

# Causal Discourse in a Game of Incomplete Information

Halbert White  
Department of Economics  
University of California, San Diego

Haiqing Xu  
Department of Economics  
University of Texas, Austin

Karim Chalak<sup>‡</sup>  
Department of Economics  
Boston College

September 14, 2012

## Abstract

Notions of cause and effect are fundamental to economic explanation. Despite the immediate intuitive content of price effects, income effects, and the like, rigorous foundations justifying well-posed discussions of cause and effect in the wide range of settings relevant to economics are still lacking. We illustrate the need for these foundations using the familiar context of an  $N$ -bidder private-value auction, posing a variety of relevant causal questions that cannot be formally addressed within existing causal frameworks. We extend the causal frameworks of Pearl (2000) and White and Chalak (2009) to introduce *topological settable systems*, a causal framework capable of delivering the missing answers. In particular, our framework can accommodate choices that are elements of general function spaces. Our analysis suggests how topological settable systems can be applied to support causal discourse in more general games and in other areas of economic inquiry.

**Acknowledgements** The authors are grateful to Douglas Bernheim, Kalyan Chatterjee, James Heckman, Vijay Krishna, Brad Larsen, Robert Marshall, Philip Reny, Joel Sobel, Maxwell Stinchcombe, and Quang Vuong for helpful comments and suggestions.

## 1 Introduction and Motivation

Causal discourse – that is, discussion of cause and effect – is fundamental to economic explanation. It appears naturally and unselfconsciously throughout Adam Smith’s *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776) and in all the major economic contributions of the nineteenth century (e.g., Mill, 1848; Marshall, 1890) and a good part of the twentieth (e.g., Hicks, 1939; Samuelson, 1947). As twentieth-century economists began to think carefully about systems of structural or simultaneous equations, work began on formalizing notions of

---

\*Corresponding author: Karim Chalak, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA. Email: chalak@bc.edu.

<sup>‡</sup>Our coauthor, Halbert White, passed away after completion of a recent draft of this paper. This manuscript is a slight modification of that draft.

causality and structure. Classical efforts in this area include the work of Haavelmo (1943, 1944), Marschak (1950), Simon (1953), Strotz and Wold (1960), and Granger (1969). Unfortunately, no clear consensus emerged from this work. Causal notions remained murky, in part due to the causal paradoxes associated with simultaneity, which nevertheless plays an indispensable role in describing economic phenomena. This lack of clarity contributed to economists tending to avoid not only formal discussion of causality, but even informal discussion, as Hoover (2004) documents. Nevertheless, causal discourse is so central to economics and the social sciences and so intuitive that explicit causal discussion has re-emerged in the last few decades, together with renewed and deep consideration of causal foundations (e.g., Holland, 1986; Heckman, 2005).

This renewal has led to significant advances, particularly in program and policy evaluation.<sup>1</sup> Nevertheless, causal discourse still occupies an ambivalent status in a variety of areas of economics. Although it is intuitive and natural to speak about income effects and price effects and the like, rigorous foundations justifying well-posed discussions of cause and effect in the wide range of settings relevant to economics, including game theory, are still lacking. Intuition can only go so far. Without firm foundations, it is easy to go astray when describing economic theories, when analyzing the identification or estimation of causal effects (particularly in the increasingly sophisticated structures analyzed nowadays), and especially when attempting to draw policy conclusions or economic insight from model estimates. There remains a clear need to find broadly applicable rigorous foundations for causal discourse in economics.

We demonstrate this need for a suitable causal framework using the familiar context of an  $N$ -bidder private-value auction, a game of incomplete information (Harsányi, 1967). This game is simple enough to allow straightforward analysis, yet rich enough for us to illustrate all of the central issues by posing a variety of relevant causal questions. As we discuss, these basic questions cannot be addressed using existing causal frameworks. We then provide a framework that delivers the missing answers by introducing *topological settable systems*, an extension of the causal frameworks of Pearl (2000) and White and Chalak (2009, “WC”). Applying this framework to the  $N$ -bidder private-value auction permits delivering answers to causal questions there, but also suggests how topological settable systems can be applied to support causal discourse in more general games and in other areas of economic inquiry.

In game theory, the lack of formal foundations for causal discourse leads, not surprisingly, either to informal discussion of cause and effect, which may be limited or misleading, or, more commonly, to the avoidance of such discussion altogether. As an example of informal causal discourse in game theoretic settings, consider the Wikipedia entry on *complete information*,<sup>2</sup>

---

<sup>1</sup>See e.g., Heckman (2005), Heckman and Vytlačil (2005), Imbens and Wooldridge (2009), and Heckman (2010).

<sup>2</sup>[http://en.wikipedia.org/wiki/Complete\\_information](http://en.wikipedia.org/wiki/Complete_information)

where we find the following:

If a game is not of complete information, then the individual players would not be able to predict the effect that their actions would have on the others players (even if the actor presumed other players would act rationally).

Is this statement correct? Does it even make sense? With only intuition as a guide, it is hard to know. Indeed, this statement raises an array of salient questions about causality for incomplete information games: In what sense is a player’s strategy or action causally affected by other players’ strategies or actions? How is the *ceteris paribus* “effect” of bidder  $j$ ’s strategy on bidder  $i$ ’s strategy defined? How do rationality in behavior and belief matter for causal discourse? What is the causal role of Harsányi’s (1967) agent types ( $X$ ), if any? Do the number of players ( $N$ ) and the distribution of types ( $F$ ) have effects? If so, how? If not, why not? What are the structural equations here? Specifically, are the simultaneous equations of, e.g., Bayesian-Nash equilibrium structural? How about the equilibrium “reduced form”? Is it structural? In particular, do equilibrium strategies and actions have structural meaning and/or causal content?

Giving sensible answers to these questions requires a suitable causal framework. Outside of economics, the Pearl causal model (PCM; Pearl, 2000, def. 7.1.1, p. 203) has emerged as a leading paradigm for understanding cause and effect. The PCM has been applied usefully to address certain causal inquiries (see e.g. Pearl, 2000; Halpern and Pearl, 2005(a,b)). In particular, the PCM has been productively applied to game theory, and, specifically, to games of incomplete information. Unfortunately, the PCM does not apply to answer the questions above. In seminal work applying the PCM to games, Koller and Milch (2001, 2003) build on “probabilistic graphical models” (e.g. Pearl, 2000) to introduce Multi-Agent Influence Diagrams (MAIDs) for representing and computing equilibrium in non-cooperative games. However, there is no mention of causality in Koller and Milch’s (2001, 2003) careful work. Probabilistic graphical models and related PCM notions have also been explicitly applied to incomplete information games by Penalva-Zuasti and Ryall (2003), Jiang and Leyton-Brown (2010), and Wellman, Hong, and Page (2011), among others. Nevertheless, there are generic limitations of the PCM for causal discourse in games (see WC): among other things, the PCM cannot support causal discourse in games with non-unique equilibrium (see also Halpern, 2000); by ruling out a causal role for “background” variables, the PCM does not permit discussion of the causal role played by structurally exogenous variables, such as agent types,  $X$ ; and the PCM is not explicit about attributes, i.e. non-varying objects that play a role in characterizing systems. As WC illustrate, the PCM also does not apply to complete information pure- and mixed-strategy static games or to infinitely repeated dynamic games. Thus, the PCM does not provide a satisfactory foundation for game-theoretic causal

discourse.

