A revealed preference theory of

monotone choice and strategic complementarity

Natalia Lazzati, John K.-H. Quah, and Koji Shirai[*]

December 24, 2015

Abstract: We develop revealed preference characterizations of (1) monotone choice in the context of individual decision making and (2) strategic complementarity in the context of simultaneous games. We first consider the case where the observer has access to panel data and then extend the analysis to the case where data sets are cross sectional and preferences heterogenous. Lastly, we apply our techniques to investigate the possibility of spousal influence in smoking decisions.

Keywords: monotone comparative statics, single crossing differences, interval dominance, supermodular games, lattices

JEL classification numbers: C6, C7, D7

1 Introduction

Economists are often interested in knowing when the action chosen by an agent will increase (according to some ordering) with another variable, so that the two may be regarded as complements. The theory of monotone comparative statics provides conditions on preferences, such as single crossing differences, that guarantee this behavior. The objective of this paper is to provide a revealed preference analysis of monotone comparative statics. The starting point of our investigation is a data set collected from an agent where each observation consists of a choice problem and the action taken by the agent; we ask what conditions on these observations are necessary and sufficient for them to be consistent with the hypothesis that the agent is choosing according to a preference that obeys single crossing differences. This is an important question because, if the hypothesis is supported, then there are grounds for believing that any complementarity observed in the data will hold more generally, even outside the set of observations.

This introduction sets out the themes in this paper and summarizes our conclusions. The results discussed in Section 1.n are treated in detail in Section $n + 2$ (for $n = 1$, 2, and 3). Section 2 summarizes some of the key results in monotone comparative statics and games with strategic complementarity that motivate our analysis.

1.1 Single Crossing Differences and Revealed Complementarity

Consider an agent i who, after observing the realization of some exogenous variable, chooses an action from a feasible set. There is a binary relation $\succsim_i$ on $(x_i, \xi_i) \in \mathbb{R} \times \mathbb{R}$, where x_i is a possible action for agent i and ξ_i is some exogenous variable that may affect the agent's choice. With some abuse of terminology, we call $\succsim_i$ a *preference* if, for any fixed ξ_i, the restriction of $\succsim_i$ to the set $\{(x_i, \xi_i) : a_i \in \mathbb{R}\}$ is a complete, reflexive, and symmetric relation. Given ξ_i and a feasible action set $A_i \subset \mathbb{R}$, agent i's optimal choice (or best response) is

$$\mathrm{BR}(\xi_i, A_i, \succsim_i) = \{x_i' \in A_i : \ (x_i', \xi_i) \succsim_i (x_i, \xi_i) \text{ for all } x_i \in A_i\}. \tag{1}$$

What conditions guarantee that $\mathrm{BR}(\xi_i, A_i)$ is increasing in ξ_i, in the sense that every element in $\mathrm{BR}(\xi_i'', A_i)$ is greater than every element in $\mathrm{BR}(\xi_i', A_i)$, when $\xi_i'' > \xi_i'$? A fundamental result in

monotone comparative statics says that, for this to hold on every set A_i, it is necessary and sufficient that $\succsim_i$ obeys *strict single crossing differences* (Milgrom and Shannon, 1994). This property says that for every $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$,

$$(x_i'', \xi_i') \succsim_i (x_i', \xi_i') \implies (x_i'', \xi_i'') \succ_i (x_i', \xi_i''),$$

where $\succ_i$ is the strict preference induced by $\succsim_i$. In the case where we restrict the feasible action sets A_i to intervals of $\mathbb{R}$, then strict single crossing differences can be weakened and replaced by the *strict interval dominance* property (Quah and Strulovici, 2009), which says that

$$(x_i'', \xi_i') \succsim_i (x_i, \xi_i') \text{ for all } x_i \in [x_i', x_i''] \implies (x_i'', \xi_i'') \succ_i (x_i', \xi_i'').$$

These basic results motivate the following revealed preference problem. Suppose an observer has access to a data set with T observations, $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t=1}^T$, where a_i^t is the action chosen by agent i under the treatment (ξ_i^t, A_i^t), when the exogenous variable is ξ_i^t and the feasible action set is A_i^t, which we assume is a compact interval of $\mathbb{R}$. What condition on $\mathcal{O}_i$ will guarantee that there exists a binary relation $\succsim_i$ defined over $(x_i, \xi_i) \in \mathbb{R} \times \mathbb{R}$ that obeys the interval dominance order and rationalizes the agent's behavior in the sense that $a_i^t \in \mathrm{BR}(\xi_i^t, A_i^t, \succsim_i)$? It turns out that this hinges on an easy-to-check and easy-to-understand property on $\mathcal{O}_i$ we call the *axiom of revealed complementarity* (ARC). Suppose that, through his choices, the agent reveals a preference for a'' over a', at a given realization of the exogenous variable. This can be a direct revelation in the sense that a'' was chosen when a' was feasible at some observation, or it could be revealed indirectly via transitive closure (for example, if a'' was chosen when b was available at some observation and b was chosen when a' was available at another observation). ARC says the following: if the agent reveals a preference for a'' over a' with $a' < a''$ when $\xi_i = \beta'$ then the agent cannot reveal a preference for a' over a'' at some $\xi_i = \beta'' > \beta'$. We show that any data set collected from an agent choosing with a preference obeying strict interval dominance must obey ARC and any data set that obeys ARC is rationalizable by a preference obeying strict single crossing differences.[1]

---

[1]Readers familiar with Afriat's Theorem may notice a parallel in the following sense: the general axiom of revealed preference (GARP) is necessary whenever the consumer is maximizing a locally nonsatiated preference while GARP is sufficient to guarantee a *stronger* conclusion: that there is a continuous, strictly increasing, and concave function rationalizing the data. In our case, ARC is necessary for strict interval dominance and sufficient for strict single

1.2 Games with strategic complementarity

An important application of monotone comparative statics is to the study of games with strategic complementarity (see Milgrom and Roberts (1990) and Vives (1990)). These are games where players' strategies are complements in the sense that an agent's best response increases with the action of other players in the game. These games are known to be very well-behaved: they always have pure strategy Nash equilibria; in fact, there is always a largest and a smallest pure strategy Nash equilibrium and a parameter change that leads to one agent having a greater best response will raise both the largest and smallest equilibrium.

In this context it is natural to pose a revealed preference question analogous to the one we posed in the single agent case. For each player i ($i = 1, 2, ..., n$), suppose we observe the feasible action set A_i^t (assumed to be a compact interval), the action chosen by the player, $a_i^t \in A_i^t$, and an exogenous variable y_i^t (drawn from a poset) that affects player i's action. An observation t may be succinctly written as (a^t, y^t, A^t) (where $a^t = (a_i^t)_{i=1}^n$, etc.) such that a^t is the observed action profile in the treatment (y^t, A^t) and the data set is $\mathcal{O} = \{(a^t, y^t, A^t)\}_{t=1}^T$. Then we can ask whether the data set is consistent with the hypothesis that the observations constitute Nash equilibria in games with strategic complementarity. Notice that this hypothesis is at least internally consistent since we know that these games always have pure strategy Nash equilibria. The answer to our question is straightforward given the single-agent results: all we need to do is to check that each player's choices obey ARC, in the sense that, for all i, $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t=1}^T$, where $\xi_i = (a_{-i}^t, y_i^t)$, obeys ARC. (From player i's perspective, the variables affecting his preference are the realized value of y_i and the actions of other players.)

When the data set $\mathcal{O}$ obeys ARC (in the sense that every player obeys ARC), it would be natural to exploit this data to make predictions of the outcome in a new game, with different feasible action sets $A^0 = (A_i^0)_{i=1}^n$ and different exogenous variables $y^0 = (y_i^0)_{i=1}^n$, assuming that the players' preferences obey single crossing differences and remain unchanged. We provide a procedure for working out the set of all possible Nash equilibria in this new game. We also show that this set has properties that echo those of a set of Nash equilibria in a game with stratgic complementarity: while the set itself may not have a largest or smallest element, its closure does have a largest and

crossing differences (which is a stronger property).

a smallest element and these extremal elements increase with y^0.

1.3 Revealed preference tests on cross sectional data

So far we have considered an observer who records the behavior of an agent or a group of agents across a sequence of different treatments. It is not always possible to obtain data of this type in empirical settings. Suppose instead that, at each treatment, we observe the joint actions taken by a large population of n-player groups. Formally, the data set is $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t=1}^T$, where μ^t is a distribution on A^t. In this case, the natural generalization of our notion of rationalization is to require that the population can be decomposed into segments such that all groups within a segment have the same equilibrium play across treatments and the equilibrium play is consistent with strategic complementarity. This rationalization concept captures the idea that treatments have been randomly assigned across the whole population of groups, so that the distribution of 'group types' is the same across treatments; it allows for preference heterogeneity in the population but requires every group type to be consistent with strategic complementarity. We show that it is possible to check whether $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t=1}^T$ is consistent with strategic complementarity in this sense by solving a certain system of linear equations. When a data set passes this test, we provide a procedure to estimate the distribution of equilibrium responses in the population under a new treatment, again by solving an appropriate linear program.

1.4 Application: spousal influence in smoking behavior

To illustrate the use of our techniques, we apply them to investigate whether spouses influence each other in their cigarette smoking behavior. The US census provides information on tobacco use in married couples and smoking policies at their workplaces (whether it is permitted or not).[2] The latter plays the role of the exogenous variable in our analysis and couples are modeled as playing a 2×2 game, where the action is either 'not smoke' or 'smoke'. Strategy sets of both players do not vary in this application and so there are precisely four treatments, corresponding to the different combinations of workplace policies for the couple. The hypothesis is that couples in the population

---

[2]The data is taken from the years 1992-93, when there were still significant numbers of workplaces that permitted smoking.

are playing games of strategic complementarity, where a husband's (wife's) smoking decision is nondecreasing with respect to the spouse's smoking behavior and the workplace smoking policy (with the ordering being the intuitive one). Under each treatment, there are four pure-strategy outcomes, so there are $4^4 = 256$ ways a couple could vary its behavior across the four treatments. It can be shown that 64 of these are consistent with strategic complementarity, so the hypothesis is that the entire population consists of groups belonging to one of these 64 types. We show that the data set does not pass the rationalizability test exactly; however, the failure is not statistically significant, so the strategic complementarity hypothesis cannot be rejected.

1.5 Literature Review

Topkis (1998, Theorem 2.8.9) reports a revealed preference-type result in a monotone choice environment. He considers a correspondence $\varphi : T \to \mathbb{R}^\ell$ that maps elements of a totally ordered set T to compact sublattices of the Euclidean space $\mathbb{R}^l$. He shows that this correspondence is increasing in the strong set order if and only if there is a function $f : \mathbb{R}^\ell \times T \to \mathbb{R}$ such that $\varphi(t) = \arg\max_{x \in \mathbb{R}^\ell} f(x, t)$ where f is supermodular in x and has increasing differences in (x, t). Notice that the rationalizability concept used by Topkis is more stringent than the one we employ since the optimal choices under f must coincide with (rather than simply contain) $\varphi(t)$. In the case where φ is a choice *function*, it is not hard to see that such a rationalization is possible even when T is a partially (rather than totally) ordered set; this has been noted by Carvajal (2004) who also applies it to a game setting. In our paper, we confine ourselves to the case where actions are totally ordered (rather than elements of a Euclidean space) and allow observations of choices made from *different subsets* of the set of all possible actions. Consequently, at a given parameter value, the observer may have partial information on the agent's ranking over different actions rather than simply the globally optimal action. In this respect, the problem is more complicated than the one posed by Topkis, because the rationalizing preference we construct has to agree with this wider range of preference information (in addition to obeying single crossing differences).

The extension of our revealed preference tests to cross-sectional data sets with unobserved heterogeneity follows an approach that has been taken by other authors (see McFadden and Richter (1991), McFadden (2005), and Manski (2007)). Manski (2007) also discusses making predictions

in unobserved treatments, subject to a particular theory of behavior, and our approach to this issue is in essence the same as his. (He did not, however, consider the specific theory relating to single crossing differences discussed here). Echenique and Komunjer (2009) develop a structural model that could be used to test for strategic complementarity in certain special classes of games, including two person games. Their test relies on a stochastic equilibrium selection rule that places strictly positive probability on the extremal elements of the set of Nash equilibria and checks certain observable properties implied by strategic complementarity; the sufficiency of those properties (for rationalizability) is not addressed. Aradillas-Lopez (2011) provides nonparametric probability bounds for Nash equilibrium actions for a class of games with characteristics that are similar to, but distinct from, games with strategic complementarity. There are also papers where actions are assumed to be strategic complements or substitutes in order to sharpen inference or predictions of one type or another. For example, Kline and Tamer (2012) employ such assumptions to provide nonparametric bounds for best-replies in the context of binary games; other papers of this type include Molinari and Rosen (2008), Uetake and Watanabe (2013), and Lazzati (2015). By and large, the emphasis in these papers is not to test for strategic complementarity but to exploit it as an assumption; indeed the model may not include the type of exogenous treatment variation that makes the assumption refutable.

For our empirical implementation at the end of the paper, we test for strategic complementarity in smoking behavior among married couples by taking advantage of the variation in workplace smoking policies. Cutler and Glaeser (2010) also exploit this variation for essentially the same purpose but their work differs from ours in that they use a reduced form parametric model of smoking behavior; like us, they find evidence of complementarity in smoking behavior among married couples. While our theoretical results on testing for complementarity are developed in an idealized setting where population distributions are known, we must necessarily account for sampling variation in the application; for this we rely on the econometric procedure devised by Kitamura and Stoye (2013). Those authors use their procedure to implement the test for the strong axiom of revealed preference on cross sectional data sets (as developed by McFadden and Richter (1991)), but it applies equally well to our model and to others with a similar structure (such as those discussed in Manski (2007)).

2 Basic concepts and theory

Our objective in this section is to give a quick review of some basic concepts and results in monotone comparative statics and of their application to games with strategic complementarities. This will motivate the revealed preference theory developed later in the paper.

2.1 Monotone choice on intervals

Let $X_i \subset \mathbb{R}$ be the set of all conceivable actions of an agent i. A *feasible action set* of agent i is a subset A_i of X_i. We assume that A_i is compact in $\mathbb{R}$ and that it is an *interval of X_i*. We say that a set $A_i \subseteq X_i$ is an *interval of X_i* if, whenever x'', $x' \in A_i$, with $x'' > x'$, then, for any element $\tilde{x} \in X_i$ such that $x'' > \tilde{x} > x'$, $\tilde{x} \in A_i$. Given that A_i is both compact and an interval, we can refer to it as a *compact interval*. It is clear that there must be $\underline{a}_i$ and $\bar{a}_i$ in A_i such that $A_i = \{x_i \in X_i : \underline{a}_i \leqslant x_i \leqslant \bar{a}_i\}$ and it is sometimes convenient to denote A_i by $[\underline{a}_i, \bar{a}_i]$. We denote by $\mathcal{A}_i$ the collection of all compact intervals of X_i. We assume that agent i's choice over different actions in a feasible action set A_i is affected by a parameter ξ_i, where ξ_i is drawn from a partially ordered set (or poset, for short) $(\Xi_i, \geqslant)$; ξ_i may include certain exogenous variables and/or the actions of other agents (when we extend the analysis to a game). For the sake of notational simplicity, we are using the same notation for the orders on X_i and Ξ_i and for any other ordered sets; we do not anticipate any danger of confusion.

A binary relation $\succsim_i$ on $X_i \times \Xi_i$ is said to be a *preference* of agent i if, for every fixed $\xi_i \in \Xi_i$, $\succsim_i$ is a complete, reflexive and transitive relation on X_i. We call a preference $\succsim_i$ *regular* if, for all $A_i \in \mathcal{A}_i$ and ξ_i, the set $\mathrm{BR}_i(\xi_i, A_i, \succsim_i)$ (which we may shorten to $\mathrm{BR}_i(\xi_i, A_i)$ when there is no danger of confusion), as defined by (1), is nonempty and compact in $\mathbb{R}$. We refer to $\mathrm{BR}_i(\xi_i, A_i)$ as agent i's *best response* or *optimal choice* at (ξ_i, A_i). The best response of agent i is said to be *monotone* or *increasing* in ξ_i if, for every $\xi_i'' > \xi_i'$,

$$a_i'' \in \mathrm{BR}_i(\xi_i'', A_i) \text{ and } a_i' \in \mathrm{BR}_i(\xi_i', A_i) \implies a_i'' \geqslant a_i'. \tag{2}$$

The preference $\succsim_i$ is said to obey *strict interval dominance* (SID) if, for every $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$,

$$(x_i'', \xi_i') \succsim_i (x_i, \xi_i') \text{ for all } x \in [x_i', x_i''] \Longrightarrow (x_i'', \xi_i'') \succ_i (x_i', \xi_i''), \tag{3}$$

where $\succ_i$ is the asymmetric part of $\succsim_i$, i.e., $(x_i, \xi_i) \succ_i (y_i, \xi_i)$ if $(x_i, \xi_i) \succsim_i (y_i, \xi_i)$ and $(y_i, \xi_i) \not\succsim_i (x_i, \xi_i)$. We denote the symmetric part of $\succsim_i$ by $\sim_i$, i.e., $(x_i, \xi_i) \sim_i (y_i, \xi_i)$ if $(x_i, \xi_i) \succsim_i (y_i, \xi_i)$ and $(y_i, \xi_i) \succsim_i (x_i, \xi_i)$. The following result is a straightforward adaptation of Theorem 1 in Quah and Strulovici (2009). We shall re-prove it here because of its central role in this paper.

