbf{b}}^* + (1 - \beta)\widehat{\mathbf{b}}$. We conclude that $\{\widehat{\mathbf{b}}^*, \widehat{\mathbf{b}}\}$ improves $\mathcal{B}_0$ given problem (20). □

References

Apesteguia, Jose, Miguel A. Ballester, and Jay Lu (2017), “Single-crossing random utility models.” *Econometrica*, 85, 661–674. [0237]

Aradillas-Lopez, Andres (2011), “Nonparametric probability bounds for Nash equilibrium actions in a simultaneous discrete game.” *Quantative Economics*, 2, 135–171. [0238, 0251]

Beresteanu, Arie (2005), “Nonparametric analysis of cost complementarities in the telecommunications industry.” *RAND Journal of Economics*, 36, 870–889. [0237]

Beresteanu, Arie (2007), “Nonparametric estimation of regression functions under restrictions on partial derivatives.” http://www.pitt.edu/~arie/PDFs/shape.pdf. [0237]

Berry, Steven T. (1992), “Estimation of a model of entry in the airline industry.” *Econometrica*, 60, 889–917. [0236, 0257]

Bresnahan, Timothy F. and Peter C. Reiss (1990), “Entry in monopoly markets.” *Review of Economic Studies*, 57, 531–553. [0236]

Carvajal, Andres (2004), “Testable restrictions of Nash equilibrium with continuous domains.” Discussion paper series, Royal Holloway College, University of London, 2004–2026. [0238]

Chen, Xiaohong, Timothy M. Christensen, and Elie Tamer (2018), “Monte Carlo confidence sets for identified sets.” *Econometrica*, 86, 1965–2018. [0236, 0245]

Chernozhukov, Victor, Iván Fernández-Val, Jinyong Hahn, and Whitney Newey (2013), “Average and quantile effects in nonseparable panel models.” *Econometrica*, 81, 535–580. [0237]

Ciliberto, Federico and Elie Tamer (2009), “Market structure and multiple equilibria in airline markets.” *Econometrica*, 77, 1791–1828. [0236, 0240]

Deb, Rahul, Yuichi Kitamura, John K.-H. Quah, and Jörg Stoye (2023), “Revealed price preference: Theory and empirical analysis.” *Review of Economic Studies*, 90, 707–743. [0237, 0238, 0245, 0251, 0255]

Dubey, Pradeep, Ori Haimanko, and Andry Zapechelnyuk (2006), “Strategic complements and substitutes, and potential games.” *Games and Economic Behavior*, 64, 77–94. [0245]

Evans, William and Ioannis Kessides (1994), “Living by the ‘golden rule’: Multimarket contact in the U.S. airline industry.” *Quarterly Journal of Economics*, 109, 341–366. [0260]

Hoderlein, Stefan and Jörg Stoye (2014), “Revealed preferences in a heterogeneous population.” *Review of Economics and Statistics*, 96, 197–213. [0237]

Honoré, Bo E. and Elie Tamer (2006), “Bounds on parameters in panel dynamic discrete choice models.” *Econometrica*, 74, 611–629. [0237]

Jensen, Martin K. (2010), "Aggregative games and best-reply potentials." *Economic Theory*, 43, 45–66. [0245]

Kitamura, Yuichi and Jörg Stoye (2018), "Nonparametric analysis of random utility models." *Econometrica*, 86, 1883–1909. [0237, 0245, 0246, 0251, 0252, 0258]

Kline, Brendan and Elie Tamer (2016), "Bayesian inference in a class of partially identified models." *Quantative Economics*, 7, 329–366. [0237, 0239, 0242, 0256, 0257, 0260, 0261]

Lazzati, Natalia, John K.-H. Quah, and Koji Shirai (2018), "Nonparametric analysis of monotone choice." Discussion Paper Series 184, School of Economics, Kwansei Gakuin University. [0238]

Lazzati, Natalia, John K.-H. Quah, and Koji Shirai (2024), "Supplement to 'An ordinal approach to the empirical analysis of games with monotone best responses'." *Quantitative Economics Supplemental Material*, 15, https://doi.org/10.3982/QE2192. [0239]

Manski, Charles F. (2007), "Partial identification of counterfactual choice probabilities." *International Economic Review*, 48, 1393–1410. [0237]

Marschak, Jacob (1960), *Binary Choice Constraints on Random Utility Indicators* (Kenneth Arrow, ed.), Stanford Symposium on Mathematical Methods in the Social Sciences. Stanford University Press. [0237]

Matzkin, Rosa L. (2007), *Nonparametric Identification* (James J. Heckman and Edward E. Leamer, eds.), Handbook of Econometrics, vol. 6b. Elsevier Science. [0237]

McFadden, Daniel L. (2005), "Revealed stochastic preference: A synthesis." *Economic Theory*, 26, 245–264. [0237]

McFadden, Daniel L. and Marcel K. Richter (1991), "Stochastic rationality and revealed stochastic preference." In *Preferences, Uncertainty and Rationality* (John Chipman, Daniel L. McFadden, and Marcel K. Richter, eds.), 161–186, Westview Press, Boulder. [0237]

Milgrom, Paul and John Roberts (1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica*, 58, 1255–1277. [0244]

Milgrom, Paul and Chris Shannon (1994), "Monotone comparative statics." *Econometrica*, 62, 157–180. [0236, 0240, 0243, 0244]

Smeulders, Bart, Laurens Cherchye, and Bram De Rock (2021), "Nonparametric analysis of random utility models: Computational tools for statistical testing." *Econometrica*, 89, 437–455. [0238, 0239, 0251, 0252, 0253]

Tamer, Elie (2003), "Incomplete simultaneous discrete response model with multiple equilibria." *Review of Economic Studies*, 70, 147–165. [0237]

Tebaldi, Pietro, Alexander Torgovitsky, and Hanbin Yang (2023), "Nonparametric estimates of demand in the California health insurance exchange." *Econometrica*, 91, 107–146. [0237]

Vives, Xavier (1990), "Nash equilibria with strategic complementarities." *Journal of Mathematical Economics*, 19, 305–321. [0244]

Co-editor Stéphane Bonhomme handled this manuscript.

Manuscript received 29 June, 2022; final version accepted 20 August, 2024; available online 25 September, 2024.

The replication package for this paper is available at https://doi.org/10.5281/zenodo.13322699. The Journal checked the data and codes included in the package for their ability to reproduce the results in the paper and approved online appendices.