In order to overcome these limitations, WC introduced settable systems, extending and refining the PCM to accommodate features essential to economic analysis: optimization, possibly non-unique equilibrium, and learning, while preserving the structural systems spirit of the PCM. Game theory examples where settable systems apply but not the PCM<sup>3</sup> are complete information pure- and mixed-strategy static games, infinitely repeated dynamic games with complete and perfect information, and fictitious play with continuum strategies. Other examples are static consumer demand optimization, dynamic rational expectations consumer demand optimization, stochastic dynamic optimization of consumption and saving, and adaptive dynamic rational expectations models of perfectly competitive markets, among others.

Despite these many applications, settable systems still do not apply to certain major classes of problems, such as incomplete information games. In these games, players' choices of strategy ("type-contingent plans") can be rather general functions, such as monotone functions; but WC's settable systems only admit function variables belonging to topological spaces homeomorphic<sup>4</sup> to the space of countable sequences of reals, such as Hilbert space (e.g., Anderson and Bing, 1968). WC's settable systems framework cannot handle player choices that are elements of more general function spaces. While it may be natural to speak of "cause" and "effect" in these environments, this discourse remains informal, at best, without an adequate rigorous framework. This paper fills this void by introducing topological settable systems, which permit choices to be elements of general function spaces, providing just the right framework for answering the causal questions posed above. We illustrate topological settable systems by applying them to the familiar  $N$ -bidder private-value auction. Further, this application suggests how this framework may be applied not only to more elaborate games and to other areas of economics, but even to other fields. For example, topological settable systems may apply to study causality in the spatial-temporal manifolds used to analyze neural activity in the brain (Roebroek, et al., 2011; Valdés-Sosa, Bornot-Sánchez, et al., 2011; Valdés-Sosa, Roebroek, et al., 2011).

The plan of the paper is as follows. Section 2 specifies the details of the  $N$ -bidder first-price private-value auction that provides our game-theoretic focus throughout. In Section 3, we introduce basic elements of topological settable systems and relate these to individually rational behavior in the  $N$ -bidder first-price private-value auction, resolving several of the causal questions posed above. In Section 4, we introduce further elements of topological settable systems and relate these to Bayesian-Nash equilibrium, resolving the remaining questions. In particular, we introduce the notions of comparable and compatible settable systems and employ these to distinguish mutual

---

<sup>3</sup>See WC, White, Chalak, and Lu (2011), and Chen and White (1998).

<sup>4</sup>Recall that two spaces are homeomorphic if there exists a homeomorphism between them, that is, a one-to-one function continuous in both directions.

consistency conditions from structural equations. Section 5 contains further discussion. Section 6 summarizes and concludes by providing explicit answers to each of the causal questions posed above. An Appendix contains supplementary material.

## 2 An $N$ -Bidder Private-Value Auction

We consider the first-price private-value auction studied by Guerre, Perrigne, and Vuong (2000) and others in the empirical auction literature and treated by Krishna (2010, chapter 2). There is a single object for sale, and  $1 < N < \infty$  potential buyers bid for the object. The highest bidder pays the amount they bid (first price) and gets the object. If there is a tie between two or more bidders, then the winner can be decided randomly, e.g., by drawing lots. It is assumed that there are no liquidity constraints for any bidder. Thus, each bidder is both willing and able to pay the amount bid. For simplicity, we assume the seller is non-strategic: she imposes no entry fee, reserve price, or other requirements.

Bidder  $i$  assigns a value of  $X_i$  to the object – this is the maximum amount the bidder is willing to pay. The value  $X_i$  corresponds to Harsányi’s (1967) notion of agent *type*. The  $X_i$ ’s are random and assumed to be independently and identically distributed on an interval  $[\underline{v}, \bar{v}]$  ( $\underline{v}, \bar{v} \in \mathbb{R}_+ \equiv [0, \infty)$ ) according to the strictly increasing distribution function  $F$  with full support on  $[\underline{v}, \bar{v}]$ , so that  $\mathbb{E}X_i < \infty$ . For simplicity and without essential loss of generality, we let  $[\underline{v}, \bar{v}]$  be the unit interval,  $\mathbb{I} \equiv [0, 1]$ .  $F$  is also assumed to admit a continuous density  $f \equiv F'$ . Each bidder is assumed to know that her and the other bidders’ values are independently distributed according to  $F$ . We also assume that  $N$  is commonly known.

Bidder  $i$ ’s *strategy* is a function  $\beta_i$  mapping types to *actions* (bids),<sup>5</sup> determining her bid for every realized value  $x_i$  of  $X_i$ . Bidders are risk neutral: they maximize the expected value of their surplus. This surplus is  $X_i - \beta_i(X_i)$  if player  $i$  wins and 0 otherwise. Because negative surplus can always be avoided, we require  $\beta_i(X_i) \leq X_i$ , so  $\beta_i$  maps  $\mathbb{I}$  to  $\mathbb{I}$ . As shown in the Appendix, non-decreasing (“monotone”) strategies are never worse than non-monotone strategies, so we restrict attention to strategies belonging to  $\mathcal{M}$ , the class of monotone functions from  $\mathbb{I}$  to  $\mathbb{I}$ .

First, consider ex ante individually rational behavior, where bidder  $i$  determines her best response strategy taking other bidders’ strategies as given and without knowledge of her type. To state bidder  $i$ ’s optimization problem, let  $\beta_{-i}$  denote the other bidders’ *strategy profile*; this is the  $(N - 1)$ -vector containing all strategies  $\beta_j$ ,  $j \neq i$ . Next, define

$$G_i(b; \beta_{-i}) \equiv \mathbb{P} \left\{ b > \max_{j \neq i} \beta_j(X_j) \right\} + \mathbb{E} \left( 1 \left\{ b = \max_{j \neq i} \beta_j(X_j) \right\} \times \varphi_i(b, \beta_{-i}(X_{-i})) \right).$$

---

<sup>5</sup>This is in fact a *pure* strategy; *mixed* strategies involve random choice among pure strategies. For now, all strategies are pure strategies.