THEOREM A. *Suppose $\succsim_i$ is a regular preference on $X_i \times \Xi_i$. Then agent i has a monotone best response correspondence if and only if $\succsim_i$ obeys strict interval dominance.*

Proof. To show that $\succsim_i$ obeys SID, suppose that, for some $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$, the left side of (3) holds. Letting $A_i = [x_i', x_i'']$, we obtain $x_i'' \in \mathrm{BR}_i(\xi_i', A_i)$. Hence, by (2), it also holds that $x_i'' \in \mathrm{BR}_i(\xi_i'', A_i)$. If $(x_i'', \xi_i'') \sim_i (x_i', \xi_i'')$ were to hold, then $x_i' \in \mathrm{BR}_i(\xi_i'', A_i)$. However, then we have that $x_i'' \in \mathrm{BR}_i(\xi_i', A_i)$, $x_i' \in \mathrm{BR}_i(\xi_i'', A_i)$, and $x_i' < x_i''$, which contradicts (2). Therefore, $(x_i'', \xi_i'') \succ_i (x_i', \xi_i'')$. Conversely, suppose $\xi_i'' > \xi_i'$, $x_i'' \in \mathrm{BR}_i(\xi_i'', A_i)$ and $x_i' \in \mathrm{BR}_i(\xi_i', A_i)$. If $x_i'' < x_i'$, then $(x_i', \xi_i') \succsim_i (x_i, \xi_i')$ for every $x_i \in [x_i'', x_i'] \subseteq A_i$. SID guarantees that $(x_i', \xi_i'') \succ_i (x_i'', \xi_i'')$, which contradicts the assumption that $x_i'' \in \mathrm{BR}_i(\xi_i'', A_i)$. $\square$

Readers familiar with the standard theory of monotone comparative statics will notice that our definition of monotonicity in (2) is stronger than the standard notion, which merely requires that $\mathrm{BR}_i(\xi_i'', A_i)$ dominates $\mathrm{BR}_i(\xi_i', A_i)$ in the *strong set order*. This means that, for any $x_i'' \in \mathrm{BR}_i(\xi_i'', A_i)$ and $x_i' \in \mathrm{BR}_i(\xi_i', A_i)$, $\max\{x_i'', x_i'\} \in \mathrm{BR}_i(\xi_i'', A_i)$ and $\min\{x_i'', x_i'\} \in \mathrm{BR}_i(\xi_i', A_i)$. In turn, this weaker notion of monotonicity can be characterized by preferences obeying *interval dominance* (rather than strict interval dominance), which is defined as follows: for every $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$,

$$(x_i'', \xi_i') \succsim_i (\succ_i) (x_i, \xi_i') \text{ for every } x_i \in [x_i', x_i''] \Longrightarrow (x_i'', \xi_i'') \succsim_i (\succ_i) (x_i', \xi_i''). \tag{4}$$

(The reader can verify this claim by a straightforward modification of the proof of Theorem A or by consulting Theorem 1 in Quah and Strulovici (2009).) Throughout this paper we have chosen to work with a stronger notion of monotonicity; the weaker notion does not permit meaningful revealed

preference analysis because it does not exclude the possibility that an agent is simply indifferent to all actions at every ξ_i. In this sense, our stronger assumption here is analogous to the assumption of local non-satiation made in Afriat's Theorem.[3]

The interval dominance order is Quah and Strulovici's (2009) generalization of single crossing differences, due to Milgrom and Shannon (1994). Just as there is strict interval dominance, so there is a strict version of single crossing differences. We say that a preference relation $\succsim_i$ has *strict single crossing differences (SSCD)* if, for every $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$,

$$(x_i'', \xi_i') \succsim_i (x_i', \xi_i') \Longrightarrow (x_i'', \xi_i'') \succ_i (x_i', \xi_i''). \tag{5}$$

It is clear that every preference that obeys SSCD will also satisfy SID. Hence, it is obvious from Theorem A that if $\succsim_i$ is a regular preference on $X_i \times \Xi_i$ that obeys SSCD, then agent i has a monotone best response correspondence $\mathrm{BR}_i(\xi_i, A_i)$ for every interval $A_i \in \mathcal{A}_i$.[4]

2.2 Strategic complementarity

An important application of monotone comparative statics is to the study of games with strategic complementarity. Let $N = \{1, 2, ..., n\}$ be the set of agents in a game, and let $X_i \subset \mathbb{R}$ be the set of all conceivable actions of agent i. We assume that i has a feasible action set A_i that is a compact interval of X_i; as before, the family of compact intervals of X_i is denoted by $\mathcal{A}_i$. Agent i's choice over different feasible actions is affected by the actions of other players and also by an exogenous variable y_i, which we assume is drawn from a poset $(Y_i, \geqslant)$. Let $\Xi_i = X_{-i} \times Y_i$, where $X_{-i} := \times_{j \neq i} X_j$. A typical element of Ξ_i is denoted by $\xi_i = (x_{-i}, y_i)$ and Ξ_i is a poset if we endow it with the product order. We assume that agent i has a preference $\succsim_i$ on $X_i \times \Xi_i$, in the sense defined in Section 2.1.

Given a profile of regular preferences $\{\succsim_i\}_{i \in N}$, a *joint feasible action set* $A \in \mathcal{A} = \times_{i \in N} \mathcal{A}_i$, and

---

[3]It is clear that without such an assumption, any type of consumption data is rationalizable since one could simply suppose that the consumer is indifferent across all consumption bundles. For a statement and proof of Afriat's Theorem see Varian (1982).

[4]In fact, SSCD of a preference ensures more than that: it is necessary and sufficient for the monotonicity of a best response correspondence on *arbitrary* feasible action sets and not only interval feasible action sets. On the relationship between single crossing differences and the interval dominance order, see Quah and Strulovici (2009).

a profile of exogenous variables $y \in Y = \times_{i \in N} Y_i$, we can define a game

$$\mathcal{G}(y, A) = \left[(y_i)_{i \in N}, (A_i)_{i \in N}, (\succsim_i)_{i \in N} \right].$$

We say that the family of games $\mathbb{G} = \{\mathcal{G}(y, A)\}_{(y, A) \in Y \times \mathcal{A}}$ exhibits *strategic complementarity* if, for every $A \in \mathcal{A}$, the best response of each agent i (as given by (1)) is monotone in $\xi_i = (x_{-i}, y_i)$. It is clear from Theorem A that *the family of games* $\mathbb{G} = \{\mathcal{G}(y, A)\}_{(y, A) \in Y \times \mathcal{A}}$ *exhibits strategic complementarity if and only if* $\succsim_i$ *is an SID preference for every agent* i.

EXAMPLE 1. Consider a Bertrand oligopoly with n firms, with each firm producing a single differentiated product. Assume that firm i has constant marginal cost $c_i > 0$, faces the demand function $D_i(p_i, p_{-i}) : \mathbb{R}_{++} \times \mathbb{R}_{++}^{n-1} \to \mathbb{R}_+$, and chooses its price $p_i > 0$ to maximize profit $\Pi_i(p_i, p_{-i}, c_i) = (p_i - c_i) D_i(p_i, p_{-i})$. Suppose that the firms' products are substitutes in the sense that the own-price elasticity of demand,

$$-\frac{p_i}{D_i(p_i, p_{-i})} \frac{\partial D_i}{\partial p_i}(p_i, p_{-i})$$

is strictly falling with respect to p_{-i} (the prices charged by other firms). Then, the profit of each firm has SSCD in $(p_i; p_{-i}, c_i)$. Hence, on any compact interval of prices, firm i's set of profit-maximizing prices is monotone in (p_{-i}, c_i).[5] If this property holds for every firm in the industry, the collection of Bertrand games generated by different feasible price sets to each firm and different exogenous variables, $c = (c_i)_{i \in N}$, will constitute a collection of games exhibiting strategic complementarity.

It is known that the set of Nash equilibria of a game with strategic complementarity (even in the weaker sense of best responses increasing in the strong set order) is particularly well-behaved. The following result summarizes some of its properties. For our purposes, the most important feature of these games is that they always have pure strategy Nash equilibria, so it is not a priori unreasonable to hypothesize that players are playing a pure strategy Nash equilibrium in such a game.

THEOREM B. *Suppose* $\mathbb{G} = \{\mathcal{G}(y, A)\}_{(y, A) \in Y \times \mathcal{A}}$ *exhibits strategic complementarity. Then, for every game* $\mathcal{G}(y, A) \in \mathbb{G}$, *the set of pure strategy Nash equilibria* $E(y, A)$ *is a nonempty complete lattice*

---

[5]Specifically, they guarantee that for any $p_i'' > p_i'$, $\ln \Pi(p_i'', p_{-i}, c_i) - \ln \Pi(p_i', p_{-i}, c_i)$ is strictly increasing in (p_{-i}, c_i), which implies SSCD (see, Milgrom and Shannon (1994)).

and, in particular, it has a largest and a smallest Nash equilibrium. Furthermore, both the largest and smallest Nash equilibria are increasing in y.

The set of Nash equilibria of $\mathcal{G}(y, A)$ coincides with the fixed points of the joint best response correspondence $\mathrm{BR}(\cdot, y, A) : A \rightrightarrows A$, where, denoting (x_{-i}, y_i) by ξ_i,

$$\mathrm{BR}(x, y, A) = (\mathrm{BR}_1(\xi_1, A_1), \mathrm{BR}_2(\xi_2, A_2), ..., \mathrm{BR}_n(\xi_n, A_n)).$$

Both the non-emptiness and structure of $E(y, A)$ flow from the fact that this is a very well-behaved correspondence. Indeed, under strategic complementarity, $\mathrm{BR}_i(\xi_i, A_i)$ is increasing in ξ_i (in the sense of (2), for all i) and so $\mathrm{BR}(x, y, A)$ is increasing in (x, y).[6]

3   Revealed monotone choice

Consider an observer who collects a finite data set from agent i, where each observation consists of the action chosen by the agent, the set of feasible actions, and the value of the parameter. Formally, the data set is $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$, where $\mathcal{T} = \{1, 2, ..., T\}$. This means that, at observation t, the agent is subjected to the *treatment* $(\xi_i^t, A_i^t) \in \Xi_i^t \times \mathcal{A}_i$ and chooses the action $a_i^t \in A_i^t$. We say that $\mathcal{O}_i$ (or simply, agent i) is *consistent with monotonicity* or *monotone-rationalizable* if there is a regular and SID preference $\succsim_i$ on $X_i \times \Xi_i$ such that for every $t \in \mathcal{T}$, $(a_i^t, \xi_i^t) \succsim_i (x_i, \xi_i^t)$ for every $x_i \in A_i^t$. The motivation for this definition is clear given Theorem A: if $\mathcal{O}_i$ is monotone-rationalizable then we have found a preference that (i) accounts for the observed behavior of the agent and (ii) guarantees that the agent's optimal choice based on this preference is increasing in ξ_i, on any feasible action set that is an interval. Our principal objective in this section is to characterize those data sets that are monotone-rationalizable.

---

[6]See Topkis (1998) for the proof of Theorem B. The complete lattice structure of $E(A, y)$ was first pointed out in Zhou (1994), and the monotone comparative statics of extremal equilibria is found in Milgrom and Roberts (1990).

3.1 The axiom of revealed complementarity

We first introduce the revealed preference relations induced by $\mathcal{O}_i$. The *direct revealed preference* relation $\succsim_i^R$ is defined in the following way: $(x_i'', \xi_i) \succsim_i^R (x_i', \xi_i)$ if $(x_i'', \xi_i) = (a_i^t, \xi_i^t)$ and $x_i' \in A_i^t$ for some $t \in \mathcal{T}$. The *indirect revealed preference* relation $\succsim_i^{RT}$ is the transitive closure of $\succsim_i^R$, i.e., $(x_i'', \xi_i) \succsim_i^{RT} (x_i', \xi_i)$ if there exists a finite sequence $z_i^1, z_i^2, ..., z_i^k$ in X_i such that

$$(x_i'', \xi_i) \succsim_i^R (z_i^1, \xi_i) \succsim_i^R (z_i^2, \xi_i) \succsim_i^R \ ... \ \succsim_i^R (z_i^k, \xi_i) \succsim_i^R (x_i', \xi_i). \tag{6}$$

The motivation for this terminology is clear. If we observe, at some treatment (ξ_i, A_i), agent i playing x_i'' when $x_i' \in A_i$, then it must be the case that $(x_i'', \xi_i) \succsim_i (x_i', \xi_i)$ if agent i is optimizing with respect to the preference $\succsim_i$. Furthermore, given that $\succsim_i$ is transitive, if $(x_i'', \xi_i) \succsim_i^{RT} (x_i', \xi_i)$ then $(x_i'', \xi_i) \succsim_i (x_i', \xi_i)$.[7]

A relation $\mathcal{R}$ on $X_i \times \Xi_i$ said to have the *interval property* if, whenever $(x_i, \xi_i) \, \mathcal{R} \, (\tilde{x}_i, \xi_i)$, for x_i, $\tilde{x}_i$ in X_i, then $(x_i, \xi_i) \, \mathcal{R} \, (z_i, \xi_i)$ for any z_i between x_i and $\tilde{x}_i$, i.e., $x_i < z_i < \tilde{x}_i$ or $\tilde{x}_i < z_i < x_i$. This property plays an important role in our results. The lemma below uses the assumption that feasible action sets are compact intervals to guarantee that $\succsim_i^{RT}$ has the interval property.

LEMMA 1. *The relation $\succsim_i^{RT}$ in $X_i \times \Xi_i$ induced by $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t=1}^T$ has the interval property.*

Proof. If $(x_i'', \xi_i) \succsim_i^R (x_i', \xi_i)$, then there is A_i^t such that $x_i'' = a_i^t$ and $x_i' \in A_i^t$. Since A_i^t is an interval, it is clear that $(x_i'', \xi_i) \succsim_i^R (x_i, \xi_i)$ for any x_i between x_i'' and x_i'. Now suppose $(x_i'', \xi_i) \succsim_i^{RT} (x_i', \xi_i)$, but $(x_i'', \xi_i) \not\succsim_i^R (x_i', \xi_i)$. Then, we have a sequence like (6). Suppose also that $x_i'' > x_i'$ and consider x_i such that $x_i'' > x_i > x_i'$. (The case where $x_i'' < x_i'$ can be handled in a similar way.) Letting $z_i^0 = x_i''$ and $z_i^{k+1} = x_i'$, we know that there exists at least one $0 \leqslant m \leqslant k$ such that $z_i^m \geqslant x_i \geqslant z^{m+1}$. Since $(z_i^m, \xi_i) \succsim_i^R (z_i^{m+1}, \xi_i)$, it must hold that $(z_i^m, \xi_i) \succsim_i^R (x_i, \xi_i)$. This in turn implies that $(x_i'', \xi_i) = (z_i^0, \xi_i) \succsim_i^{RT} (x_i, \xi_i)$, since $(z_i^0, \xi_i) \succsim_i^{RT} (z_i^m, \xi_i)$. $\qquad\square$

DEFINITION 1. *The data set $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t=1}^T$ obeys the* Axiom of Revealed Complementarity

---

[7]Note, however, that $\succsim_i^R$ and $\succsim_i^{RT}$ are not generally complete on X_i for every fixed ξ_i; as such, these relations are not preferences as we have defined them.

(ARC) *if, for every $s, t \in \mathcal{T}$,*

$$\xi_i^t > \xi_i^s, \ a_i^t < a_i^s, \ \text{and} \ (a_i^s, \xi_i^s) \succsim_i^{RT} (a_i^t, \xi_i^s) \implies (a_i^t, \xi_i^t) \nsucceq_i^{RT} (a_i^s, \xi_i^t). \tag{7}$$

It is clear that ARC is a non-vacuous restriction on data. So long as the number of observations $\mathcal{O}_i$ is finite (as it is by assumption), checking whether two elements (a_i^s, ξ_i^s) and (a_i^t, ξ_i^s) are related by $\succsim_i^{RT}$ is a finite procedure and, consequently, so is checking for ARC. It is also clear that there are no computational difficulties, whether theoretical or practical, associated with the implementation of this test.

For a data set to obey monotone-rationalizability, it is *necessary* that it obeys ARC. Indeed, suppose there are observations s and t such that $\xi_i^t > \xi_i^s$, $a_i^t < a_i^s$, and $(a_i^s, \xi_i^s) \succsim_i^{RT} (a_i^t, \xi_i^s)$. By Lemma 1, $\succsim_i^{RT}$ has the interval property, and so $(a_i^s, \xi_i^s) \succsim_i^{RT} (x_i, \xi_i^s)$ for all $x_i \in [a_i^t, a_i^s]$. Since $\mathcal{O}_i$ is SID-rationalizable, there is an SID preference $\succsim_i$ on $X_i \times \Xi_i$ such that $(a_i^s, \xi_i^s) \succsim_i (x_i, \xi_i^s)$ for all $x_i \in [a_i^t, a_i^s]$. The SID property on $\succsim_i$ guarantees that $(a_i^s, \xi_i^t) \succ_i (a_i^t, \xi_i^t)$, which means $(a_i^t, \xi_i^t) \nsucceq_i^{RT} (a_i^s, \xi_i^t)$.