This is the probability that bidder  $i$  wins by bidding  $b$ , given the other bidders' strategy profile  $\beta_{-i}$  and the tie-breaking rule  $\varphi_i$ , which gives the probability that bidder  $i$  wins if there is a tie. Then bidder  $i$ 's ex ante problem is

$$\max_{\beta_i \in \mathcal{M}} \psi_i(\beta_i; \beta_{-i}) \equiv \mathbb{E} [G_i(\beta_i(X_i); \beta_{-i}) \times (X_i - \beta_i(X_i))]. \quad (1)$$

To ensure existence of a best response, we restrict the profiles to have strictly monotone elements, so that  $\beta_{-i} \in \times_{j \neq i} \mathcal{M}_s$ , where  $\mathcal{M}_s$  is the set of strictly increasing functions from  $\mathbb{I}$  to  $\mathbb{I}$ .<sup>6</sup> To verify existence, we first note that when profiles  $\beta_{-i}$  belong to  $\times_{j \neq i} \mathcal{M}_s$ , ties occur with probability zero, and solving the problem (1) is equivalent to solving

$$\max_{\beta_i \in \mathcal{M}} \mathbb{E} \left[ \mathbb{P} \left\{ \beta_i(X_i) \geq \max_{j \neq i} \beta_j(X_j) \mid X_i \right\} \times (X_i - \beta_i(X_i)) \right]. \quad (2)$$

By lemma 7 of van Zandt and Vives (2007) or from Milgrom and Shannon's (1994) Monotonicity Theorem, this problem has a solution for every  $\beta_{-i} \in \times_{j \neq i} \mathcal{M}$ , and in particular for every  $\beta_{-i} \in \times_{j \neq i} \mathcal{M}_s$ .<sup>7</sup> When the solution is not unique, there is a measurable selection, as we verify in the Appendix. We let  $\mathbf{r}_i$  denote this selection. We call  $\mathbf{r}_i$  the *ex ante response function*, and we call  $\beta_i^* = \mathbf{r}_i(\beta_{-i})$  the *best response*.

Note the distinction between the strategies  $\beta_j \in \mathcal{M}_s$ ,  $j \neq i$ , which need not be optimal, and the optimal<sup>8</sup>  $\beta_i^* \in \mathcal{M}$ . That is,  $\beta_i^*$  is  $i$ 's individually rational strategy, whereas  $\beta_{-i}$  is a vector of others' possibly suboptimal strategies, which bidder  $i$  knows (i.e., correctly believes).

To express the bid determined by the best response, we observe that  $\mathbf{r}_i$  satisfies the following: for  $x_i \in \mathbb{I}$  almost surely (*a.s.*),

$$\mathbf{b}_i(\beta_{-i}, x_i) \equiv \mathbf{e}(\mathbf{r}_i(\beta_{-i}), x_i) \in \arg \max_{b_i \in \mathbb{I}} \mathbb{P} \left\{ b_i \geq \max_{j \neq i} \beta_j(X_j) \right\} \times (x_i - b_i),$$

where  $\mathbf{e}$  is the evaluation functional,  $\mathbf{e}(\beta, x) = \beta(x)$ . This represents bidder  $i$ 's optimal bid,  $b_i^* = \mathbf{b}_i(\beta_{-i}, x_i)$ , once she knows that her type is actually  $x_i$  and given other bidders' strategies  $\beta_{-i}$ . As bidder  $i$  never observes the other players' realized types, her payoff is not directly a function of these – hence the designation "private value."

Although ex ante the other bidder's strategies may or may not represent their best responses, in *Bayesian-Nash equilibrium*, each bidder's chosen strategy must be the best response to the other bidders' equilibrium strategies. This equilibrium condition is the fixed point requirement,

$$\beta_i^\dagger = \mathbf{r}_i(\beta_{-i}^\dagger), \quad i = 1, \dots, N, \quad (3)$$

<sup>6</sup>If one replaces  $\mathcal{M}_s$  with  $\mathcal{M}$ , ties can occur with non-zero probability, and best responses need not exist. We thank Philip Reny for his guidance.

<sup>7</sup>The solutions to (2) for  $\beta_{-i} \notin \times_{j \neq i} \mathcal{M}_s$  are not best responses and are not of interest here, apart from their serving to allow easy application of available existence results.

<sup>8</sup>In the PCM, the same symbol would be used to denote both of these very different objects. (See, e.g., Pearl (2000, p.203).) Our use of the superscript "\*" is intended to avoid this confusion.

where  $\beta_i^\dagger \in \mathcal{M}_s$  denotes bidder  $i$ 's equilibrium strategy. In Bayesian-Nash equilibrium, each bidder correctly believes that the other players will best respond. Thus, Bayesian-Nash equilibrium is a *rational belief* equilibrium. Equilibrium strategies yield equilibrium bids  $b_i^\dagger = \beta_i^\dagger(x_i)$ .

For this  $N$ -bidder independent private value auction, the existence of a unique strictly monotone symmetric Bayesian-Nash equilibrium is well known. Among others, Riley and Samuelson (1981) show that the equilibrium responses satisfy

$$\beta_i^\dagger(x_i) = x_i - \frac{1}{(F(x_i))^{N-1}} \int_0^{x_i} (F(u))^{N-1} du.$$

This auction game is now sufficiently specified that all the questions in the Introduction apply.

### 3 Topological Settable Systems and Individual Rationality

To address the causal questions of interest, we extend WC's settable systems to topological settable systems. This supports causal discourse for phenomena involving random variables taking values in general topological spaces, rather than just the real line. This makes it possible to talk sensibly about the causal effects of one agent's strategy on other agents' best responses. Some of the material below is unavoidably abstract, but we promptly follow each general concept with discussion relating these to the auction of Section 2. We draw on and refer the interested reader to Corbae, Stinchcombe, and Zeman (2009, ch.10) for topological background.

Recall that a *topological space* is a pair  $(\mathfrak{X}, \tau)$ , where  $\mathfrak{X}$  is a non-empty set and  $\tau$  is a collection of subsets of  $\mathfrak{X}$  containing  $\mathfrak{X}$  and  $\emptyset$ , closed under arbitrary unions and finite intersections. The elements of  $\tau$  are *open sets*. Topological spaces may or may not have associated norms or metrics.

For example, consider the strategy set  $\mathcal{M}$ , the set of non-decreasing functions from  $\mathbb{I}$  to  $\mathbb{I}$ . This set has a topology (described next), say  $\tau_{\mathcal{M}}$ . The topological space  $(\mathcal{M}, \tau_{\mathcal{M}})$  is known as *Helly space*, named after mathematician Eduard Helly (see, e.g., Kelley, 1975, p.164).  $\mathcal{M}$  is a closed subset of  $\mathbb{I}^{\mathbb{I}}$ , the set of all mappings from  $\mathbb{I}$  to  $\mathbb{I}$ . That set has an associated topology, say  $\tau^{\mathbb{I}}$ . We can take  $\tau_{\mathcal{M}}$  to be the *relative* topology, defined as the collection of all sets of the form  $\mathcal{M} \cap A$ , where  $A$  belongs to  $\tau^{\mathbb{I}}$ .  $\mathcal{M}_s$  is a subset of  $\mathcal{M}$  that is neither open nor closed, as the limit of a sequence of strictly increasing functions can be strictly increasing, but need not be. As  $\mathcal{M}$  is the closure of  $\mathcal{M}_s$ ,  $\mathcal{M}_s$  is dense in  $\mathcal{M}$ . Helly space  $\mathcal{M}$  is not a metric space.<sup>9</sup>

To define random elements of topological spaces, we let  $\mathcal{B}(\mathfrak{X}, \tau)$  denote the Borel  $\sigma$ -field generated by the topology; this is the smallest collection of subsets of  $\mathfrak{X}$  that includes all elements of  $\tau$ , the complement of any set in  $\mathcal{B}(\mathfrak{X}, \tau)$ , and the countable union of any collection of sets in

---

<sup>9</sup>It is also known that Helly space is compact convex Hausdorff, first (but not second) countable, and separable (Kelley, 1975, p.164).

$\mathcal{B}(\mathfrak{X}, \tau)$ . Let  $(\Omega, \mathcal{F})$  be a measurable space. Random elements of  $(\mathfrak{X}, \tau)$  are mappings, say  $Z : \Omega \rightarrow \mathfrak{X}$ , that are measurable— $\mathcal{F}/\mathcal{B}(\mathfrak{X}, \tau)$ , i.e., for all  $B \in \mathcal{B}(\mathfrak{X}, \tau)$ ,  $Z^{-1}(B) \equiv \{\omega : Z(\omega) \in B\} \in \mathcal{F}$ .

Random scalars and vectors are simple examples of this concept. For example, for types,  $X_i : \Omega \rightarrow \mathbb{I}$ , take  $\mathfrak{X} = \mathbb{I}$  and let  $\tau_{\mathbb{I}}$  be the usual open sets of  $\mathbb{I}$ . Another example is that of a *random strategy*. This is a mapping  $\mathfrak{B}$  from  $\Omega$  to  $\mathcal{M}$ , say, such that  $\beta = \mathfrak{B}(\omega)$  belongs to  $\mathcal{M}$  for each  $\omega$  in  $\Omega$  with the property that for all “strategy events”  $M \in \mathcal{B}(\mathcal{M}, \tau_{\mathcal{M}})$ ,<sup>10</sup>  $\mathfrak{B}^{-1}(M) \equiv \{\omega : \mathfrak{B}(\omega) \in M\}$  belongs to  $\mathcal{F}$ . If  $P$  is a probability measure on  $(\Omega, \mathcal{F})$ ,  $M$  is a strategy event, and  $\mathfrak{B}$  is a random strategy, then we can define the probability of  $M$  under  $\mathfrak{B}$  as

$$\mathbb{P}_{\mathfrak{B}}[M] \equiv P[\mathfrak{B} \in M] = P\{\omega : \mathfrak{B}(\omega) \in M\}.$$

For example, a *measurable selection* from a strategy event  $M$ , say  $s_M : \Omega \rightarrow \mathcal{M}$ , gives a random strategy  $\mathfrak{B}$  such that  $\mathfrak{B}(\omega) = s_M(\omega) \in M$  for all  $\omega \in \Omega$ . Thus, even when there are multiple individually rational or equilibrium solutions, whenever a measurable selection exists, we can apply it to the solution set to obtain a single solution, a *randomized* strategy. This is especially relevant for games with correlated equilibria. As Nash equilibrium suffices for our auction, we will not immediately require these, but they suggest the scope of topological settable systems.

We also use the concept of the product topology. Let  $\mathbb{N}^+$  denote the positive integers, and define  $\bar{\mathbb{N}}^+ \equiv \mathbb{N}^+ \cup \{\infty\}$ . When  $n = \infty$ , we interpret  $i = 1, \dots, n$  as  $i = 1, 2, \dots$ . Let  $n \in \bar{\mathbb{N}}^+$ , and for  $i = 1, \dots, n$ , let  $(\mathfrak{X}_i, \tau_i)$  be topological spaces. Define the Cartesian product  $\mathfrak{X}^n \equiv \times_{i=1}^n \mathfrak{X}_i$ ; the projection mapping  $\pi_i : \mathfrak{X}^n \rightarrow \mathfrak{X}_i$  is the mapping such that  $\pi_i(x^n) = x_i$  for each  $x^n = (x_1, \dots, x_n) \in \mathfrak{X}^n$ . The product topology, say  $\tau^n = \times_{i=1}^n \tau_i$ , is the smallest topology that makes each  $\pi_i, i = 1, \dots, n$ , continuous, i.e.,  $\pi_i^{-1}(\tau_i) \subset \tau^n$ . When  $n$  is countable infinity,  $\mathcal{B}(\mathfrak{X}^\infty, \tau^\infty)$  denotes the Borel  $\sigma$ -field generated by the measurable finite-dimensional product cylinders, i.e. the smallest collection of subsets of  $\mathfrak{X}^\infty$  that includes: (i) sets of the form  $\times_{i=1}^\infty B_i$ , where  $B_i \in \mathcal{B}(\mathfrak{X}_i, \tau_i)$ ,  $i = 1, 2, \dots$ , and  $B_i = \mathfrak{X}_i$  except for finitely many  $i$ ; (ii) the complement of any set in  $\mathcal{B}(\mathfrak{X}^\infty, \tau^\infty)$ ; and (iii) the countable union of any collection of sets in  $\mathcal{B}(\mathfrak{X}^\infty, \tau^\infty)$ .

For example, consider the ex ante response function,  $r_i : \times_{j \neq i} \mathcal{M}_s \rightarrow \mathcal{M}$ . We will require the domain (here,  $\mathcal{M}_s^{N-1} \equiv \times_{j \neq i} \mathcal{M}_s$ ) to be a Borel subset of the product space  $(\mathcal{M}^{N-1}, \tau_{\mathcal{M}}^{N-1})$ , say. This holds, as  $\mathcal{M}_s$  is a measurable subset of  $\mathcal{M}$ .<sup>11</sup> For the evaluation functional,  $e : \mathcal{M} \times \mathbb{I} \rightarrow \mathbb{I}$ , we require the domain  $\mathcal{M} \times \mathbb{I}$  to be a Borel set of the product space  $(\mathcal{M} \times \mathbb{I}, \tau_{\mathcal{M}} \times \tau_{\mathbb{I}})$ , which it obviously is. Below, we will require  $r_i$  and  $e$  to be measurable functions. We noted above that  $r_i$  is a measurable selection. The joint measurability of  $e$  is ensured by results of Aumann (1960), as  $\mathcal{M}$  is a bounded Banach class, as verified in the Appendix.

<sup>10</sup>We use the term "event" to denote a Borel measurable set. A "strategy set" could be any subset of  $\mathcal{M}$ .

<sup>11</sup>The complement of  $\mathcal{M}_s$  in  $\mathcal{M}$  can be represented as a countable union of closed (thus measurable) subsets of  $\mathcal{M}$ , namely the collection of functions in  $\mathcal{M}$  with flat area of width at least  $r_n \in \mathbb{Q} \cap \mathbb{I}$ , where  $\mathbb{Q}$  is the set of rational numbers. Thus,  $\mathcal{M} - \mathcal{M}_s$  is measurable, and so is  $\mathcal{M}_s$ .