Our more substantial claim is that ARC is also *sufficient* for monotone-rationalizability. In fact, an even stronger property is true: whenever a data set obeys ARC, it is rationalizable by an SSCD (and not just SID) preference.[8] The next result summarizes our main findings.

THEOREM 1. *The following statements on the data set $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ are equivalent:*

(a) $\mathcal{O}_i$ *is monotone-rationalizable.*

(b) $\mathcal{O}_i$ *obeys ARC.*

(c) $\mathcal{O}_i$ *is rationalizable by a regular and SSCD preference relation on $X_i \times \Xi_i$.*

Since every SSCD preference is also an SID preference, (c) implies (a), and we have just shown that (a) implies (b). It remains for us to show that (b) implies (c). Our proof involves first working

---

[8]This phenomenon, which may seem surprising, is not unknown to revealed preference analysis; for example, it is present in Afriat's Theorem. In that context, the data consist of observations of consumer's consumption bundles at different linear budget sets. If the agent is maximizing a locally non-satiated preference, then the data set must obey a property called the generalized axiom of revealed preference (GARP, for short); conversely, if a data set obeys GARP then it can be rationalized by a preference that is not just locally non-satiated but also obeys continuity, strong monotonicity, and convexity.

out the (incomplete) revealed preference relations on $X_i \times \Xi_i$ that *must* be satisfied by any SID preference that rationalizes the data and then explicitly constructing a rationalizing preference on $X_i \times \Xi_i$ that completes that incomplete relation and obeys SSCD.

Given $\mathcal{O}_i$, the *single crossing extension* of the indirect revealed preference relation $\succsim_i^{RT}$ is another binary relation $\succ_i^{RTS}$ defined in the following way:

(i) for $x_i'' > x_i'$, $(x_i'', \xi_i) \succ_i^{RTS} (x_i', \xi_i)$ if there is $\xi_i' < \xi_i$ such that $(x_i'', \xi_i') \succsim_i^{RT} (x_i', \xi_i')$;

(ii) for $x_i'' < x_i'$, $(x_i'', \xi_i) \succ_i^{RTS} (x_i', \xi_i)$, if there is $\xi_i'' > \xi_i$ such that $(x_i'', \xi_i'') \succsim_i^{RT} (x_i', \xi_i'')$.

Let $\succsim_i^{RTS}$ be the binary relation given by $\succsim_i^{RTS} = \succsim_i^{RT} \cup \succ_i^{RTS}$. (Note that $\succ_i^{RTS}$ is *not* the asymmetric part of $\succsim_i^{RTS}$.) It follows immediately from its definition that $\succsim_i^{RTS}$ also has strict single crossing differences, in the following sense: if $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$ *or* $x_i'' < x_i'$ and $\xi_i'' < \xi_i'$, then

$$(x_i'', \xi_i') \succsim_i^{RTS} (x_i', \xi_i') \implies (x_i'', \xi_i'') \succ_i^{RTS} (x_i', \xi_i''). \tag{8}$$

In addition, let $\succsim_i^{RTST}$ be the transitive closure of $\succsim_i^{RTS}$, i.e., $(x_i'', \xi_i) \succsim_i^{RTST} (x_i', \xi_i)$ if there exists a sequence $z_i^1, z_i^2, ..., z_i^k$ such that

$$(x_i'', \xi_i) \succsim_i^{RTS} (z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^k, \xi_i) \succsim_i^{RTS} (x_i', \xi_i). \tag{9}$$

If we can find at least one strict relation $\succ_i^{RTS}$ in the sequence (9), then, we let $(x_i'', \xi_i) \succ_i^{RTST} (x_i', \xi_i)$ (which, once again, is *not* the asymmetric part of $\succsim_i^{RTST}$). The relevance of the binary relations $\succsim_i^{RTST}$ and $\succ_i^{RTST}$ flows from the following result, which says that any rationalizing preference for agent i must respect the ranking implied by them.

PROPOSITION 1. *Suppose that the preference $\succsim_i$ obeys SID and rationalizes $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$. Then $\succsim_i$ extends $\succsim_i^{RTST}$ and $\succ_i^{RTST}$ in the following sense:*

$$(x_i'', \xi_i) \succsim_i^{RTST} (\succ_i^{RTST}) (x_i', \xi_i) \implies (x_i'', \xi_i) \succsim_i (\succ_i) (x_i', \xi_i) \tag{10}$$

Proof. Without loss of generality, we may let $x_i'' > x_i'$. Since $\succsim_i$ is transitive, it is clear that we need only show that $(x_i'', \xi_i) \succsim_i (\succ_i) (x_i', \xi_i)$ whenever $(x_i'', \xi_i) \succsim_i^{RTS} (\succ_i^{RTS}) (x_i', \xi_i)$. If

15

$(x_i'', \xi_i) \succsim_i^{RTS} (\succ_i^{RTS}) (x_i', \xi_i)$ then there exists some $\xi_i' \leqslant (<) \xi_i$ such that $(x_i'', \xi_i') \succsim_i^{RT} (x_i', \xi_i')$. By the interval property of $\succsim_i^{RT}$, we obtain $(x_i'', \xi_i') \succsim_i^{RT} (x_i, \xi_i')$ for all $x_i \in [x_i', x_i'']$. Since $\succsim_i$ rationalizes $\mathcal{O}_i$, we also have $(x_i'', \xi_i') \succsim_i (x_i, \xi_i')$ for all $x_i \in [x_i', x_i'']$. By SID of $\succsim_i$, we obtain $(x_i'', \xi_i) \succsim_i (\succ_i) (x_i', \xi_i)$ for $\xi_i' \leqslant (<) \xi_i$. $\qquad\square$

At this point, it is reasonable to ask if we could go beyond the revealed preference relations we have already constructed and consider the single crossing extension of $\succsim_i^{RTST}$, the transitive closure of that extension, and so on. The answer to that is 'no' because, as we shall show in Lemma 2, $\succsim_i^{RTST}$ obeys SSCD, so it does not admit a nontrivial single crossing extension. By Proposition 1, it is clear that, in order for $\mathcal{O}_i$ to be monotone rationalizable, the binary relation $\succsim_i^{RTST}$ must have the following property: for any (x_i', ξ_i) and (x_i'', ξ_i) in $X_i \times \Xi_i$,

$$(x_i', \xi_i) \succsim_i^{RTST} (x_i'', \xi_i) \implies (x_i'', \xi_i) \not\succ_i^{RTST} (x_i', \xi_i). \tag{11}$$

If not, we obtain simultaneously, $(x_i', \xi_i) \succsim_i (x_i'', \xi_i)$ and $(x_i'', \xi_i) \succ_i (x_i', \xi_i)$, which is impossible.

The following lemma summarizes our observations on $\succsim_i^{RTST}$.

LEMMA 2. *Suppose that $\mathcal{O}_i$ obeys ARC. Then $\succsim_i^{RTST}$ obeys SSCD and property (11).*

Since $\succsim_i^R \subseteq \succsim_i^{RTST}$, it is clear that Proposition 1 has a converse: if there is a regular and SID preference $\succsim_i$ on $X_i \times \Xi_i$ that obeys (10), then this preference rationalizes $\mathcal{O}_i$. This observation, together with Lemma 2, suggest that a reasonable way of constructing a rationalizing preference is to *begin with* $\succsim_i^{RTST}$ and $\succ_i^{RTST}$ and then complete these incomplete relations in a way that gives a preference with the required properties, which is precisely the approach we take. Define the binary relation $\succsim_i^*$ on $X_i \times \Xi_i$ in the following manner:

$$\begin{aligned}(x_i'', \xi_i) \succsim_i^* (x_i', \xi_i) \text{ if } &(x_i'', \xi_i) \succsim_i^{RTST} (x_i', \xi_i) \\ &\text{or } (x_i'', \xi_i) \parallel_i^{RTST} (x_i', \xi_i) \text{ and } x_i' \geqslant x_i'',\end{aligned} \tag{12}$$

where $(x_i'', \xi_i) \parallel_i^{RTST} (x_i', \xi_i)$ means neither $(x_i'', \xi_i) \succsim_i^{RTST} (x_i', \xi_i)$ nor $(x_i', \xi_i) \succsim_i^{RTST} (x_i'', \xi_i)$. The following result (which we prove in the Appendix with the help of Lemma 2) completes our argument that (b) implies (c) in Theorem 1.

LEMMA 3. *Suppose that $\mathcal{O}_i$ obeys ARC. The binary relation $\succsim_i^*$ is an SSCD preference that ratio-nalizes $\mathcal{O}_i$. On every set $K \subset X_i$ that is compact in $\mathbb{R}$ and for every $\xi_i \in \Xi_i$, $\mathrm{BR}_i(\xi_i, K, \succsim_i^*)$ is nonempty and finite; in particular, $\succsim_i^*$ is a regular preference*

3.2 ARC and SSCD

Theorem 1 tells us that when an agent has an SID preference, then any data set collected from this agent must obey ARC. It also says that if a data set obeys ARC, then the agent's actions can be accounted for by an SID preference, and moreover, we can explicitly construct a preference consistent with those observations that obey the *stronger* property of SSCD. It is known that SSCD is sufficient (and, in fact, also necessary in some sense) for an agent's optimal action to be increasing with the parameter ξ_i on all *arbitrary* constraint sets drawn from X_i (see Edlin and Shannon (1998)). It follows that when a data set $\mathcal{O}_i$ is monotone-rationalizable, we can find a preference that both explains the data and guarantees that the optimal choices based on this preference is monotone, on any arbitrary feasible action set (and not just intervals).

So far we have maintained the assumption that the observed feasible action sets A_i^t are intervals. Now consider a data set $\mathcal{O}_i = \{(a_i^t, \xi_i^t, B_i^t)\}_{t \in \mathcal{T}}$, where a_i^t is the observed choice from B_i^t, and B_i^t is a compact subset of X_i that is not necessarily an interval. It is easy to check that if $\mathcal{O}_i$ is rationalizable by an SSCD preference then it must obey ARC and, given the characterization of SSCD preferences, we may be tempted to think that the converse is also true. However, as the following example shows, that is *not* the case and so a revealed preference theory built around arbitrary observed feasible action sets and SSCD must involve a data set property different from ARC; we leave this interesting issue to further research.

EXAMPLE 2. Let $X_i = \{u_i, v_i, w_i\}$ with $u_i < v_i < w_i$, and let $A_i^1 = \{u_i, w_i\}$, $A_i^2 = \{u_i, v_i\}$, and $A_i^3 = \{v_i, w_i\}$. Note that A_i^1 is not an interval of X_i. Suppose that $\xi_i^1 < \xi_i^2 < \xi_i^3$, and that $a_i^1 = w_i$, $a_i^2 = u_i$, and $a_i^3 = v_i$. Then $(w_i, \xi_i^1) \succsim_i^R (u_i, \xi_i^1)$, $(u_i, \xi_i^2) \succsim_i^R (v_i, \xi_i^2)$, and $(v_i, \xi_i^3) \succsim_i^R (w_i, \xi_i^3)$. The indirect revealed preference relation $\succsim_i^{RT}$ is equal to the direct revealed preference relation $\succsim_i^R$ in this example and, clearly, this set of three observations obeys ARC. However, it cannot be rationalized by an SSCD preference. Suppose, instead that an SSCD preference $\succsim_i$ rationalizes the data. Then, it must hold that $(w_i, \xi_i^1) \succsim_i (u_i, \xi_i^1)$ and, by SSCD, $(w_i, \xi_i^2) \succ_i (u_i, \xi_i^2)$. In addition, we have

$(u_i, \xi_i^2) \succsim_i (v_i, \xi_i^2)$ and so $(w_i, \xi_i^2) \succ_i (v_i, \xi_i^2)$. Since $\succsim_i$ obeys SSCD, we obtain $(w_i, \xi_i^3) \succ_i (v_i, \xi_i^3)$, which contradicts the direct revealed preference $(v_i, \xi_i^3) \succsim_i (w_i, \xi_i^3)$.

3.3   Out-of-sample predictions

Suppose an observer collects a data set $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ that is monotone rationalizable, and then, maintaining that hypothesis, asks the following question: what do the observations in $\mathcal{O}_i$ say about the *set of possible choices* of agent i in some treatment $(\xi^0, A^0) \in \Xi_i \times \mathcal{A}$?[9] If $\mathcal{O}_i$ obeys ARC, then we know that the set of all SID preferences that rationalize $\mathcal{O}_i$, call it $\mathcal{P}_i^\star$, is nonempty. For each $\succsim_i \in \mathcal{P}_i^\star$, the set of best responses at (ξ_i^0, A_i^0) is $\mathrm{BR}_i(\xi_i^0, A_i^0, \succsim_i)$, and hence the set of *possible best responses* at (ξ_i^0, A_i^0) is given by

$$\mathrm{PR}_i(\xi^0, A^0) := \bigcup_{\succsim_i \in \mathcal{P}_i^\star} \mathrm{BR}_i\left(\xi_i^0, A_i^0, \succsim_i\right). \tag{13}$$

It follows from Theorem 1 that,

$$\mathrm{PR}_i(\xi_i^0, A_i^0) = \left\{\tilde{x}_i \in A_i^0 : \overline{\mathcal{O}_i} = \mathcal{O}_i \cup \{(\tilde{x}_i, \xi_i^0, A_i^0)\} \text{ obeys ARC}\right\}, \tag{14}$$

where $\overline{\mathcal{O}_i}$ is the data set $\mathcal{O}_i$ augmented by the (fictitious) observation $\{(\tilde{a}_i, \xi_i^0, A_i^0)\}$. The following proposition shows that $\mathrm{PR}_i(\xi_i^0; A_i^0)$ coincides with the undominated elements with respect to $\succsim_i^{RTST}$.

PROPOSITION 2. *Suppose that $\mathcal{O}_i$ obeys ARC. For any $\xi^0 \in \Xi_i$, it holds that*

$$\mathrm{PR}_i(\xi_i^0, A_i^0) = \{x_i \in A_i^0 : \nexists \hat{x}_i \in A_i^0 \text{ such that } (\hat{x}_i, \xi_i^0) \succ_i^{RTST} (x_i, \xi_i^0)\}. \tag{15}$$

Proof. It follows from (14) that (15) holds provided we can show the following: $\overline{\mathcal{O}_i} = \mathcal{O}_i \cup \{\tilde{x}_i, \xi_i^0, A_i^0)\}$ violates ARC if and only if there is $\hat{x}_i \in A_i^0$ such that $(\hat{x}_i, \xi_i^0) \succ_i^{RTST} (\tilde{x}_i, \xi_i^0)$. Let $\succsim_i^{\overline{R}}$, $\succsim_i^{\overline{RT}}$, $\succsim_i^{\overline{RTS}}$, and $\succsim_i^{\overline{RTST}}$ be the revealed preference relations derived from $\overline{\mathcal{O}_i} = \mathcal{O}_i \cup \{(\tilde{x}_i, \xi_i^0, A_i^0)\}$; by definition, these must contain the analogous revealed preference relations of $\mathcal{O}_i$, i.e., $\succsim_i^R$, $\succsim_i^{RT}$, $\succsim_i^{RTS}$, and $\succsim_i^{RTST}$. Suppose there is $\hat{x}_i \in A_i^0$ such that $(\hat{x}_i, \xi_i^0) \succ_i^{RTST} (\tilde{x}_i, \xi_i^0)$ and so $(\hat{x}_i, \xi_i^0) \succ_i^{\overline{RTST}} (\tilde{x}_i, \xi_i^0)$.

---

[9]The environment (ξ^0, A^0) may – or may not – be distinct from the ones already observed in the data set; the latter can still be an interesting question since optimal choices are not unique.
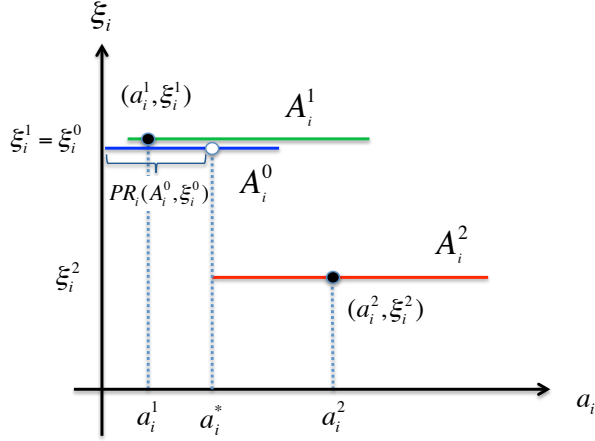
Figure 1: $\mathcal{E}(A^0)$ in Example 3

On the other hand, since $\hat{x}_i \in A_i^0$, we have $(\tilde{x}_{-i}, \xi_i^0) \succsim_i^{\overline{R}} (\hat{x}_i, \xi_i^0)$. This is a violation of the property (11) and, by Lemma 2, $\overline{\mathcal{O}}_i$ violates ARC. Conversely, suppose that $\overline{\mathcal{O}}_i = \mathcal{O}_i \cup \{(\tilde{x}_i, \xi_i^0, A_i^0)\}$ violates ARC. Since $\mathcal{O}_i$ obeys ARC, this violation can only occur in two ways: there is $\hat{x}_i \in X_i$ such that $(\tilde{x}_i, \xi_i^0) \succsim_i^{\overline{RT}} (\hat{x}_i, \xi_i^0)$ and $(\hat{x}_i, \bar{\xi}_i) \succsim_i^{RT} (\tilde{x}_i, \bar{\xi}_i)$ with either (1) $\hat{x}_i < \tilde{x}_i$ and $\bar{\xi}_i > \xi_i^0$ or (2) $\hat{x}_i > \tilde{x}_i$ and $\bar{\xi}_i < \xi_i^0$. We need to show that $\tilde{x}_i$ is dominated (with respect to $\succ_i^{RTST}$) by some element in A_i^0. In either cases (1) or (2), since $(\hat{x}_i, \bar{\xi}_i) \succsim_i^{RT} (\tilde{x}_i, \bar{\xi}_i)$, we obtain $(\hat{x}_i, \xi_i^0) \succ_i^{RTS} (\tilde{x}_i, \xi_i^0)$. If $\hat{x}_i \in A_i^0$, we are done. If $\hat{x}_i \notin A_i^0$ then, given that $(\tilde{x}_i, \xi_i^0) \succsim_i^{\overline{RT}} (\hat{x}_i, \xi_i^0)$, there exists $\bar{x}_i \in A_i^0$ such that $(\bar{x}_i, \xi_i^0) \succsim_i^{RT} (\hat{x}_i, \xi_i^0)$. Thus $(\bar{x}_i, \xi_i^0) \succ_i^{RTST} (\tilde{x}_i, \xi_i^0)$. $\qquad\square$

It is very convenient to have Proposition 2 because computing $\succsim_i^{RTST}$ is straightforward and thus it is also straightforward to obtain the set of possible responses at a given treatment.