### 3.1 Topological Settable Systems: Formal Definition

We now have sufficient background to introduce topological settable systems. We begin by providing a formal definition. We then discuss the component of this definition and illustrate topological settable systems using our auction example. We let  $\mathbb{N} \equiv \{0, 1, \dots\}$ , and define  $\bar{\mathbb{N}} \equiv \mathbb{N} \cup \{\infty\}$ . When  $m = 0$ , we ignore references to  $j = 1, \dots, m$ .

**Definition 3.1 (Partitioned Topological Settable System)** *A triple  $\mathcal{S} \equiv \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, \mu_{\mathbf{a}}), (\Pi, \mathcal{X}^{\Pi})\}$  is a **partitioned topological settable system** with components  $(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, \mu_{\mathbf{a}})$ , and  $(\Pi, \mathcal{X}^{\Pi})$ , defined as follows.*

*Let  $\mathbf{A}$  be a non-empty set and let **attributes**  $\mathbf{a} \in \mathbf{A}$  be given. Let  $(\Omega, \mathcal{F}, \mu_{\mathbf{a}})$  be a complete signed measure space such that  $\Omega$  contains at least two elements and  $\mu_{\mathbf{a}}$  is  $\sigma$ -finite and countably additive. For  $n \in \bar{\mathbb{N}}^+$ , let **units** have indexes  $i = 1, 2, \dots, n$ , and let  $\Pi = \{\Pi_b\}$  be a partition of  $\{1, \dots, n\}$ , with cardinality  $B \in \bar{\mathbb{N}}^+$ .*

*To define  $\mathcal{X}^{\Pi} \equiv (\mathcal{X}_0, \mathcal{X}_1^{\Pi}, \dots, \mathcal{X}_n^{\Pi})$ , let  $m \in \bar{\mathbb{N}}$ , let  $(\mathfrak{X}_{0,j}, \tau_{0,j})$  be topological spaces,  $j = 1, \dots, m$ , let  $\mathbb{S}_0$  be a multi-element Borel subset of the product topological space  $(\mathfrak{X}_0^m, \tau_0^m)$ , and let the **fundamental settings**  $Z_0 : \Omega \rightarrow \mathbb{S}_0$  be measurable  $(-\mathcal{F}/\mathcal{B}(\mathfrak{X}_0^m, \tau_0^m))$ .*

*For units  $i = 1, 2, \dots, n$ , let  $(\mathfrak{X}_i, \tau_i)$  be a topological space, let  $\mathbb{S}_i$  be a multi-element subset of  $\mathcal{B}(\mathfrak{X}_i, \tau_i)$ , and let **settings**  $Z_i^{\Pi} : \Omega \rightarrow \mathbb{S}_i$  be measurable functions into  $\mathbb{S}_i$ . For **blocks**  $b = 1, \dots, B$ , let  $Z_{(b)}^{\Pi}$  be the vector containing  $Z_0$  and  $Z_i^{\Pi}, i \notin \Pi_b$ , and taking values in the non-empty Borel set  $\mathbb{S}_{(b)}^{\Pi} \subseteq \mathbb{S}_0 \times_{i \notin \Pi_b} \mathbb{S}_i$ . For  $b = 1, \dots, B$ , let  $\mathbb{S}_{[b]}^{\Pi} \subseteq \times_{i \in \Pi_b} \mathbb{S}_i$  be a non-empty Borel set, and suppose there exists a measurable **joint response function**  $r_{[b]}^{\Pi}(\cdot; \mathbf{a}) : \mathbb{S}_{(b)}^{\Pi} \rightarrow \mathbb{S}_{[b]}^{\Pi}$  such that **responses**  $Y_{[b]}^{\Pi}(\cdot) \equiv (Y_i^{\Pi}(\cdot), i \in \Pi_b)$ , are jointly determined as*

$$Y_{[b]}^{\Pi} = r_{[b]}^{\Pi}(Z_{(b)}^{\Pi}; \mathbf{a}).$$

*For  $i = 1, \dots, n$ , define the **settable variables**  $\mathcal{X}_i^{\Pi} : \{0, 1\} \times \Omega \rightarrow \mathbb{S}_i$  as*

$$\mathcal{X}_i^{\Pi}(0, \omega) \equiv Y_i^{\Pi}(\omega) \quad \text{and} \quad \mathcal{X}_i^{\Pi}(1, \omega) \equiv Z_i^{\Pi}(\omega) \quad \omega \in \Omega,$$

*and let  $\mathcal{X}_0(0, \omega) \equiv Y_0(\omega) \equiv \mathcal{X}_0(1, \omega) \equiv Z_0(\omega), \omega \in \Omega$ . Put  $\mathcal{X}^{\Pi} \equiv \{\mathcal{X}_0, \mathcal{X}_1^{\Pi}, \mathcal{X}_2^{\Pi}, \dots\}$ .*

The definition of topological settable systems includes settable systems, originally defined by WC, as the special case where all the topological spaces are  $(\mathbb{R}, \tau_{\mathbb{R}})$ , where  $\tau_{\mathbb{R}}$  is the usual collection of open sets of  $\mathbb{R}$ . We call WC's settable systems *real* settable systems to distinguish them from topological settable systems. Because real settable systems also admit a countable number of units, they can accommodate settable variables taking values in topological spaces homeomorphic to  $(\mathbb{R}^{\infty}, \tau^{\infty})$ , the topological space of countable sequences of reals, a metrizable

space. The auction game of Section 2 involves  $(\mathcal{M}, \tau_{\mathcal{M}})$ . This space is not metrizable, so there can be no homeomorphism between  $(\mathcal{M}, \tau_{\mathcal{M}})$  and  $(\mathbb{R}^{\infty}, \tau^{\infty})$ .<sup>12</sup> That is, there is no way to work with real settable systems and replicate the topological properties of  $(\mathcal{M}, \tau_{\mathcal{M}})$ . WC's real settable systems therefore do not apply. But as we now show, our auction maps directly to topological settable systems.

A topological settable system is a triple  $\mathcal{S} \equiv \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, \mu_{\mathbf{a}}), (\Pi, \mathcal{X}^{\Pi})\}$ . The first component of  $\mathcal{S}$  specifies attributes  $\mathbf{a} \in \mathbf{A}$ . Attributes implicitly contain all fixed features of the system of interest, apart from  $\Pi$  discussed in what follows. For games, attributes can exhaustively specify the rules, including the number of players, their possible actions, their type spaces, and their utility/payoff functions,<sup>13</sup> together with other fixed aspects of the game, such as the form of equilibrium and any equilibrium selection mechanisms, among other things. In our auction, the given number of players,  $N$ , and the common distribution of types,  $F$ , are salient elements of  $\mathbf{a}$ .

The second component of  $\mathcal{S}$  is the measure space  $(\Omega, \mathcal{F}, \mu_{\mathbf{a}})$ . It is often convenient to take the signed measure  $\mu_{\mathbf{a}}$  to be a probability measure, in which case we write  $\mu_{\mathbf{a}}$  as  $P_{\mathbf{a}}$ . In our auction game, the probability measure  $P_{\mathbf{a}}$  ( $= \mathbb{P}$ ) is determined by  $F$  and  $N$  and by the assumed independence of types. When dependence of types is of interest, this dependence can be expressed as a component of  $\mathbf{a}$ , expressed, for example, as a specific choice of copula (Sklar, 1959; Nelsen, 1999).

The third component of  $\mathcal{S}$  is  $(\Pi, \mathcal{X}^{\Pi})$ . There are  $n$  units in  $\mathcal{S}$ , indexed by  $i = 1, 2, \dots, n$ , which  $\Pi$  partitions into blocks. For each unit  $i$ , there is a settable variable  $\mathcal{X}_i^{\Pi}$ . Roughly speaking, each settable variable  $\mathcal{X}_i^{\Pi}$  is formed of a setting  $\mathcal{X}_i^{\Pi}(1, \omega) \equiv Z_i^{\Pi}(\omega)$  and a response  $\mathcal{X}_i^{\Pi}(0, \omega) \equiv Y_i^{\Pi}(\omega)$ , with  $\omega \in \Omega$ , and  $Y_i^{\Pi}$  may depend on settings of other system variables specified by the partition  $\Pi$ . Further,  $\mathcal{S}$  may admit fundamental settable variables  $\mathcal{X}_0$  whose responses do not depend on settings of other system variables. Nevertheless, the fundamental settings  $Z_0$  of  $\mathcal{X}_0$  can influence the responses of other system variables. Next, we discuss settings and responses of settable variables in more detail and illustrate these using our auction example.

In the auction example, the vector of types,  $X \equiv (X_1, \dots, X_N)$ , corresponds to the fundamental settings,  $Z_0$ , as these are determined outside the system. For the agent types, the spaces  $(\mathfrak{X}_{0,j}, \tau_{0,j}) = (\mathbb{I}, \tau_{\mathbb{I}})$  are implicitly components of  $\mathbf{a}$ ; the product topological space for the types is  $(\mathfrak{X}_0^m, \tau_0^m) = (\mathbb{I}^N, \tau_{\mathbb{I}}^N)$ . Thus, in this example  $m = N$ . The admissible values for the fundamental settings coincide with  $\mathbb{S}_0 = \mathbb{I}^N$ . The responses  $Y_0$  of  $\mathcal{X}_0$  are defined by

<sup>12</sup>If there were a homeomorphism, it could be used together with the metric on  $(\mathbb{R}^{\infty}, \tau^{\infty})$  to define a metric for  $(\mathcal{M}, \tau_{\mathcal{M}})$ . But this is impossible.

<sup>13</sup>It is standard to represent the structure of Bayesian games as  $\Gamma^b \equiv (\mathbb{N}, (\mathcal{A}_i)_{i \in \mathbb{N}}, (\mathcal{T}_i)_{i \in \mathbb{N}}, F, (u_i)_{i \in \mathbb{N}})$  (e.g., Myerson, 1997), where  $\mathbb{N}$  is the set of players (here,  $\mathbb{N} = \{1, \dots, N\}$ ),  $\mathcal{A}_i$  is player  $i$ 's set of possible actions (here,  $\mathcal{A}_i = \mathbb{I}$ ),  $\mathcal{T}_i$  is player  $i$ 's type space (here  $\mathcal{T}_i = \mathbb{I}$ ), and  $u_i$  is player  $i$ 's payoff/utility function. Thus, the components of  $\Gamma^b$  are components of  $\mathbf{a}$ .

$\mathcal{X}_0(0, \omega) \equiv Y_0(\omega) \equiv \mathcal{X}_0(1, \omega) \equiv Z_0(\omega)$ , with  $\mathcal{X}_0$  not depending on settings of other variables.

In our auction with  $N$  bidders, we designate two units for each player, a “strategy” unit and a “bid” unit.<sup>14</sup> Thus, here  $n = 2N$ . For each unit  $i$ , there is a setting  $Z_i^\Pi : \Omega \rightarrow \mathbb{S}_i$  with  $\mathbb{S}_i$  a subset of  $\mathcal{B}(\mathfrak{X}_i, \tau_i)$ . In the auction, for strategy units ( $i = 1, \dots, N$ ), the topological space is  $(\mathfrak{X}_i, \tau_i) = (\mathcal{M}, \tau_{\mathcal{M}})$ .<sup>15</sup> For bid units ( $i = N + 1, \dots, 2N$ ), the topological space is  $(\mathfrak{X}_i, \tau_i) = (\mathbb{I}, \tau_{\mathbb{I}})$ . These spaces are also implicitly components of  $\mathbf{a}$ . Here,  $\mathbb{S}_i = \mathcal{M}$  includes the admissible strategies, and  $\mathbb{S}_i = \mathbb{I}$  represents the admissible bids, implicitly components of  $\mathbf{a}$  and clearly elements of  $\mathcal{B}(\mathcal{M}, \tau_{\mathcal{M}})$  and  $\mathcal{B}(\mathbb{I}, \tau_{\mathbb{I}})$ , respectively.

The partition  $\Pi$  acts to specify blocks of units that jointly respond to settings of system units outside their block. The *elementary* partition is  $\Pi^e = \{\{1\}, \dots, \{n\}\}$ . The elementary settable system  $\mathcal{S}^e$  has this partition and represents the response of each individual unit to every other unit of the system. To examine the individually rational case in the auction, we can use the *agent* partition,  $\Pi^a = \{\Pi_b^a, b = 1, \dots, N\}$ , where block  $b = i$  has  $\Pi_b^a = \{i, i + N\}$ . The  $i$ th element of this partition groups together all responses governed by agent  $i$ . The corresponding settable system describes how each agent’s individually rational optimal strategy and optimal bid jointly depend on settings for all other agents. Other partitions, discussed below, correspond to jointly rational equilibrium.

The *joint response function*  $r_{[b]}^\Pi(\cdot; \mathbf{a})$  specifies how the settings outside block  $b$ ,  $Z_{(b)}^\Pi$ , determine the joint responses inside block  $b$ ,  $Y_{[b]}^\Pi$ . The distinction between the structurally determined responses (left-hand-side variables,  $Y_{[b]}^\Pi$ ) and the settings (right-hand-side variables,  $Z_{(b)}^\Pi$ ), which can take any admissible value, enforces the Strotz-Wold (1960) dichotomy between left-hand side and right-hand side variables in structural equations: the same variables never appear on both the right- and left-hand sides of the system’s structural equations. This rules out simultaneity.

Settings, response functions, and responses are partition specific. For the agent partition in the auction,  $\Pi^a$ , block  $b$  references bidder  $i$ , and  $r_{[i]}^a = r_{[i]}^{\Pi^a} = (r_i^a, r_{i+N}^a)$ , say, where strategy units have responses

$$r_i^a(z_{(i)}^a; \mathbf{a}) = \mathbf{r}_i(\beta_{-i}), \quad i \in \{1, \dots, N\},$$

and bid units have responses

$$r_{i+N}^a(z_{(i)}^a; \mathbf{a}) = \mathbf{b}_i(\beta_{-i}, x_i), \quad i \in \{1, \dots, N\}.$$

By definition,  $z_{(i)}^a = z_{(i)}^{\Pi^a}$  contains setting values outside block  $i$ , that is, setting values for other agents’ arbitrary strategies and bids, and setting values for types, say,<sup>16</sup>  $z_{(i)}^a = (\beta_{-i}, b_{-i}, x)$ .