EXAMPLE 3. Consider two observations as depicted in Figure 1, where A_i^1 and A_i^2 are the feasible sets of agent i at observations 1 and 2 respectively, while ξ_i^1 and ξ_i^2 are the parameter values at each observation. Let A_i^0 be the blue segment in the figure. It is easy to check that observations 1 and 2 obey ARC, and that the set of possible best responses, $\mathrm{PR}_i(a_i^0; A_i^0)$, is the set indicated in the figure. Notice that this set is not closed since $a_i^* \notin \mathrm{PR}_i(a_i^0, A_i^0)$. Indeed, $(a_i^2, \xi_i^1) \succ_1^{RTS} (a_i^*, \xi_i^1)$ since $(a_i^2, \xi_i^2) \succsim_i^{R} (a_i^*, \xi_i^2)$. Furthermore, $(a_i^1, \xi_i^1) \succsim_i^{R} (a_i^2, \xi_i^1)$ and so we obtain $(a_i^1, \xi_i^1) \succ_i^{RTST} (a_i^*, \xi_i^1)$.

19

4 Revealed strategic complementarity

Let $\mathbb{G} = \{\mathcal{G}(y, A)\}_{(y,A) \in Y \times \mathcal{A}}$ be a collection of games, as defined in the Section 2.2. We consider an observer who has a set of observations drawn from this collection. Each observation consists of a triple (a^t, y^t, A^t), where a^t is the action profile observed at the treatment $(y^t, A^t) \in Y \times \mathcal{A}$. The set of observations is finite and is denoted by $\mathcal{O} = \{(a^t, y^t, A^t)\}_{t \in \mathcal{T}}$, where $\mathcal{T} = \{1, 2, ..., T\}$.

DEFINITION 2. *A data set* $\mathcal{O} = \{(a^t, y^t, A^t)\}_{t \in \mathcal{T}}$ *is consistent with strategic complementarity (or SC-rationalizable) if there exists a profile of regular and SID preferences* $(\succsim_i)_{i \in N}$ *such that each observation constitutes a Nash equilibrium, i.e., for every* $t \in \mathcal{T}$, $(a_i^t, a_{-i}^t, y_i^t) \succsim_i (x_i, a_{-i}^t, y^t)$ *for all* $x_i \in A_i^t$.

The motivation for this definition is clear. If $\mathcal{O}$ is SC-rationalizable then we have found a profile of preference $(\succsim_i)_{i \in N}$ such that (i) a^t is a Nash equilibrium of $\mathcal{G}(A^t, y^t)$ and (ii) the family of games $\mathbb{G} = \{\mathcal{G}(y, A)\}_{(y,A) \in Y \times \mathcal{A}}$, where $\mathcal{G}(y, A) = [(y_i)_{i \in N}, (A_i)_{i \in N}, (\succsim_i)_{i \in N}]$ exhibits strategic complementarity (in the sense defined in Section 2.2).

For each agent i, we can define the *agent data set* $\mathcal{O}_i = \{(a_i^t, \xi_i^t, A_i^t)\}_{t=1}^T$ *induced by* $\mathcal{O}$, where $\xi_i^t = (a_{-i}^t, y_i^t)$. We say that $\mathcal{O} = \{(a^t, A^t, y^t)\}_{t \in \mathcal{T}}$ obeys ARC if $\mathcal{O}_i$ obeys ARC, for every agent i. It is clear that $\mathcal{O}$ is SC-rationalizable if and only if $\mathcal{O}_i$ is monotone-rationalizable for every agent i. This leads to the following result, which is an immediate consequence of Theorem 1 and provides with us with an easy-to-implement test of SC-rationalizability.

THEOREM 2. *The data set* $\mathcal{O} = \{a^t, y^t, A^t\}_{t \in \mathcal{T}}$ *is SC-rationalizable if and only if it obeys ARC.*

We turn now to the issue of out-of-sample equilibrium predictions. Given an SC-rationalizable data set $\mathcal{O} = \{(a^t, A^t, y^t)\}_{t \in \mathcal{T}}$, the agent data set $\mathcal{O}_i$ obeys ARC and so the set of regular and SID preferences that rationalize $\mathcal{O}_i$, i.e., $\mathcal{P}_i^\star$, is nonempty. The observed strategy profile a^t in $\mathcal{O}$ is supported as a Nash equilibrium by any preference profile $(\succsim_i)_{i \in N}$ in $\mathcal{P}^\star := \times_{i \in N} \mathcal{P}_i^\star$. For each $(\succsim_i)_{i \in N} \in \mathcal{P}^\star$, we know from Theorem B that the set of pure strategy Nash equilibria at another game $\mathcal{G}(y^0, A^0)$, which we shall denote by $E(y^0, A^0, (\succsim_i)_{i \in N})$, is nonempty and hence

$$\mathcal{E}(y^0, A^0) := \bigcup_{(\succsim_i)_{i \in N} \in \mathcal{P}^\star} E\left(y^0, A^0, (\succsim_i)_{i \in N}\right)$$

is also nonempty. $\mathcal{E}(y^0, A^0)$ is the set of *possible Nash equilibria of the game* $\mathcal{G}(y^0, A^0)$. This gives rise to two related questions that we shall answer in this section: [1] how can we compute $\mathcal{E}(y^0, A^0)$ from the data? and [2] what can we say about the structure of $\mathcal{E}(y^0, A^0)$?

4.1 Computable characterization of $\mathcal{E}(y^0, A^0)$

Recall that $\mathrm{PR}_i(\xi_i, A_i^0)$ denotes the possible best responses of player i in A_i^0 to $\xi_i = (a_{-i}, y_i^0)$ (see (13)). Given this, we define the *joint possible response correspondence* $\mathrm{PR}(\cdot, y^0, A^0) : A^0 \rightrightarrows A^0$ by

$$\mathrm{PR}(a, y^0, A^0) = \left(\mathrm{PR}_1(a_{-1}, y_1^0; A_1^0), \mathrm{PR}_2(a_{-2}, y_2^0, A_2^0), ..., \mathrm{PR}_n(a_{-n}, y_n^0, A_n^0) \right). \tag{16}$$

The crucial observation to make in computing $\mathcal{E}(y^0, A^0)$ is that just as the set of Nash equilibria in a game coincides with the fixed points of its joint best response correspondence, so the set of possible Nash equilibria, $\mathcal{E}(y^0, A^0)$, coincides with the fixed points of $\mathrm{PR}(\cdot, y^0, A^0)$. Equivalently, one could think of $\mathcal{E}(y^0, A^0)$ as the intersection of the graphs of each player's possible response correspondence, i.e., $\mathcal{E}(y^0, A^0) = \bigcap_{i \in N} \Gamma_i(y^0, A^0)$, where

$$\Gamma_i(y^0, A^0) = \{(a_i, a_{-i}) \in A^0 : a_i \in \mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)\}. \tag{17}$$

Therefore, the computation of $\mathcal{E}(A^0; y^0)$ hinges on the computation of $\mathrm{PR}_i(\cdot, y_i^0, A^0) : A_{-i} \rightrightarrows A_i^0$. Two features of this correspondence together make it possible for us to compute it explicitly.

First, we know from Proposition 2 that, for any a_{-i}, the set $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ coincides exactly with those elements in A_i^0 that are not dominated (with respect to $\succ_i^{RTST}$) by another element in A_i^0. Since the data set is finite, $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ can be constructed after a finite number of steps and, in fact, one could also show that it consists of a finite number of intervals.

Second, the correspondence $\mathrm{PR}_i(\cdot, y_i^0, A^0)$ takes only finitely many distinct values. For $j \neq i$, let

$$A_j^{\mathcal{T}} = \{a_j \in X_j : \exists\, a_{-j} \text{ such that } (a_j, a_{-j}) = a^t \text{ for some } t \in \mathcal{T}\}$$

We denote by $\mathcal{I}_j$ the collection consisting of *all* subsets of A_j^0 of the following two types: the singleton sets $\{\tilde{a}_j\}$, where $\tilde{a}_j$ is in the set $\underline{A}_j^0 = \left(A_j^{\mathcal{T}} \cap A_j^0 \right) \bigcup \max A_j^0 \bigcup \min A_j^0$ and the interval sets

$\{a \in A_j^0 : \tilde{a} < a < \tilde{b}\}$, where $\tilde{a} \in \underline{A}_j^0$ and $\tilde{b}$ is the element in $\underline{A}_j^0$ immediately above $\tilde{a}$. We denote by $\mathcal{H}_i$ the collection of hyper-rectangles

$$I_1 \times I_2 \times ... \times I_{i-1} \times I_{i+1} \times ... \times I_N$$

where $I_j \in \mathcal{I}_j$, for $j \neq i$; note that these hyper-rectangles are subsets of $\times_{j \neq i} A_j^0$. Then one could show that for any hyper-rectangle $H_i \in \mathcal{H}_i$, the following property holds:

$$a'_{-i}, \; a''_{-i} \in H_i \Longrightarrow \mathrm{PR}_i(a'_{-i}, y_i^0; A_i^0) = \mathrm{PR}_i(a''_{-i}, y_i^0; A_i^0). \tag{18}$$

In other words, the correspondence $\mathrm{PR}_i(\cdot, y_i^0; A_i^0)$ is constant within each hyper-rectangle H_i. Therefore, to compute this correspondence we need only find its value via (15) for a typical element within each hyper-rectangle H_i in the *finite* collection $\mathcal{H}_i$.

It follows from these two observations that the graph of player i's possible response correspondence (as defined by (17)) is also given by

$$\Gamma_i(y^0, A^0) = \{(a_i, a_{-i}) \in A^0 : \not\exists \, \hat{a}_i \in A_i^0 \text{ such that } (\hat{a}_i, a_{-i}, y_i^0) \succ_i^{RTST} (a_i, a_{-i}, y_i^0)\} \tag{19}$$

and can be explicitly constructed. Furthermore, because $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ consists of a finite union of intervals of A_i^0, $\Gamma_i(y^0, A^0)$ is a finite union of hyper-rectangles in A^0. The following theorem, which we prove in the Appendix, summarizes these observations.

THEOREM 3. *Suppose a data set $\mathcal{O} = \{(a^t, y^t, A^t)\}_{t=1}^T$ obeys ARC and let $(y^0, A^0) \in Y \times \mathcal{A}$.*

(i) *$\mathrm{PR}_i(\cdot, y_i^0, A_i^0)$ obeys (15) and (18) and, for any $a_{-i} \in \times_{j \neq i} A_j^0$, $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ consists of a finite union of intervals of A_i^0.*

(ii) *The graph of $\mathrm{PR}_i(\cdot, y_i^0, A_i^0)$, $\Gamma_i(y^0, A^0)$, is a finite union of hyper-rectangles in A^0. Consequently, the set of possible Nash equilibria, $\mathcal{E}(y^0, A^0) = \bigcap_{i \in N} \Gamma_i(y^0, A^0)$, is also a finite union of hyper-rectangles in A^0.*

EXAMPLE 4. Figure 2(a) depicts two observations, $\{(a^1, A^1) \text{ and } (a^2, A^2)\}$, drawn from two games involving two players. This data set obeys ARC and we would like to compute $\mathcal{E}(A^0)$, where
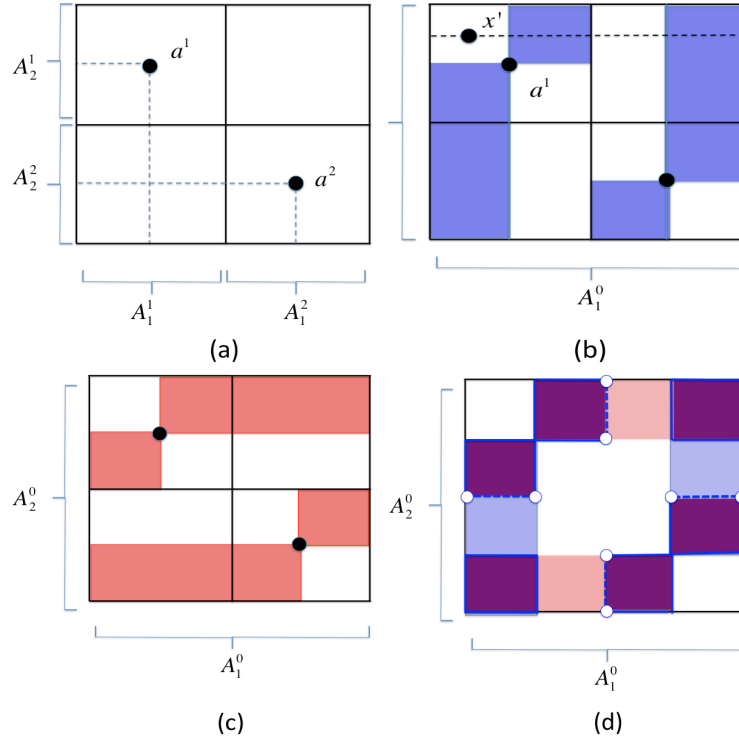
22

Figure 2: $\mathcal{E}(A^0)$ in Example 4

$A_i^0 = A_i^1 \cup A_i^2$ (for $i = 1, 2$). First, we claim that the unshaded area in Figure 2(b) cannot be contained in $\Gamma_1(A^0)$. Indeed, consider the point $x' = (x_1', x_2')$ in the unshaded area, at which $x_1' < a_1^1$, $x_2' > a_2^1$, and $x_1' \in A_1^1$. Therefore, $(a_1^1, a_2^1) \gtrsim_1^R (x_1', a_2^1)$ and so $(a_1^1, a_2^1) \gtrsim_1^{RT} (x_1', a_2^1)$. Since $x_2' > a_2^1$, $(a_1^1, a_2^1) \succ_1^{RTS} (x_1', a_2^1)$, which means that $(x_1', x_2') \notin \Gamma_1(A^0)$. Using (19), it is easy to check that $\Gamma_1(A^0)$ corresponds precisely to the shaded area in Figure 2(b). Similarly, $\Gamma_2(A^0)$ consists of the shaded area in Figure 2(c). The common shaded area, as depicted with the darker shade in Figure 2(d), represents $\mathcal{E}(A^0) = \Gamma_1(A^0) \cap \Gamma_2(A^0)$. Note that the dashed lines are *excluded* from $\mathcal{E}(A^0)$, so this set is not closed.

4.2 The structure of $\mathcal{E}(y^0, A^0)$

As we have pointed out in Section 2.2, the set of pure strategy Nash equilibria in a game with strategic complementarity admits a largest and smallest Nash equilibrium, both of which exhibit monotone comparative statics with respect to exogenous parameters. In this subsection, we show that these properties are largely inherited by the set of predicted pure strategy Nash equilibria

23

$\mathcal{E}(y^0, A^0)$. The next result (which we prove in the Appendix) lists the main structural properties of $\mathcal{E}(y^0, A^0)$; we have consciously presented them in a way that is analogous to Theorem B.

THEOREM 4. *Suppose a data set $\mathcal{O} = \{(a^t, y^t, A^t)\}_{t \in \mathcal{T}}$ obeys ARC and let $(y^0, A^0) \in Y \times \mathcal{A}$. Then $\mathcal{E}(y^0, A^0)$, the set of possible pure strategy Nash equilibria of the game $\mathcal{G}(y^0, A^0)$, is nonempty. Its closure admits a largest and a smallest element, both of which are increasing in $y^0 \in Y$.*

Since A^0 is a subcomplete sublattice of $(\mathbb{R}^n, \geqslant)$, *any* set in A^0 will have a supremum and an infimum in A^0. Therefore, the principal claim in Theorem 4 is that the supremum and infimum of the closure of $\mathcal{E}(y^0, A^0)$ are *contained in that set* (and thus arbitrarily close to elements of $\mathcal{E}(y^0, A^0)$): to all intents and purposes, we could speak of a largest and a smallest possible Nash equilibrium. Note that the analogous statement in Theorem B is stronger since it says that the set of pure strategy Nash equilibria (even when it is not closed) has a largest and a smallest element; however, Example 3 in Section 3.3 shows that the conclusion in Theorem 4 cannot be strengthened since in that case the possible response set does not contain its supremum.