<sup>14</sup>We explicitly consider both strategies and bids to provide a complete causal account, permitting us to address and resolve all the causal questions raised in the introduction.

<sup>15</sup>Below, we suitably restrict the domains and co-domains.

<sup>16</sup>For clarity and notational simplicity, we write  $z_{(i)}^a = (\beta_{-i}, b_{-i}, x)$ , omitting the agent partition superscript “ $a$ ”

Observe that  $r_{[i]}^a(\cdot; \mathbf{a})$  makes explicit the dependence of the response on  $\mathbf{a}$ , whereas the dependence of  $\mathbf{r}_i$  and  $\mathbf{b}_i$  on  $N$  and  $F$  is implicit. Also, observe that whereas  $r_{[i]}^a(z_{(i)}^a; \mathbf{a})$  allows the response to potentially depend on settings for every unit except those in  $\Pi_i^a$ , the economics of the auction game dictate that the optimal choice for strategy unit  $i$  depends only on the settings of strategies for the other strategy units,  $\beta_{-i}$ , and not on any of the agent values (types) or other agents' bid settings. Similarly, the optimal choice for bid unit  $i + N$  depends only on  $\beta_{-i}$  and the own type,  $x_i$ . The economics of the game thus impose variable exclusion structure.

For block  $b$ , the joint response function is given by  $r_{[b]}^\Pi(\cdot; \mathbf{a}) : \mathbb{S}_{(b)}^\Pi \rightarrow \mathbb{S}_{[b]}^\Pi$  and the sets  $\mathbb{S}_{(b)}^\Pi$  and  $\mathbb{S}_{[b]}^\Pi$  can impose partition-specific restrictions on the admissible values for  $Z_{(b)}^\Pi$  and  $Y_{[b]}^\Pi$ , respectively. Under the agent partition in the auction, we have  $z_{(i)}^a = (\beta_{-i}, b_{-i}, x) \in \mathbb{S}_{(i)}^{\Pi^a} = \mathcal{M}_s^{N-1} \times \mathbb{I}^{N-1} \times \mathbb{I}^N$ , enforcing strict monotonicity for the profile faced by a bidder. Because  $\mathcal{M}_s$  is a measurable subset of  $\mathcal{M}$ ,  $\mathbb{S}_{(i)}^{\Pi^a}$  is measurable as required. The joint responses take values in the unrestricted  $\mathbb{S}_{[i]}^{\Pi^a} = \mathcal{M} \times \mathbb{I}$ .

Thus, each aspect of individually rational behavior in the  $N$ -bidder first-price private-value auction maps to the agent partition of the settable systems framework, permitting us to define agent partition-specific settable variables  $\mathcal{X}^a \equiv \{\mathcal{X}_0, \mathcal{X}_1^a, \dots, \mathcal{X}_{2N}^a\}$  for this game.<sup>17</sup> Because of the structure imposed by the settable system, we refer to any of its elements or aspects as *structural*.

In particular, the response equations  $Y_{[b]}^\Pi = r_{[b]}^\Pi(Z_{(b)}^\Pi; \mathbf{a})$  are *structural equations*. In our auction game, the agent partition structural equations are

$$\begin{aligned} \beta_i^{a,*} &= \mathbf{r}_i(\beta_{-i}) \\ b_i^{a,*} &= \mathbf{b}_i(\beta_{-i}, x_i), \quad i = 1, \dots, N. \end{aligned} \tag{4}$$

We write  $\beta_i^{a,*}$  and  $b_i^{a,*}$  to make it clear that these are *agent partition* response values; these were written  $\beta_i^*$  and  $b_i^*$  in Section 2.

The fundamental settings,  $Z_0$ , are determined outside the system; they are thus *structurally exogenous*.<sup>18</sup> Here, the types  $X$  are structurally exogenous. The responses  $Y_{[b]}^\Pi$  are determined within the system; they are therefore *structurally endogenous*. The individually rational strategies and bids  $\beta_i^{a,*}$  and  $b_i^{a,*}$  are structurally endogenous.

---

from the formally correct  $(\beta_{-i}^a, b_{-i}^a, x)$ .

<sup>17</sup>It would be formally correct and more explicit to write  $\mathcal{X}_{\mathbf{a}}^a \equiv \{\mathcal{X}_{\mathbf{a},0}, \mathcal{X}_{\mathbf{a},1}^a, \dots, \mathcal{X}_{\mathbf{a},2N}^a\}$  instead of  $\mathcal{X}^a$ , making clear the dependence of the settable variables on the attributes. We leave this implicit to keep the notation simpler.

<sup>18</sup>This notion of structural exogeneity has no necessary relation to econometric notions of exogeneity, which involve various forms of stochastic orthogonality (e.g., independence or non-correlation) between observed regressors and unobserved "errors".

### 3.2 Cause and Effect in Topological Settable Systems

In topological settable systems, definitions of cause and effect are relative to a system and partition specific.

**Definition 3.2 (Intervention, Direct Cause, Direct Effect)** *Let  $\mathcal{S} \equiv \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, \mu_{\mathbf{a}}), (\Pi, \mathcal{X}^{\Pi})\}$  be a partitioned topological settable system. For  $b \in \{1, \dots, B\}$  and  $j \notin \Pi_b$ , let  $z_{(b)}$  and  $\tilde{z}_{(b);j}$  be distinct and admissible, i.e.,  $z_{(b)}, \tilde{z}_{(b);j} \in \mathbb{S}_{(b)}^{\Pi}$ , with  $z_{(b)}$  and  $\tilde{z}_{(b);j}$  identical, except that  $z_{(b)}$  has a component  $z_j$ , whereas  $\tilde{z}_{(b);j}$  has  $\tilde{z}_j \neq z_j$  instead. Then we say the ordered pair  $(z_{(b)}, \tilde{z}_{(b);j})$ , denoted  $z_{(b)} \rightarrow \tilde{z}_{(b);j}$ , is an **intervention to  $Z_{(b)}^{\Pi}$** .*

*Let  $i \in \Pi_b, j \notin \Pi_b$ . Then  $\mathcal{X}_i^{\Pi}$  **directly causes  $\mathcal{X}_j^{\Pi}$  in  $\mathcal{S}$**  if there exists  $z_{(b)} \rightarrow \tilde{z}_{(b);j}$  such that*

$$r_i^{\Pi}(\tilde{z}_{(b);j}; \mathbf{a}) \neq r_i^{\Pi}(z_{(b)}; \mathbf{a}),$$