In the special but important case where A^0 is finite, every subset of A^0 is closed and so it follows immediately from Theorem 4 that $\mathcal{E}(y^0, A^0)$ is a closed set with a largest and smallest element. The conclusion of Theorem 4 may also be strengthened in the case where the feasible action set of every agent is unchanged throughout the observations, i.e. $A^t = A^0 \in \mathcal{A}$ for all $t \in \mathcal{T}$. By (14), a necessary and sufficient condition for $\tilde{a}_i \in A^0$ to be contained in $\mathrm{PR}_i(a_{-i}, y_i^0; A_i^0)$ is that $\overline{\mathcal{O}}_i = \mathcal{O}_i \cup \{(\tilde{a}_i, (a_{-i}, y_i^0), A_i^0)\}$ obeys ARC. If $A^0 = A^t$ for all $t \in \mathcal{T}$, then it is straightforward to check that this is equivalent to $\tilde{a}_i$ having the following property:

$$\text{for all } t \in \mathcal{T}, \ \tilde{a}_i \geqslant a_i^t \text{ if } (a_{-i}, y_i^0) > \xi_i^t \text{ and } \tilde{a}_i \leqslant a_i^t \text{ if } (a_{-i}, y_i^0) < \xi_i^t. \tag{20}$$

It follows that $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ must be a closed interval in A_i^0 and (by Theorem 3) its graph $\Gamma_i(y^0, A^0)$ is a finite union of *closed* hyper-rectangles. Therefore, $\mathcal{E}(y^0; A^0) = \bigcap_{i \in N} \Gamma_i(y^0, A^0)$ is also closed and, by Theorem 4, it must contain its largest and smallest element.

5 Testing for complementarity with cross sectional data

So far in this paper we have assumed that the observer has access to panel data that gives the actions of the same agent (or, in the case of a game, the same group of agents) across different treatments. Oftentimes, data of this type is not available; instead, we only observe the actions of different agents, with presumably heterogeneous preferences, subject to different treatments. It is possible to extend our revealed preference analysis to this setting, provided we assume that the distribution of preferences is the same across populations subject to different treatments or, put another way, the assignment of agents or groups to treatments is random.

5.1 Stochastic monotone rationalizability

Suppose we observe a population of agents, whom we shall call population i, choosing actions from a subset of a chain X_i. Throughout this section (and unlike previous sections), we shall require that X_i be a *finite* chain. As usual, we assume agents choose from feasible sets that are intervals of X_i. Preferences are potentially affected by a set of parameters Ξ_i. At each observation t, all agents in population i are subject to the same treatment $(\xi_i^t, A_i^t) \in \Xi_i \times \mathcal{A}_i$, though they may choose different actions because they have different preferences. We assume that the true distribution of actions is observable and given by μ_i^t, where $\mu_i^t(x_i)$ denotes the fraction of agents who choose action x_i; we require $\mu_i^t(x_i) = 0$ for all $x_i \notin A_i^t$. The (cross sectional) data set for population i is a collection of triples $(\mu_i^t, \xi_i^t, A_i^t)$, i.e., $\mathcal{O}_i = \{(\mu_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$, where $\mathcal{T} = \{1, 2, ..., T\}$. Given $\mathcal{O}_i$, we denote the set of observed treatments by E_i, i.e., $E_i = \{(\xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$. We allow for the *same* treatment to occur at different observations; it is possible that $\mu^t \neq \mu^s$ even though the treatments at observations t and s are identical since we do not require agents to have unique optimal actions.[10] We adopt the convention of allowing the same treatment to be repeated in the set E_i if it occurs at more than one observation.

We call $\mathbf{a}_i = (a_i^1, a_i^2, ..., a_i^T) \in \times_{t \in \mathcal{T}} A_i^t$ a *monotone rationalizable path* on E_i if the induced 'panel' data set $\{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ is monotone-rationalizable (in the sense defined in Section 3) and denote

---

[10]If it helps, one could think of the index t itself to be part of the treatment, which may influence an agent's selection rule amongst optimal choices, though it has no impact on the agent's preference or the feasible alternatives, which depend only on the 'real' treatment.

the set of monotone rationalizable paths by $\mathbf{A}_i$. Since we allow for non-unique optimal choices, two distinct monotone rationalizable paths may be rationalized by the same SID preference.

DEFINITION 3. *A data set* $\mathcal{O}_i = \{(\mu^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ *is* stochastically monotone rationalizable *if there exists a probability distribution* Q_i *on* $\mathbf{A}_i$, *the set of monotone rationalizable paths on* E_i, *such that* $\mu_i^t(x_i) = \sum_{\mathbf{a}_i \in \mathbf{A}_i} Q_i(\mathbf{a}_i) 1(a_i^t = x_i)$ *for all* $t \in \mathcal{T}$ *and* $x_i \in X_i$.

When there is no danger of confusion, we shall simply refer to a data set as *monotone rationalizable* when it is stochastically monotone rationalizable. The definition says that the population i can be decomposed into types corresponding to different monotone rationalizable paths, so that the observed behavior of each type (across treatments) is consistent with maximizing an SID preference; it captures the idea that treatments have been randomly assigned across the entire population by requiring that the distribution of types is the same across treatments.[11]

Theorem 1 tells us that a path $\mathbf{a}_i$ on E_i is monotone rationalizable if and only if it is *ARC-consistent* in the sense that the data set $\{(a_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ obeys ARC. Therefore, we have the following result.

THEOREM 5. *A data set* $\mathcal{O}_i = \{(\mu^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$ *is monotone rationalizable if and only if there exists a probability distribution* Q_i *on* $\mathbf{A}_i^*$, *the set of ARC-consistent paths on* E_i, *such that*

$$\mu_i^t(x_i) = \sum_{\mathbf{a}_i \in \mathbf{A}_i^*} Q_i(\mathbf{a}_i) 1(a_i^t = x_i) \ \ \textit{for all } t \in \mathcal{T} \textit{ and } x_i \in X_i. \tag{21}$$

This theorem sets out a procedure that could, in principle, allow us to determine the monotone-rationalizability of a stochastic data set: first, we need to list all the ARC-consistent paths, and then we solve the linear equations given by (21). Of course, whether or not this procedure is implementable in practice will depend crucially on the number of observations, treatments, and possible actions, which determines the size of the set of ARC-consistent paths.

---

[11]While our definition of monotone rationalizable paths excludes the possibility that some group in the population may decide among non-unique optimal actions stochastically, the large population assumption means that this is without loss of generality. If, say, 10% of the population is indifferent between two optimal actions a' and a'' at some observation t, and decides between them by flipping a fair coin, then it simply means that 5% will belong to a type that chooses a' at t and another 5% to a type that choose a'' at t. A data set drawn from a large population of agents with heterogenous SID preferences who use stochastic selection rules (when there are multiple optimal actions) will still be stochastically monotone rationalizable in the sense defined here.

Given a monotone-rationalizable stochastic data set $\mathcal{O}_i = \{(\mu_i^t, \xi_i^t, A_i^t)\}_{t \in \mathcal{T}}$, we may wish to predict behavior at some given treatment, (ξ_i^0, A_i^0). The prediction consists of all those distributions on A_i^0 that are compatible with the data set. Formally, a distribution μ_i^0 on X_i is a *possible response distribution* at (ξ_i^0, A_i^0) if the augmented stochastic data set $\mathcal{O} \cup \{\mu_i^0, \xi_i^0, A_i^0\}$ is monotone rationalizable. It follows immediately from Theorem 5 that μ_i^0 is a possible response distribution at (ξ_i^0, A_i^0) if and only if there exists a probability distribution $\tilde{Q}_i$ on $\mathbf{A}_i^{**}$, the set of ARC-consistent paths on the set of environments $E_i \cup \{(\xi_i^0, A_i^0)\}$, such that for every $t \in \mathcal{T} \cup \{0\}$ and $x_i \in X_i$,

$$\mu_i^t(x_i) = \sum_{\mathbf{a}_i \in \mathbf{A}_i^{**}} \tilde{Q}_i(\mathbf{a}_i) 1(a_i^t = x_i) \ \text{ for all } t \in \mathcal{T} \cup \{0\} \text{ and } x_i \in X_i. \text{[12]} \tag{22}$$

It is worth noting that we allow for $(\xi_i^0, A_i^0) = (\xi_i^{t'}, A_i^{t'})$ for some observation t' and, indeed, it is instructive to consider that case. Then $\mu_i^{t'}$ is clearly a possible response distribution but since multiple optimal actions are permitted, the set of all such distributions can be strictly larger. In other words, in determining whether or not a distribution is a possible response distribution, we allow for the possibility that agents in the population with multiple optimal actions at the treatment $(\xi_i^{t'}, A_i^{t'})$ could switch to a different optimal action than the one taken at t'.

Let us denote the set of possible response distributions by $\text{PRD}_i(\xi_i^0, A_i^0)$. All the elements of $\text{PRD}_i(\xi_i^0, A_i^0)$ can be obtained by solving the equations (22). The unknown variables in this system are $\tilde{Q}_i(\mathbf{a}_i)$ for all $\mathbf{a}_i \in \mathbf{A}_i^{**}$ and $\mu_i^0(x_i)$ for all $x_i \in A_i^0$, and the equations are linear in these variables. Very conveniently, this implies that $\text{PRD}_i(\xi_i^0, A_i^0)$ *is a convex set*. We may be interested in establishing the possible fraction of agents who will choose a particular action $\tilde{x}_i$ in the treatment (ξ_i^0, A_i^0). Since $\text{PRD}_i(\xi_i^0, A_i^0)$ is a convex set, this is given precisely by the closed interval

$$\left[\min\{\mu_i^0(\tilde{x}_i) : \mu_i^0 \in \text{PRD}_i(\xi_i^0, A_i^0)\} \ , \ \max\{\mu_i^0(\tilde{x}_i) : \mu_i^0 \in \text{PRD}_i(\xi_i^0, A_i^0)\}\right]$$

The value of $\max\{\mu_i^0(\tilde{x}_i) : \mu_i^0 \in \text{PRD}_i(\xi_i^0, A_i^0)\}$ can be obtained by solving the following linear program:

$$\max \mu_i^0(\tilde{x}_i) \ \text{ subject to } \{\tilde{Q}_i(\mathbf{a}_i)\}_{\mathbf{a}_i \in \mathbf{A}_i^{**}} \text{ and } \{\mu_i^0(x_i)\}_{x_i \in A_i^0} \text{ satisfying (22)}.$$

---

[12] It follows from this definition that $\mu_i^t(x_i) = 0$ if $x_i \notin A_i^0$.

In a similar vein, we can calculate $\min\{\mu_i^0(\tilde{a}_i) : \mu_i^0 \in \mathrm{PRD}_i(\xi_i^0, A_i^0)\}$.

5.2 Stochastic strategic complementarity

The results on stochastic monotone rationalizability have an analog in a game-theoretic framework. In this case, we assume that the population consists of groups of n players, with each group choosing an action profile from their joint feasible set $A = \times_{i \in N} A_i$, where A_i is an interval of a finite chain X_i. The player in role i takes an action in A_i; the player's preference over his/her actions is affected by the actions of other players in that group and by some exogenous variable drawn from Y_i. We assume that the observer can distinguish amongst players in different roles in the game and can observe their actions separately; for example, in a population of heterosexual couples, the observer can distinguish between the 'husband' player and the 'wife' player and can observe their actions separately.

At observation t, each group in the population chooses an action profile from the joint feasible action set $A^t \in A$, with the exogenous parameter being $y^t \in Y = \times_{i \in N} Y_i$.; thus all groups in the population are subject to the same treatment $(y^t, A^t) \in Y \times \mathcal{A}$, with observed differences in action profiles stemming from heterogenous preferences amongst players within each group and possibly different equilibrium selection rules. We observe a probability distribution μ^t, with support on A^t, where $\mu^t(x)$ denotes the fraction of groups in which the action profile $x \in X$ is played. Therefore, the data set can be written as $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t \in \mathcal{T}}$. We denote the set of observed treatments by E, i.e., $E = \{(y^t, A^t)\}_{t \in \mathcal{T}}$. The possibility of multiple equilibria means that it is both meaningful and interesting to allow for the same treatment to appear at more than one observation. We have explained this at length in Section 5.1 and we shall not repeat it here. We allow identical treatments to appear more than once in E if they correspond to different observations. We refer to $\mathbf{a} = (a^1, a^2, ..., a^T) \in \times_{t \in \mathcal{T}} A^t$ as an *SC-rationalizable path on E* if the induced 'panel' data set $\{a^t, y^t, A^t\}_{t \in \mathcal{T}}$ is SC-rationalizable (in the sense defined in Section 4). The set of SC-rationalizable paths on E is denoted by $\mathbf{A}$.

DEFINITION 4. *A data set* $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t \in \mathcal{T}}$ *is* stochastically SC-rationalizable *if there is a probability distribution on* $\mathbf{A}$ *such that* $\mu^t(x) = \sum_{\mathbf{a} \in \mathbf{A}} Q(\mathbf{a}) 1(a^t = x)$ *for all* $t \in \mathcal{T}$ *and* $x \in X$.

Unless there is danger of confusion, we shall simply refer to a data set as *SC-rationalizable* when it is stochastically SC-rationalizable. This definition says that the population can be decomposed into 'group types' corresponding to different SC-rationalizable paths, so that we could interpret the action profile for each group as a Nash equilibrium, with players having SID preferences that are the same across observations; it captures the idea that treatments are randomly assigned across the large population of groups, so that the distribution of types is identical across treatments. As in the single agent case discussed in Section 5.1, the definition allows for groups belonging to different types to have members with the same preferences, because of the possibility of multiple equilibria. It is also worth emphasizing that the definition imposes no restrictions on what groups can be formed; for example, if a data set consists of a population of heterosexual couples, then the set of SC-rationalizable paths $\mathbf{A}$ allows for all possible matchings between different types of male and female players.

By Theorem 2, a path on E is SC-rationalizable if and only if it is *ARC-consistent* in the sense that $\{(a^t, y^t, A^t)\}_{t \in \mathcal{T}}$ obeys ARC. This leads immediately to the following result.

THEOREM 6. *A data set $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t \in \mathcal{T}}$ is SC-rationalizable if and only if there exists a probability distribution Q on $\mathbf{A}^*$, the set of ARC-consistent paths on E, such that*

$$\mu^t(x) = \sum_{\mathbf{a} \in \mathbf{A}^*} Q(\mathbf{a}) 1(a^t = x) \ \text{ for all } t \in \mathcal{T} \text{ and } x \in X. \tag{23}$$

Given an SC-rationalizable data set $\mathcal{O} = \{(\mu^t, y^t, A^t)\}_{t \in \mathcal{T}}$, the out-of-sample predictions at some given environment $(y^0, A^0) \in Y \times \mathcal{A}$ can be obtained by identifying those distributions μ^0 (which must have their support on A^0) such that the augmented stochastic data set $\mathcal{O} \cup \{\mu^0, y^0, A^0\}$ is SC-rationalizable. We refer to μ^0 as a *possible (Nash) equilibrium distribution* and denote the set of these distributions by $\text{PED}(y^0, A^0)$. It follows immediately from Theorem 6 that μ^0 is a possible equilibrium distribution if and only if there exists a probability distribution $\widetilde{Q}$ on $\mathbf{A}^{**}$, the set of ARC-consistent paths on the set of environments $E \cup \{(y^0, A^0)\}$, such that for every $t \in \mathcal{T} \cup \{0\}$ and $x \in X$,

$$\mu^t(x) = \sum_{\mathbf{a} \in \mathbf{A}^{**}} \widetilde{Q}(\mathbf{a}) 1(a^t = x) \text{ for all } t \in \mathcal{T} \cup \{0\} \text{ and } x \in X. \tag{24}$$

All the elements of PED(ξ^0, A^0) can be obtained by solving the equations (24). The unknown variables in this system are $\widetilde{Q}(\mathbf{a})$ for all $\mathbf{a} \in \mathbf{A}^{**}$ and $\mu^0(x)$ for all $x \in A^0$, and the equations are linear in these variables, which implies that PED(ξ^0, A^0) *is a convex set.* It follows that the possible fraction of the population playing a particular strategy profile $\tilde{x}$ at (ξ^0, A^0) will take values in an interval, with its limits obtained by solving the appropriate linear programs. (See the analogous result at the end of Section 5.1.)

If we wish, we can also form set estimates of the fraction of players in a particular role who choose a given action. Formally, a distribution μ^0 on X induces a distribution ν_i^0 on the equilibrium actions of player i; for each $\tilde{x}_i \in X_i$,

$$\nu_i^0(\tilde{x}_i) = \sum_{\{x \in A^0 : x_i = \tilde{x}_i\}} \mu^0(x). \tag{25}$$

The *set* of possible distributions on player i's equilibrium actions, which we shall denote by PED$_i(\xi^0, A^0)$ is also convex. This follows immediately from the convexity of PED(ξ^0, A^0). Since PED$_i(\xi^0, A^0)$ is a convex set, the predicted fraction of players in role i who choose a particular action $\tilde{x}_i$ from A_i^0 is given precisely by the closed interval

$$\left[\min\{\nu_i^0(\tilde{x}_i) : \nu_i^0 \in \text{PED}_i(\xi_i^0, A_i^0)\} , \ \max\{\nu_i^0(\tilde{x}_i) : \nu_i^0 \in \text{PED}_i(\xi_i^0, A_i^0)\} \right].$$

By (25), the value of $\max\{\nu_i^0(\tilde{x}_i) : \nu_i^0 \in \text{PED}_i(\xi_i^0, A_i^0)\}$ can be easily obtained by solving the following linear program:

$$\sum_{\{x \in A^0 : x_i = \tilde{x}_i\}} \mu_i^0(x) \text{ subject to } \{\widetilde{Q}(\mathbf{a})\}_{\mathbf{a} \in \mathbf{A}^{**}} \text{ and } \{\mu^0(x_i)\}_{x_i \in A_i^0} \text{ satisfying (24)}.$$

In a similar vein, we can calculate $\min\{\nu_i^0(\tilde{x}_i) : \nu_i^0 \in \text{PED}_i(\xi_i^0, A_i^0)\}$.

6 Application: Smoking Decisions in Married Couples

We now apply the results of Section 5.2 to the analysis of smoking decisions among married couples. Each married couple is modeled as a group whose members decide whether or not to smoke, with the smoking decision of each person depending on both the smoking decision of his/her partner

Smoking rates		Smoking rates	
Pr (Husband Smokes | Wife Smokes)	59.2%	Pr (Husband Smokes | Smoking Allowed)	31.2%
Pr (Husband Smokes | Wife Doesn't Smoke)	15.6%	Pr (Husband Smokes | Smoking Prohibited)	21.9%
Pr (Wife Smokes | Husband Smokes)	46.6%	Pr (Wife Smokes | Smoking Allowed)	24.9%
Pr (Wife Smokes | Husband Doesn't Smoke)	10.0%	Pr (Wife Smokes | Smoking Prohibited)	17.6%

Figure 3: Conditional Smoking Rates

and the smoking policy at his/her workplace. We use a data set that provides us with the smoking decision and the workplace smoking policy for each member of a large population of married couples. Differing workplace smoking policies provide the treatment variation needed for testing the presence of strategic complementarity. Similar data has also been used by Cutler and Glaeser (2010). As we do, they test for (and find) the presence of interaction in smoking behavior among married couples, using the exogenous variation in workplace smoking policies as an instrument. Their work differs from ours in that they use a reduced form parametric model of smoking behavior.

6.1 Data

We employ the Tobacco Use Supplement of the Current Population Survey (TUS-CPS) to get information on both smoking decisions and workplace smoking policies. This is an NCI-sponsored survey of tobacco use that has been administered as part of the US Census Bureau's Current Population Survey every 2 to 3 years since 1992. We focus on the period 1992-1993 because, in contrast to more recent years, a significant proportion of workplaces then did not have smoking restrictions, which guarantees that we have enough treatment variation. While the smoking information is obtained from everyone in our population of interest, the question on workplace smoking policy is posed only to indoor workers. Thus, we restrict attention to married couples where both members work indoors. After eliminating from our sample all couples where at least one member did not reply to all the questions of interest, we have 5,363 married couples across the US.

Within this sample, the smoking rate is 23.8% among the men and 18.7% among the women. Smoking is permitted in 19.7% of husbands' workplaces and 15% of wives' workplaces. Figure 3 displays the conditional probabilities of smoking given partner's smoking behavior (left panel) and smoking policy at work (right panel). As we can see, irrespective of gender, the probability of

Workplace smoking policy = (0,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 8.6%	μ (S,S) = 14.9%
Non-smoking (N)	μ (N,N) = 64.6%	μ (S,N) = 11.9%

Workplace smoking policy = (1,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 11.3%	μ (S,S) = 16.3%
Non-smoking (N)	μ (N,N) = 59.0%	μ (S,N) = 13.4%

Workplace smoking policy = (0,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 6.9%	μ (S,S) = 9.4%
Non-smoking (N)	μ (N,N) = 71.8%	μ (S,N) = 11.9%

Workplace smoking policy = (1,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 9.1%	μ (S,S) = 14.9%
Non-smoking (N)	μ (N,N) = 59.1%	μ (S,N) = 16.9%

Figure 4: Joint Distribution of Smoking Choices Across Smoking Policies

smoking is larger when either the partner smokes or when smoking is permitted in the workplace. Overall, the fraction of spouses that make the same smoking choice —either both smoke or do not smoke— is around 80% of the whole sample. These figures are at least suggestive of the influence of spousal behavior and workplace policy on smoking decisions. To examine this issue more closely, we now apply the test developed in Section 5.2.

6.2   Findings

Figure 4 displays the distribution of joint choices regarding smoking decisions for four different workplace smoking policies, which serve as treatments in our analysis. As before, we use μ to indicate the probability of each action profile for each workplace smoking policy. The first argument of μ takes the value of S if the husband smokes and N otherwise; the second argument indicates the smoking decision of his wife. Similarly, the first argument in Workplace Smoking Policy takes the value of 1 if smoking is permitted in the working place of the husband and 0 otherwise; the second argument indicates the smoking policy at the wife's workplace. In this application, the choice set of each person is $\{N, S\}$ and it remains the same across observations.

We use Theorem 6 to test if this data set is SC-rationalizable. (Appendix II gives a fuller description of the procedure.) Notice that in this application there are a priori $4^4 = 256$ group paths, since for each of the four possible treatment values, there are four joint choices that a married couple can make. One could check that 64 of these paths are SC-rationalizable. For our test to be valid, we must assume that the population is randomly assigned to these four treatments,

Workplace smoking policy = (0,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 8.8%	μ (S,S) = 14.8%
Non-smoking (N)	μ (N,N) = 64.5%	μ (S,N) = 11.8%

Workplace smoking policy = (1,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 11.3%	μ (S,S) = 16.3%
Non-smoking (N)	μ (N,N) = 59.0%	μ (S,N) = 13.4%

Workplace smoking policy = (0,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 6.9%	μ (S,S) = 9.4%
Non-smoking (N)	μ (N,N) = 71.8%	μ (S,N) = 11.9%

Workplace smoking policy = (1,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 8.8%	μ (S,S) = 15.0%
Non-smoking (N)	μ (N,N) = 59.2%	μ (S,N) = 16.9%

Figure 5: Closest SC-rationalizable distribution of smoking choices

so that the distribution of the 64 types is the same across treatments. For now at least, let us also ignore issues of sample size and treat the observations in Figure 4 as the true distribution of joint actions across the four treatments. In that case, we can test for SC-rationalizability by checking if there is a positive solution to the linear system (23), where the solution vector, if it exists, gives the proportion of the population belonging to each of the 64 types. Performing this test, we find that there is in fact no solution to the linear system, so the data set is not SC-rationalizable.

This may come as a surprise, since the number of unknowns (64) far exceeds the number of linear constraints and it is tempting to think that the conditions are very permissive. In fact, there is at least one easy-to-understand reason why the data set displayed in Figure 4 is not SC-rationalizable. Notice from Figure 4 that $\mu\left(N,S|1,0\right) = 9.1\% > 8.6\% = \mu\left(N,S|0,1\right)$. This is impossible because, to be consistent with strategic complementarity, any couple type that selects (N,S) under the smoking policy $(1,0)$ must select (N,S) again under the smoking policy $(0,1)$.[13] Interestingly, if we solve for the data set that is SC-rationalizable and closest (as measured by the sum of square deviations) to the one actually observed, the solution, as displayed in Figure 5, sets $\mu\left(N,S|1,0\right) = \mu\left(N,S|0,1\right) = 8.8\%$.

If we compare the entries in Figures 4 and 5, we see immediately that they are quite close, which naturally makes us wonder whether the observed violation of SC-rationalizability is in fact significant. To address this issue, we adopt the approach recently proposed by Kitamura and Stoye

---

[13]Let $\succsim_h$ be the husband's preference and $\succsim_w$ the wife's preference. Then $(N,S|1,0) \succsim_h (S,S|1,0)$ implies that $(N,S|0,1) \succ_h (S,S|0,1)$, so $(S,S|0,1)$ is ruled out as an equilibrium. Furthermore, $(N,S|1,0) \succsim_h (S,S|1,0)$ implies $(N,N|0,1) \succ_h (S,N|0,1)$, so $(S,N|0,1)$ is impossible as well. Turning now to the wife, since $(N,S|1,0) \succsim_w (N,N|1,0)$, we obtain $(N,S|0,1) \succ_w (N,N|0,1)$, so $(N,N|0,1)$ cannot be an equilibrium.

(2013); they develop a method of evaluating the statistical significance of a data set violating a set of linear constraints that directly applies to our framework.[14] Roughly speaking, the test assumes that the closest compatible distribution displayed in Figure 5 is the true population distribution, and uses a bootstrap procedure to calculate the likelihood of getting a sample like the one we observe. By applying their test, we find that the probability of getting our sample (or a more extreme one), assuming that our modelling restrictions are true, is 0.3795. The latter corresponds to the p-value for the null hypothesis that our modelling assumptions are true. This means that we cannot reject SC-rationalizability at a significance level of 5% or 10%. (See Appendix II for a fuller description of the Kitamura-Stoye procedure and our implementation.)

To examine the issue more closely, we also divided the entire sample into three smaller sub-samples, according to the educational attainment of the couples: (i) both spouses have high education levels (measured as having at least some college education); (ii) both have low education levels; and (iii) one spouse has high education level and the other a low education level. We find that the choice probabilities of the group where both spouses have high education levels are directly consistent with our modelling restrictions, i.e., the observed choice probabilities are SC-rationalizable. The other two groups are not directly consistent with the model, with the p-values being 0.509 and 0.127 respectively for the couples with low education and couples with mixed education levels respectively.

Appendix I

We have shown in Lemma 1 that $\succsim_i^{RT}$ has the interval property. The following extension of that result is needed for the proofs of Lemmas 2 and 3.

LEMMA A1: *The binary relations $\succsim_i^{RTS}$, $\succ_i^{RTS}$, and $\succsim_i^{RTST}$ on $X_i \times \Xi_i$ have the interval property.*

Proof. Let $x_i'' > x_i > x_i'$. (The case where $x_i'' < x_i < x_i'$ can be proved in a similar way.) If $(x_i'', \xi_i) \succsim_i^{RTS} (\succ_i^{RTS}) (x_i', \xi_i)$ holds, there exists some $\xi_i' \leqslant (<) \xi_i$ such that $(x_i'', \xi_i') \succsim_i^{RT} (x_i', \xi_i')$. By the interval property of $\succsim_i^{RT}$, we obtain $(x_i'', \xi_i') \succsim_i^{RT} (x_i, \xi_i')$. Since $x_i'' > x_i$ and $\xi_i' \leqslant (<) \xi_i$, we have that $(x_i'', \xi_i) \succsim_i^{RTS} (\succ_i^{RTS}) (x_i, \xi_i)$. So we have shown that $\succsim_i^{RTS}$ and $\succ_i^{RTS}$ have the interval

---

[14]Kitamura and Stoye (2013) apply their test to the consumer utility-maximization problem.

property.

If $(x_i'', \xi_i) \succsim_i^{RTST} (x_i', \xi_i)$, there exists a sequence $z_i^1, z_i^2, ..., z_i^k$ such that

$$(x_i'', \xi_i) \succsim_i^{RTS} (z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^k, \xi_i) \succsim_i^{RTS} (x_i', \xi_i).$$

Letting $z_i^0 = x_i''$ and $z_i^{k+1} = x_i'$, since $x_i'' > x_i > x_i'$, we can find some $0 \leqslant m \leqslant k$ such that $z_i^m \geqslant x_i \geqslant z_i^{m+1}$. By the interval property of $\succsim_i^{RTS}$, we obtain $(z_i^m, \xi_i) \succsim_i^{RTS} (x_i, \xi_i)$. Thus $(x_i'', \xi_i) \succsim_i^{RTST} (x_i, \xi_i)$ since $(x_i'', \xi_i) \succsim_i^{RTST} (z_i^m, \xi_i) \succsim_i^{RTS} (x_i, \xi_i)$. $\qquad\square$

Proof of Lemma 2: We first prove that (11) holds. (11) is equivalent to $\succsim_i^{RTS}$ being *cyclically consistent*, i.e.,

$$(z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^k, \xi_i) \implies (z_i^k, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i). \tag{26}$$

Cyclical consistency can in turn be equivalently re-formulated as the following:

$$(z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^k, \xi_i) \succsim_i^{RTS} (z_i^1, \xi_i) \tag{27}$$

$$\implies (z_i^1, \xi_i) \nsucc_i^{RTS} (z_i^2, \xi_i) \nsucc_i^{RTS} ... \nsucc_i^{RTS} (z_i^k, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i)$$

Thus, whenever there is a cycle like (27), it *must* be the case that

$$(z_i^1, \xi_i) \succsim_i^{RT} (z_i^2, \xi_i) \succsim_i^{RT} ... \succsim_i^{RT} (z_i^k, \xi_i) \succsim_i^{RT} (z_i^1, \xi_i)$$

We prove (11) by induction on the length of the chain, k, on the left side of (26). Whenever (26) holds for chains of length k or less (equivalently, whenever the cycles in (27) have length k or less), we say that $\succsim_i^{RTS}$ is k-consistent. For 2-consistency, we need to show that

$$(z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \implies (z_i^2, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i).$$

Suppose that $z_i^1 > z_i^2$; the case of $z_i^1 < z_i^2$ can be dealt with in a similar way. By definition, if $(z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i)$ then there is $\xi_i' \leqslant \xi_i$ such that $(z_i^1, \xi_i') \succsim_i^{RT} (z_i^2, \xi_i')$. On the other hand,

if $(z_i^2, \xi_i) >_i^{RTS} (z_i^1, \xi_i)$, then there is $\xi_i'' > \xi_i$ such that $(z_i^2, \xi_i'') \succsim_i^{RT} (z_i^1, \xi_i'')$ and so we obtain a violation of ARC.

Suppose that $\succsim_i^{RTS}$ is k-consistent for all $k < \bar{k}$. To show that $\bar{k}$-consistency holds, suppose the left side of (26) holds for $k = \bar{k}$ and $z_i^1 < z_i^{\bar{k}}$. Clearly, there must be $m < \bar{k}$ such that $z_i^m < z_i^{\bar{k}}$ and $z_i^{m+1} \succsim_i z_i^{\bar{k}}$. We consider two cases separately: (A) $z_i^m \geqslant z_i^1$ and (B) $z_i^m < z_i^1$. In case (A), by the interval property of $\succsim_i^{RTS}$, we obtain $(z_i^m, \xi_i) \succsim_i^{RTS} (z_i^{\bar{k}}, \xi_i)$. By way of contradiction, suppose also that $(z_i^{\bar{k}}, \xi_i) >_i^{RTS} (z_i^1, \xi_i)$. Then the interval property of $>_i^{RTS}$ guarantees that $(z_i^{\bar{k}}, \xi_i) >_i^{RTS} (z_i^m, \xi_i)$ and so we obtain a violation of 2-consistency. For case (B), since $(z_i^m, \xi_i) \succsim_i^{RTS} (z_i^{m+1}, \xi_i)$, the interval property guarantees that $(z_i^m, \xi_i) \succsim_i^{RTS} (z_i^1, \xi_i)$. So we obtain the cycle

$$(z_i^1, \xi_i) \succsim_i^{RTS} (z_i^2, \xi_i) \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^m, \xi_i) \succsim_i^{RTS} (z_i^1, \xi_i) \tag{28}$$

which has length strictly lower than $\bar{k}$. By the induction hypothesis, we obtain

$$(z_i^1, \xi_i) \nsucc_i^{RTS} (z_i^2, \xi_i) \nsucc_i^{RTS} ... \nsucc_i^{RTS} (z_i^m, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i)$$

and so we can replace each $\succsim_i^{RTS}$ in (28) by $\succsim_i^{RT}$. Furthermore, $(z_i^m, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i)$ guarantees that $(z_i^m, \xi_i) \nsucc_i^{RTS} (z_i^{m+1}, \xi_i)$, by the interval property of $>_i^{RTS}$. Therefore, $(z_i^1, \xi_i) \succsim_i^{RT} (x_i^{m+1}, \xi_i)$ and, by the interval property of $\succsim_i^{RT}$, we obtain $(z_i^1, \xi_i) \succsim_i^{RT} (x_i^{\bar{k}}, \xi_i)$. 2-consistency then ensures that $(z_i^{\bar{k}}, \xi_i) \nsucc_i^{RTS} (z_i^1, \xi_i)$. This completes the proof that (11) holds.

By definition, $\succsim_i^{RTST}$ obeys SSCD if whenever $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$ *or* $x_i'' < x_i'$ and $\xi_i'' < \xi_i'$, then

$$(x_i'', \xi_i') \succsim_i^{RTST} (x_i', \xi_i') \implies (x_i'', \xi_i'') >_i^{RTST} (x_i', \xi_i'').$$

We shall concentrate on the case where $x_i'' > x_i'$; the other case has a similar proof. If $(x_i'', \xi_i) \succsim_i^{RTS} (x_i', \xi_i)$, then we know that there is z_i^j (for $j = 1, 2, ..., k$) such that

$$(x_i'', \xi_i') \succsim_i^{RTS} (z_i^1, \xi_i') \succsim_i^{RTS} (z_i^2, \xi_i') \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^k, \xi_i') \succsim_i^{RTS} (x_i', \xi_i'). \tag{29}$$

We can also choose a chain with the property that (writing $z_i^0 = x_i''$ and $z_i^{k+1} = x_i'$) $(z_i^m, \xi_i') \nsucc_i^{RTS} (z_i^{m'}, \xi_i')$ for $m' > m + 1$; in other words, no link in the chain can be dropped. We claim that, for

such a chain, we must have

$$x_i'' > z_i^1 > z_i^2 > ... > z_i^k > x_i'. \tag{30}$$

Once this is established, the rest is straightforward: since $\succsim_i^{RTS}$ obeys SSCD, (29) and (30) imply

$$(x_i'', \xi_i'') \succ_i^{RTS} (z_i^1, \xi_i'') \succ_i^{RTS} (z_i^2, \xi_i'') \succ_i^{RTS} ... \succ_i^{RTS} (z_i^k, \xi_i'') \succ_i^{RTS} (x_i', \xi_i'')$$

and so $(x_i'', \xi_i'') \succ_i^{RTST} (x_i', \xi_i'')$.

It remains for us to establish (30). If this is false then there is m such that $z_i^{m+1} > z_i^m$. Let z_i^{m+n} be the first time after z_i^{m+1} such that $z_i^{m+n} \leqslant z_i^m$. Then we have $z_i^{m+n} \leqslant z_i^m < z_i^{m+n-1}$. Since $(z_i^{m+n-1}, \xi_i') \succsim_i^{RTS} (z_i^{m+n}, \xi_i')$, the interval property of $\succsim_i^{RTS}$ guarantees that $(z_i^{m+n-1}, \xi_i') \succsim_i^{RTS} (z_i^m, \xi_i')$. Thus we obtain a cycle

$$(z_i^m, \xi_i') \succsim_i^{RTS} (z_i^{m+1}, \xi_i') \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^{m+n-1}, \xi_i') \succsim_i^{RTS} (z_i^m, \xi_i').$$

Since $\succsim_i^{RTS}$ is cyclically consistent, this chain cannot be related by $\succ_i^{RTS}$ and must be related by $\succsim_i^{RT}$. In particular, $(z_i^{m+n-1}, \xi_i') \not\succ_i^{RTS} (z_i^m, \xi_i')$ and thus $(z_i^{m+n-1}, \xi_i') \not\succ_i^{RTS} (z_i^{m+n}, \xi_i')$ (by the interval property of $\succ_i^{RTS}$). We conclude that $(z_i^m, \xi_i') \succsim_i^{RT} (z_i^{m+n}, \xi_i')$ and thus we can shorten (29) to

$$(x_i'', \xi_i') \succsim_i^{RTS} (z_i^1, \xi_i') \succsim_i^{RTS} ... \succsim_i^{RTS} (z_i^m, \xi_i') \succsim_i^{RTS} (z_i^{m+n}, \xi_i') \succsim_i^{RTS} ... (z_i^k, \xi_i') \succsim_i^{RTS} (x_i', \xi_i')$$

which contradicts our assumption that no link in the chain can be dropped. $\square$

Proof of Lemma 3: We first show that $\succsim_i^*$ is a preference that rationalizes $\mathcal{O}_i$. Clearly, $\succsim_i^*$ is complete and reflexive, so to demonstrate that it is a preference we need only show that it is transitive. Indeed, suppose

$$(a_i, \xi_i) \succsim_i^* (b_i, \xi_i) \succsim_i^* (c_i, \xi_i) \succsim_i^* (a_i, \xi_i).$$

There are only four fundamentally distinct cases we need to consider:

Case 1. None of the three elements are related by $\succsim_i^{RTST}$. Given the definition of $\succsim_i^*$, this means that $a_i < b_i < c_i < a_i$, which is impossible.

Case 2. $a_i < b_i < c_i$, $(a_i, \xi_i) \parallel_i^{RTST} (b_i, \xi_i)$, $(b_i, \xi_i) \parallel_i^{RTST} (c_i, \xi_i)$, and $(c_i, \xi_i) \succsim_i^{RTST} (a_i, \xi_i)$. This is again impossible since the interval property of $\succsim_i^{RTST}$ will imply that $(c_i, \xi_i) \succsim_i^{RTST} (b_i, \xi_i)$.

Case 3. $a_i < b_i$, $(a_i, \xi_i) \parallel_i^{RTST} (b_i, \xi_i)$, $(b_i, \xi_i) \succsim_i^{RTST} (c_i, \xi_i) \succsim_i^{RTST} (a_i, \xi_i)$. This is also impossible because, by the transitivity of $\succsim^{RTST}$, we obtain $(b_i, \xi_i) \succsim_i^{RTST} (a_i, \xi_i)$.

Case 4. $(a_i, \xi_i) \succsim_i^{RTST} (b_i, \xi_i) \succsim_i^{RTST} (c_i, \xi_i) \succsim_i^{RTST} (a_i, \xi_i)$. By (11), this is only possible if

$$(a_i, \xi_i) \succsim_i^{RT} (b_i, \xi_i) \succsim_i^{RT} (c_i, \xi_i) \succsim_i^{RT} (a_i, \xi_i),$$

but then we also obtain, by the transitivity of $\succsim_i^{RT}$, $(a_i, \xi_i) \succsim_i^{RT} (c_i, \xi_i)$ and, hence, $(a_i, \xi_i) \succsim_i^* (c_i, \xi_i)$. Lastly, since $\succsim_i^{RTST} \subset \succsim^*$ by construction, it is clear that $\succsim_i^*$ rationalizes $\mathcal{O}_i$.

To show that $\succsim_i^*$ obeys SSCD, let $x_i'' > x_i'$ and $\xi_i'' > \xi_i'$; then

$$
\begin{aligned}
(x_i'', \xi_i') \succsim_i^* (x_i', \xi_i') &\Longrightarrow (x_i'', \xi_i') \succsim_i^{RTST} (x_i', \xi_i') \\
&\Longrightarrow (x_i'', \xi_i'') \succ_i^{RTST} (x_i', \xi_i'') \\
&\Longrightarrow (x_i'', \xi_i'') \succ_i^* (x_i', \xi_i''),
\end{aligned}
$$

in which the first implication follows from the definition of $\succsim_i^*$, the second implication from the SSCD property of $\succsim_i^{RTST}$, and the third from the fact that $\succ_i^*$ contains $\succ_i^{RTST}$ (so $\succsim_i^*$ extends $\succsim_i^{RTST}$ in the sense of (10)). The last claim is true because if $(x_i'', \xi_i) \succ_i^{RTST} (x_i', \xi_i)$, then Lemma 2 says that $(x_i', \xi_i) \not\succsim_i^{RTST} (x_i'', \xi_i)$; thus $(x_i', \xi_i) \not\succsim_i^* (x_i'', \xi_i)$ and we obtain $(x_i'', \xi_i) \succ_i^* (x_i', \xi_i)$.

It remains for us to show that, for every $\xi_i \in \Xi_i$, $\mathrm{BR}(\xi_i, K, \succsim^*)$ is nonempty and finite, where $K \subset X_i$ and K is compact in $\mathbb{R}$. If $K \not\ni a_i^t$ for every $t \in \mathcal{T}$, then, it follows from the definition of $\succsim_i^*$ that $(m, \xi_i) \succsim_i^* (z_i, \xi_i)$, where $m = \min K$ and $z_i \in K$. In this case, m is the only maximiser of $\succsim_i^*$ in K. Suppose that $K \ni a_i^t$ for some t. Since there are a finite number of observations, we can find some $a_i^s \in K$ such that $(a_i^s, \xi_i) \succsim_i^* (a_i^t, \xi_i)$ for every $a_i^t \in K$. We claim that either m or a_i^s maximises $\succsim_i^*$ in K for ξ_i, so that $\mathrm{BR}(\xi_i, K, \succsim^*)$ is indeed nonempty and finite. There are two cases to consider.

Suppose $(m, \xi_i) \succsim_i^* (a_i^s, \xi_i)$ and there is $z_i \in K$ such that $(z_i, \xi_i) \succ_i^* (m, \xi_i)$. Then, since $m < z_i$,

it must hold that $(z_i, \xi_i) \succ_i^{RTST} (m, \xi_i)$ and there is $\underline{t} \in \mathcal{T}$ such that $z_i = a_i^{\underline{t}}$. Consequently,

$$(a_i^t, \xi_i) \succ_i^* (m, \xi_i) \succsim_i^* (a_i^s, \xi_i) \succsim_i^* (a_i^{\underline{t}}, \xi_i),$$

which is a contradiction. Therefore, $(m, \xi_i) \succsim_i^* (x_i, \xi_i)$ for all $x_i \in K$. Now suppose $(a_i^s, \xi_i) \succsim_i^* (m, \xi_i)$. For every $x_i \in [m, n]$, either $(a_i^s, \xi_i) \succsim_i^{RTST} (x_i, \xi_i)$, in which case $(a_i^s, \xi_i) \succsim_i^* (x_i, \xi_i)$, or $(a_i^s, \xi_i) \parallel_i^{RTST} (x_i, \xi_i)$, in which case we have $(a_i^s, \xi_i) \succsim_i^* (m, \xi_i) \succsim_i^* (x_i, \xi_i)$. Thus $(a_i^s, \xi_i) \succsim_i^* (x_i, \xi_i)$ for all $x_i \in K$. $\qquad\square$

Proof of Theorem 3. Part (ii) follows straightforwardly from part (i), so we shall focus on proving (i), which consists of three claims. Proposition 2 says that (15) holds. To see that (18) holds, first note that $\tilde{a}_i \notin \mathrm{PR}_i(a'_{-i}, y_i^0, A_i^0)$ if and only if $\overline{\mathcal{O}}_i' = \mathcal{O}_i \cup \{(\tilde{a}_i, a'_{-i}, y_i^0, A_i^0)\}$ violates ARC. Since H_i is not a singleton, it must be an interval and so there is *no* a_i' such that $(a_i', a_{-i}'^t$ for some $t \in \mathcal{T}$. Therefore, $\overline{\mathcal{O}}_i'$ violates ARC if and only if there is $\hat{a}_i \in A_i^0$ and $\bar{a}_{-i}$ such that $(\hat{a}_i, \bar{a}_{-i}, \bar{y}_i) \succsim_i^{RT} (\tilde{a}_i, \bar{a}_{-i}, \bar{y}_i)$ with either (1) $\hat{a}_i < \tilde{a}_i$ and $(\bar{a}_{-i}, \bar{y}_i) > (a'_{-i}, y_i^0)$ or (2) $\hat{a}_i > \tilde{a}_i$ and $(\bar{a}_{-i}, \bar{y}_i) < (a'_{-i}, y_i^0)$. Note that there is $t \in \mathcal{T}$ such that $(\hat{a}_i, \bar{a}_{-i}) = a^t$; in particular, this means that $\bar{a}_{-i} \in \times_{j \neq i} \underline{A}^{\mathcal{T}}$. It follows from our definition of H_i that $(\bar{a}_{-i}, \bar{y}_i) > (a''_{-i}, y_i^0)$ if $(\bar{a}_{-i}, \bar{y}_i) > (a'_{-i}, y_i^0)$ and $(\bar{a}_{-i}, \bar{y}_i) < (a''_{-i}, y_i^0)$ if $(\bar{a}_{-i}, \bar{y}_i) < (a'_{-i}, y_i^0)$. Thus $\overline{\mathcal{O}}_i'' = \mathcal{O}_i \cup \{(\tilde{a}_i, a''_{-i}, y_i^0, A_i^0)\}$ also violates ARC. We conclude that $\tilde{a}_i \notin \mathrm{PR}_i(a''_{-i}, y_i^0, A_i^0)$ if $\tilde{a}_i \notin \mathrm{PR}_i(a'_{-i}, y_i^0, A_i^0)$, which establishes (18).

Lastly, we show that $\mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ consists of a finite union of intervals of A_i^0. This is equivalent to showing that $A_i^0 \backslash \mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ is a finite union of intervals; an element $\tilde{a}_i$ is in this set if and only if there is $t \in \mathcal{T}$ such that $a_i^t \in A_i^0$ and $(a_i^t, \xi_i^0) \succ_i^{RTST} (\tilde{a}_i, \xi_i^0)$, where $\xi_i^0 = (a_{-i}, y^0)$. This is turns holds if and only if is $s \in \mathcal{T}$ such that either (1) $(a_i^t, \xi_i^0) \succsim_i^{RTST} (a_i^s, \xi_i^0)$ and $(a_i^s, \xi_i^0) \succ_i^{RTS} (\tilde{a}_i, \xi_i^0)$ or (2) $(a_i^t, \xi_i^0) \succ_i^{RTST} (a_i^s, \xi_i^0)$ and $(a_i^s, \xi_i^0) \succsim_i^{RT} (\tilde{a}_i, \xi_i^0)$. Notice for a fixed $s \in \mathcal{T}$, the sets $\{a_i \in A_i^0 : (a_i^s, \xi_i^0) \succ_i^{RTS} (a_i, \xi_i^0)\}$ and $\{a_i \in A_i^0 : (a_i^s, \xi_i^0) \succsim_i^{RT} (a_i, \xi_i^0)\}$ both consist of intervals, because of the interval property on $\succ_i^{RTS}$ and $\succ_i^{RT}$ respectively. It follows that $A_i^0 \backslash \mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$ is a finite union of intervals. $\qquad\square$

The proof of Theorem 4 uses the following lemma.

LEMMA A2: *Suppose $\mathcal{O} = \{a^t, y^t, A^t\}_{t=1}^T$ obeys ARC and let $A^0 \in \mathcal{A}$. Then the map $p_i^{**} : A_{-i}^0 \times Y \to A_i^0$ given by*

$$p_i^{**}(a_{-i}, y_i) = \sup \mathrm{PR}_i(a_{-i}, y_i, A^0)$$

has the following properties: *(i) it is increasing in $(a_{-i}, y_i) \in A^0_{-i} \times Y_i$; (ii) for a'_{-i} and a''_{-i} in H_i, $p_i^{**}(a''_{-i}, y_i) = p_i^{**}(a'_{-i}, y_i)$; and (iii) if, for some $(\bar{a}_{-i}, \bar{y}_i)$, $p_i^{**}(\bar{a}_{-i}, \bar{y}_i) \in \mathrm{PR}_i(\bar{a}_{-i}, \bar{y}_i, A^0)$ and for some $(\hat{a}_{-i}, \hat{y}_i) > (\bar{a}_{-i}, \bar{y}_i)$, $p_i^{**}(\bar{a}_{-i}, \bar{y}_i) = p_i^{**}(\hat{a}_{-i}, \hat{y}_i)$, then $p_i^{**}(\hat{a}_{-i}, \hat{y}_i) \in \mathrm{PR}_i(\hat{a}_{-i}, \hat{y}_i, A^0)$.*

Remark: In a similar way, we define $p_i^* : A^0_{-i} \times Y_i \to A^0_i$ by $p_i^*(a_{-i}, y_i) = \inf \mathrm{PR}_i(a_{-i}, y_i, A^0)$. This function will obey properties (i) and (ii) and, instead of property (iii), it will have the following property (iii)': *if, for some $(\bar{a}_{-i}, \bar{y}_i)$, $p_i^*(\bar{a}_{-i}, \bar{y}_i) \in \mathrm{PR}_i(\bar{a}_{-i}, \bar{y}_i, A^0)$ and for some $(\hat{a}_{-i}, \hat{y}_i) < (\bar{a}_{-i}, \bar{y}_i)$, $p_i^*(\bar{a}_{-i}, \bar{y}_i) = p_i^*(\hat{a}_{-i}, \hat{y}_i)$, then $p_i^*(\hat{a}_{-i}, \hat{y}_i) \in \mathrm{PR}_i(\hat{a}_{-i}, \hat{y}_i, A^0)$.*

Proof. Since $\mathrm{PR}_i(a_{-i}, y_i, A^0)$ is the union of a collection of best response correspondences (see (13)), each of which is increasing in (a_{-i}, y_i), p_i^{**} must be increasing. Claim (ii) is an immediate consequence of (18) (which was proved in Theorem 3). Lastly, if $p_i^{**}(\bar{a}_{-i}, \bar{y}_i) \in \mathrm{PR}_i(\bar{a}_{-i}, \bar{y}_i, A^0)$ then there is $\succsim_i \in \mathcal{P}_i^*$ such that $p_i^{**}(\bar{a}_{-i}, \bar{y}_i) \in \mathrm{BR}_i(\bar{a}_{-i}, \bar{y}_i, A^0_i, \succsim_i)$. Since the best response correspondence is increasing, there is $a_i' \in \mathrm{BR}_i(\hat{a}_{-i}, \hat{y}_i, A^0_i, \succsim_i)$, and thus in $\mathrm{PR}_i(\hat{a}_{-i}, \hat{y}_i, A^0_i)$, such that $a_i' \geqslant p_i^{**}(\bar{a}_{-i}, \bar{y}_i)$. This establishes (iii). $\qquad\square$

Proof of Theorem 4: We have already explained at the beginning of Section 4 why $\mathcal{E}(y^0, A^0)$ is nonempty. We shall confine our attention to showing that $\max \overline{\mathcal{E}(y^0, A^0)}$ exists, where $\overline{\mathcal{E}(y^0, A^0)}$ refers to the closure of $\mathcal{E}(y^0, A^0)$; the proof for the other case is similar.[15] Firstly, note that the properties of p_i^{**} listed in Lemma A2 guarantee that there exists a sequence of functions $\{p_i^k(\cdot, y_i^0, A_i^0)\}_{k \in \mathbb{N}}$ selected from $\mathrm{PR}_i(\cdot, y_i^0, A_i^0)$ with the following properties: (i) for a'_{-i} and a''_{-i} in H_i, $p_i^k(a''_{-i}, y_i^0) = p_i^k(a'_{-i}, y_i^0)$; (ii) $p_i^k(a_{-i}, y_i^0, A_i^0)$ is increasing in a_{-i} and in k; (iii) $p_i^k(a_{-i}, y_i^0, A_i^0) = p_i^{**}(a_{-i}, y_i^0, A_i^0)$ if $p_i^{**}(a_{-i}, y_i^0, A_i^0) \in \mathrm{PR}_i(a_{-i}, y_i^0, A_i^0)$; and (iv) $\lim_{k \to \infty} p_i^k(a_{-i}, y_i^0, A_i^0) = p_i^{**}(a_{-i}, y_i^0, A_i^0)$. In other words, there is a sequence of increasing selections from $\mathrm{PR}_i(\cdot, y^0, A^0)$ that has $p_i^{**}(a_{-i}, y^0, A^0)$ as it limit, with the sequence being exactly equal to $p_i^{**}(a_{-i}, y_i^0, A_i^0)$ if the latter is a possible response of player i.

The function $p^k(a, y^0, A^0) = (p_i^k(a_{-i}, y_i^0, A_i^0))_{i \in N}$ is increasing in a, since p_i^k is increasing in a_{-i}. By Tarski's fixed point theorem, p^k has a largest fixed point, which we denote by $x^k(y^0, A^0)$. Since

---

[15]It is worth pointing out an obvious first approach that will *not* work. Given p_i^{**}, we can define, for each $a \in A^0$, $p^{**}(a, y^0) = (p_i^{**}(a_{-i}, y_i^0))_{i \in N}$, and since p_i^{**} is increasing in a_{-i}, so $p^{**}(a, y^0)$ is increasing in a. By Tarski's fixed point theorem, $p^{**}(\cdot, y^0)$ will have a fixed point and indeed a largest fixed point $a^\star$; thus the existence of $\max \overline{\mathcal{E}(y^0, A^0)}$ is ensured if it could be identified with $a^\star$. However, they are not generally the same points: it is straightforward to construct an increasing (but not compact-valued) correspondence such that its largest fixed point does not coincide with the largest fixed point of its supremum function. Our proof takes a different route.

$p_i^k(\cdot, y_i^0, A_i^0)$ is a selection from $\mathrm{PR}_i(\cdot, y_i^0, A_i^0)$, $x^k(y^0, A^0) \in \mathcal{E}(y^0, A^0)$. By the monotone fixed points theorem (see Section 2), the sequence $x^k(y^0, A^0)$ is increasing with k. Since A^0 is compact, this sequence must have a limit. This limit, which we denote by $a^{**}(y^0, A^0)$, lies in $\overline{\mathcal{E}(y^0, A^0)}$.

We claim that $a^{**}(y^0, A^0)) \geqslant \tilde{x}$, for any $\tilde{x} \in \mathcal{E}(y^0, A^0)$. Indeed, since $\tilde{x}_i \in \mathrm{PR}_i(\tilde{x}_{-i}, y_i^0, A_i^0)$ for all $i \in N$, for k sufficiently large, $p_i^k(\tilde{x}_{-i}, y_i^0, A_i^0) \geqslant \tilde{x}_i$. Now consider the map p^k confined to the domain $S = \times_{i \in N}\{a_i \in A_i^0 : a_i \geqslant \tilde{x}_i\}$. Since p^k is increasing, the image of p^k also falls on S; in other words, p^k can be considered as a map from S to itself. It is also an increasing map and, by Tarski's fixed point theorem will have a largest fixed point. The largest fixed point of p^k restricted to S must again be $x^k(y^0, A^0)$ and it follows from our construction that $x^k(y^0, A^0)) \geqslant \tilde{x}$. In turn this implies that $a^{**}(y^0, A^0)) \geqslant \tilde{x}$. So $a^{**}(y^0, A^0))$ is an upper bound of $\mathcal{E}(y^0, A^0)$ and thus also an upper bound of $\overline{\mathcal{E}(y^0, A^0)}$. Given that $a^{**}(y^0, A^0)) \in \overline{\mathcal{E}(y^0, A^0)}$, we conclude that $a^{**}(y^0, A^0)) = \max \overline{\mathcal{E}(y^0, A^0)}$.

To see that $a^{**}(y, A^0)$ is increasing with respect to the parameter, consider $y'' > y'$. Given the properties of p_i^{**} listed in Lemma A2, we can choose functions $\{p_i^k(\cdot, y_i, A_i^0)\}_{k \in \mathbb{N}}$ selected from $\mathrm{PR}_i(\cdot, y_i, A_i^0)$ (for $y_i = y_i'$ and y_i'') satisfying properties (i) – (iv) and, in addition, $p_i^k(a_{-i}, y_i'', A_i^0) \geqslant p_i^k(a_{-i}, y_i', A_i^0)$ for all a_{-i}. The map $p^k(\cdot, y_i'', A_i^0)$ is increasing and will have a largest fixed point $x^k(y''^0)$ which, by the monotone fixed points theorem satisfies $x^k(y''^0) \geqslant x^k(y'^0)$. Taking limits as $k \to \infty$, we obtain $a^{**}(y''^0) \geqslant a^{**}(y'^0)$. $\qquad\square$

Appendix II

AII.1 Data and Testing Procedures

We use the Tobacco Use Supplement to the Current Population Survey (TUS-CPS) to get information on both smoking decisions and workplace smoking policies. This is an NCI-sponsored survey of tobacco use that has been administered as part of the US Census Bureau's Current Population Survey every 2 to 3 years from 1992-1993. We focus on years 1992-1993 because, unlike more recent periods, there were still significant numbers of workplaces that permitted smoking. This guarantees we have enough treatment variation. While smoking information is asked of everyone in our population of interest, the smoking ban question is asked only of indoor workers. Thus, we restrict

attention to married couples where both members work indoor. After eliminating from our sample all couples where at least one member does not reply to the questions of interest, we have 5,363 married couples across the US. For 2,643 couples, both members have high education levels (the precise sense in which we explain later), for 1,422 couples, both members have low education levels, and for the remaining 1,298 couples, one spouse has a high education level and the other low.

We merge the relevant information of years 1992 and 1993. We first recover the information for September 1992 and add the information for January 1993 and July 1993 regarding spouses that do not appear in the previous period/s of time. We tabulate responses according to the variables from the Data Dictionary of the Current Population Survey for years 1992-1993 that we detail next.

Married Couples We consider as married couples all pairs where one of them is the reference person and the other one responds either *3. Husband* or *4. Wife* to question A-RRP Item 18B (Relationship to reference person).

Smoking Decisions We assign the value 0 (does not smoke) to all persons that respond either *2. No* to question A-S32 (Has... smoked at least 100 cigarettes in his/her entire life?) or *3. Not at all* to question A-S34 (Does... now smoke cigarettes every day, some days, or not at all?). We assign value 1 (smokes) to all persons that respond either *1. Every day* or *2. Some days* to question A-S34 (Does... now smoke cigarettes every day, some days, or not at all?).

Smoking Restrictions at Workplace We assign the value 0 (smoking restrictions at workplace) to all persons that respond *1. Yes* to question A-S68 (Does your place of work have an official policy that restricts smoking in anyway?). We assign value of 1 (no smoking restrictions at workplace) to all persons that respond *2. No* to question A-S68 (Does your place of work have an official policy that restricts smoking in anyway?).

Education We consider as high educated couples (HE) all those married couples (defined above) where both members report that they have high education levels; specifically these are couples where both members respond *40. Some college but not degree* or above to question A-HGA Item 18H (Education attainment). We consider as low educated couples (LE) all those married couples where both members respond strictly below *40. Some college but not degree* to question A-HGA Item 18H (Education attainment). We consider as mixed educated couples (Mix) all married couples

Workplace smoking policy = (0,1)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 6.7%	μ (S,S) = 7.7%
Non−smoking (N)	μ (N,N) = 77.4%	μ (S,N) = 8.2%

Workplace smoking policy = (1,1)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 6.7%	μ (S,S) = 9.0%
Non−smoking (N)	μ (N,N) = 70.8%	μ (S,N) = 13.5%

Workplace smoking policy = (0,0)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 4.9%	μ (S,S) = 4.6%
Non−smoking (N)	μ (N,N) = 81.8%	μ (S,N) = 8.7%

Workplace smoking policy = (1,0)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 4.6%	μ (S,S) = 10.2%
Non−smoking (N)	μ (N,N) = 69.6%	μ (S,N) = 15.5%

Figure 6: Joint Distribution of Smoking Choices Across Smoking Policies (HE)

Workplace smoking policy = (0,1)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 6.7%	μ (S,S) = 17.4%
Non−smoking (N)	μ (N,N) =63.1%	μ (S,N) = 12.8%

Workplace smoking policy = (1,1)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 12.5%	μ (S,S) = 17.5%
Non−smoking (N)	μ (N,N) = 53.8%	μ (S,N) = 16.3%

Workplace smoking policy = (0,0)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 9.6%	μ (S,S) = 11.0%
Non−smoking (N)	μ (N,N) = 65.9%	μ (S,N) = 13.5%

Workplace smoking policy = (1,0)		
Husband/Wife	Non−smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 10.0%	μ (S,S) = 10.5%
Non−smoking (N)	μ (N,N) = 61.0%	μ (S,N) = 18.5%

Figure 7: Joint Distribution of Smoking Choices Across Smoking Policies (Mix)

where one member has a high education level and the other a low education level.

We have already displayed the survey results for all couples in Figure 4 in the main body of the paper. Figures 6, 7, and 8 show the corresponding results when the survey responses are disaggregated into HE, LE, and Mix couples.

AII.2 Test and Closest Compatible Distribution

Testing whether a data set is consistent with strategic complementarity involves checking whether a system of linear equations

$$Ax = B. \tag{31}$$

has a positive solution x. We describe next all the components of this system.

Workplace smoking policy = (0,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 12.3%	μ (S,S) = 20.7%
Non-smoking (N)	μ (N,N) =52.0%	μ (S,N) = 15.1%

Workplace smoking policy = (1,1)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 14.0%	μ (S,S) = 21.1%
Non-smoking (N)	μ (N,N) = 53.5%	μ (S,N) = 11.4%

Workplace smoking policy = (0,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 9.3%	μ (S,S) = 19.5%
Non-smoking (N)	μ (N,N) = 53.0%	μ (S,N) = 18.2%

Workplace smoking policy = (1,0)		
Husband/Wife	Non-smoking (N)	Smoking (S)
Smoking (S)	μ (N,S) = 12.8%	μ (S,S) = 22.6%
Non-smoking (N)	μ (N,N) = 47.6%	μ (S,N) = 17.0%

Figure 8: Joint Distribution of Smoking Choices Across Smoking Policies (LE)

Matrix A This matrix is composed of 0's and 1's and describes the behavior (in terms of choices) of all SC-rationalizable group types. Recall that a group type specifies the profile of choices that the group makes for each possible vector of parameter values. In this case, A is a 16×64-matrix. Each row in matrix A corresponds to one of the 16 possible combinations of (joint) smoking choices and treatment values. An ijth element of matrix A takes value 1 if the jth group selects the smoking decision under the treatment corresponding to that row.

Vector B The size of this column vector is 16. It is composed of 4 conditional distributions. Each conditional distribution specifies the fraction of groups that, for a given treatment, make each of the four possible joint decisions.

Vector x This size of this column vector is 64. It represents a possible probability distribution over the set of SC-rationalizable group types.

We implement our test by using Matlab. Specifically, we use the program

$$x = \mathrm{linprog}(lb, [], [], A, B, lb, [])$$

to check whether system (31) has a positive solution in x. In this specification, inputs A and B are described as above and lb corresponds to a column vector of 64 zeros.

For those data vectors B that do not pass this test, we use program lsqnonneg in Matlab to find the positive vector x, with its components adding up to 1, that minimize $(B - Ax)^{'} (B - Ax)$.

AII.3 Small Sample Inference Procedure

As Kitamura and Stoye (2013) explain, the null hypothesis is equivalent to

$$\textbf{(H)} \quad \min_{x \in \mathbb{R}_+^{|K|}} (B - Ax)' (B - Ax) = 0$$

where $|K|$ is the number of SC-rationalizable group types (64 in our case). A natural sample counterpart of the objective function in **H** is given by

$$\left(\widehat{B} - Ax\right)' \left(\widehat{B} - Ax\right)$$

where $\widehat{B}$ estimates B by sample choice frequencies. Normalizing the latter by sample size N, we get

$$J_N = N \min_{x \in \mathbb{R}_+^{|K|}} \left(\widehat{B} - Ax\right)' \left(\widehat{B} - Ax\right).$$

Let x^* be any solution to this problem. If $Ax^* = \widehat{B}$, so that the observed choices are compatible with our restrictions, then $J_N = 0$ and the null hypothesis cannot be rejected.

Kitamura and Stoye (2013) proposes the following bootstrap algorithm to test **H**:

(i) Obtain a vector x^* that solves

$$J_N = N \min_{[x - \tau_N 1_{64}] \in \mathbb{R}_+^{64}} \left(\widehat{B} - Ax\right)' \left(\widehat{B} - Ax\right)$$

and compute $\widehat{C}_{\tau_N} = Ax^*$. In our application, we let $\tau_N = \left(\sqrt{\ln \underline{N}/\underline{N}}\right)/64$, where $\underline{N}$ is the minimum out the number of couples in each of the four treatments. (As Kitamura and Stoye (2013) explain, the tuning parameter τ_N plays the role of a similar tuning parameter in the moment selection approach.)

(ii) Calculate the boostrap estimators under the restriction

$$\widehat{B}_{\tau_N}^{(r)} = \widehat{B}^{(r)} - \widehat{B} + \widehat{C}_{\tau_N} \quad r = 1, ..., R$$

where $\widehat{C}_{\tau_N}$ derives from step (i) and $\widehat{B}^{(r)}$ is a re-sampled choice probability vector obtained via standard nonparametric boostrap. In addition, R is the number of boostrap replications. In the

paper, we let $R = 2000$.

(iii) Calculate the boostrap test statistic by solving the following problem

$$J_N^{(r)} (\tau_N) = N \min_{[x - \tau_N 1_{64}] \in \mathbb{R}_+^{64}} \left(\widehat{B}_{\tau_N}^{(r)} - Ax \right)' \left(\widehat{B}_{\tau_N}^{(r)} - Ax \right)$$

for $r = 1, ..., R$.

(iv) Use the empirical distribution of $J_N^{(r)} (\tau_N)$, $r = 1, ..., R$, to obtain the critical value of J_N.

We implement this procedure four times, namely, on all married couples, HE couples, LE couples, and Mix couples. In the case of HE couples, the vector $\widehat{B}$ is directly consistent, so the p-value of the test for this group is 1 and we cannot reject the null hypothesis. The p-values for the data set corresponding to all married couples is 0.3795, the one for LE is 0.509, and the one for Mix is 0.127.

References

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

Carvajal, A. (2004): Testable restrictions of Nash equilibrium with continuous domains. *Discussion paper series, Royal Holloway College, University of London*, 2004-26.

Cutler, D., and E. Glaeser, (2010): Social interactions and smoking. In Research Findings in the Economics of Aging, edited by D. Wise, 123-141. Chicago: University of Chicago Press.

Echenique, F., and I. Komunjer, (2009): Testing models with multiple equilibria by quantile methods. *Econometrica*, 77, 1281–1297.

Kitamura, Y., and J. Stoye, (2013): Nonparametric analysis of random utility models: testing. Cowels Foundation Discussion Paper NO. 1902.

Kline, B., and E. Tamer, (2012): Bounds for best response functions in binary games. *Journal of Econometric*, 166, 92-105.

Lazzati, N. (2015): Treatment response with social interactions: identification via monotone comparative statics. *Quantitative Economics*, 6, 49-83.

Manski, C. (2007): Partial identification of counterfactual choice probabilities. *International Economic Review*, 48, 1393-1410.

McFadden, D.L., and K. Richter (1991): Stochastic Rationality and Revealed Stochastic Preference, in *Preferences, Uncertainty and Rationality*, ed. by J. Chipman, D. McFadden, and M.K. Richter. Boulder: Westview Press, pp. 161186.

McFadden, D.L. (2005): Revealed Stochastic Preference: A Synthesis, *Economic Theory*, 26, 245264.

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

Milgrom, P. and J. Roberts, (1994): Comparing equilibria. *American Economic Review*, 84, 441-459.

Milgrom, P. and C. Shannon, (1994): Monotone comparative statics. *Econometrica*, 62, 157-180.

Molinari, F., and A. Rosen, (2008): The identification power of equilibrium in games: the supermodular case. *Journal of Business & Economic Statistics*, 26, 297-302.

Quah, J. K.-H. and B. Strulovici, (2009): Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77, 1949-1999.

Topkis, D. (1998): *Supermodularity and complementarity*. Princeton University Press.

Uetake, K., and Y. Watanabe, (2013): A note on identification and estimation of supermodular games. *Advances in Econometrics*, 31, 233-247.

Varian, H., (1982): The nonparametric approach to demand analysis. *Econometrica*, 50, 945-73.

Vives, X. (1990): Nash equilibria with strategic complementarities. *Journal of Mathematical Economics*, 19, 305-321.

Zhou, L. (1994): The set of Nash equilibria of a supermodular game is a complete lattice. *Games and Economic Behavior*, 7, 295-300.