*where  $r_{[b]}^{\Pi}(\tilde{z}_{(b);j}; \mathbf{a})$  and  $r_{[b]}^{\Pi}(z_{(b)}; \mathbf{a})$  are both admissible, i.e.,  $r_{[b]}^{\Pi}(\tilde{z}_{(b);j}; \mathbf{a}), r_{[b]}^{\Pi}(z_{(b)}; \mathbf{a}) \in \mathbb{S}_{[b]}^{\Pi}$ . Otherwise, we say  $\mathcal{X}_j^{\Pi}$  **does not directly cause  $\mathcal{X}_i^{\Pi}$  in  $\mathcal{S}$** .*

*The ordered pair  $(y_i, \tilde{y}_{i,j}) = (r_i^{\Pi}(z_{(b)}; \mathbf{a}), r_i^{\Pi}(\tilde{z}_{(b);j}; \mathbf{a}))$  is the **direct effect of  $\mathcal{X}_j^{\Pi}$  on  $\mathcal{X}_i^{\Pi}$  in  $\mathcal{S}$**  of  $z_{(b)} \rightarrow \tilde{z}_{(b);j}$ .*

According to this definition, causality is structural functional dependence, given a suitable context. This now enables us to answer several of the questions posed in the Introduction. Specifically, agent types *can* play a causal role in our auction game, as fundamental variables can indeed act as direct causes under this definition. As we see from structural equations (4), however, in the agent partition, other agent types do not directly cause either strategy or bid responses for a given agent, as the agent-partition joint response functions  $r_{[i]}^a$  do not depend on  $x_{-i}$ . Nor does an agent's own type directly cause her strategy response, as  $r_i^a$  does not depend on  $x_i$ . But an agent's own type *does* directly cause her bid response.<sup>19</sup>

Also, the individually rational strategy responses, obtained under the agent partition, formalize the intuitive directionality of best responses. In particular, we see that in the agent partition, other bidders' (suboptimal) strategies but not their actions (bids) directly cause agent  $i$ 's rational behavior, as both the strategy and bid responses depend on  $\beta_{-i}$  but not  $b_{-i}$ . Indeed, each bidder has enough information to determine precisely how her own strategy and action will causally affect every other bidders' rational responses here, despite the incomplete information and contrary to the Wikipedia entry quoted above. Only when the other players' response functions are unknown is a given bidder unable to determine the effects of her behavior. Rationality in behavior is a

<sup>19</sup>Strictly speaking, the presence of  $x_i$  in agent  $i$ 's bid response function implies only that the associated settable variable is a *potential* direct cause. In our auction game, however, this potential direct cause is generally an actual direct cause, as the bid response function generally is not constant in  $x_i$ .

sufficient but not necessary condition for other bidders' response functions to be known to every bidder.

This definition of partition-specific direct cause and effect admits the possibility of *mutual causality*, or what Strotz and Wold (1960) call *mutual causation*: we can have both  $\mathcal{X}_i^\Pi$  directly causing  $\mathcal{X}_j^\Pi$  and  $\mathcal{X}_j^\Pi$  directly causing  $\mathcal{X}_i^\Pi$ . Because of the distinction between settings and responses, however, there is no simultaneity or instantaneous causality. Framing causal statements in terms of settable variables enforces this distinction, making it explicit that causal relations do not simply hold between random variables or events (elements of  $\mathcal{F}$ ); rather, they hold between more structured objects, i.e., settable variables. In the agent partition for our auction game, we immediately see that mutual causality is present, as each agent's strategy directly causes every other agent's response. There is, however, no simultaneity.

By construction, the effects just defined are *ceteris paribus*:  $z_{(b)}$  and  $\tilde{z}_{(b);j}$  differ in only one component. This is only for simplicity and concreteness. Any pair  $(z_{(b)}, \tilde{z}_{(b)})$  of distinct admissible values is an admissible intervention, and the *joint* direct effect of this intervention is  $(y_{[b]}, \tilde{y}_{[b]}) = (r_{[b]}^\Pi(z_{(b)}; \mathbf{a}), r_{[b]}^\Pi(\tilde{z}_{(b)}; \mathbf{a}))$ . Although the term "intervention" suggests some mechanical process, there is no need for physical manipulation. As the definition makes clear, in settable systems, interventions are simply pairs of admissible values for the settings.

WC and Chalak and White (2012) define the effect of the intervention  $z_{(b)} \rightarrow \tilde{z}_{(b)}$  as the difference  $\tilde{y}_{[b]} - y_{[b]}$ . General topological spaces need not be vector spaces, so differences need not be defined for  $(\mathfrak{X}_i, \tau_i)$ , motivating our definition of effects as pairs of response values. Nevertheless, when  $\mathfrak{X}_i$  is a subset of a vector space, differences taking values in that vector space are defined and can be interpreted as effects, analogous to WC. For example, consider the strategy intervention  $\beta_{-i} \rightarrow \tilde{\beta}_{-i;j}$ , where  $\tilde{\beta}_{-i;j}$  differs from  $\beta_{-i}$  with respect to the strategy of bidder  $j$ . One effect of this intervention is  $(\mathbf{r}_i(\beta_{-i}), \mathbf{r}_i(\tilde{\beta}_{-i;j}))$ . The difference  $\mathbf{r}_i(\tilde{\beta}_{-i;j}) - \mathbf{r}_i(\beta_{-i})$  is generally not an element of  $\mathcal{M}$ , as it takes values in  $[-1, 1]$  and need not be monotonic. Nevertheless,  $\mathcal{M}$  is a subset of the set of bounded Borel-measurable functions from  $\mathbb{I}$  to  $\mathbb{R}$ , a vector space. The difference  $\Delta_{i,j} \equiv \mathbf{r}_i(\tilde{\beta}_{-i;j}) - \mathbf{r}_i(\beta_{-i})$  is thus an element of this set. The usual  $L_p$  norms can be defined for these differences, so the *magnitude* of the effect can be defined as the norm  $\|\Delta_{i,j}\|_p = [\int_0^1 |\Delta_{i,j}(x)|^p dQ(x)]^{1/p}$ , where  $Q$  is a specified probability measure on the measurable space  $(\mathbb{I}, \mathcal{B}(\mathbb{I}, \tau_{\mathbb{I}}))$ .

Observe that only variables can have effects. Constants, like  $N$  and  $F$  in the auction example, cannot take more than one value, so interventions are not possible (cf. Holland, 1986). Thus, here  $N$  and  $F$  do not have effects, resolving several more questions posed in the Introduction. The same is true for all system attributes,  $\mathbf{a}$ . Nevertheless, we emphasize that causal relations in TSS are relative to a system. If interest attaches to effects of  $F$ , say, then  $F$  must be removed from the system attributes and assigned status as a *variable* taking values in a suitable topological

space. For example,  $F$  can be a fundamental variable belonging to a collection of distributions defining measures absolutely continuous with respect to a given  $\sigma$ -finite non-negative measure  $\mu_{\mathbf{a}}$ . These distributions can be indexed by a finite- or infinite-dimensional parameter. Causality is therefore not only partition specific, but relative to the attributes chosen to describe a given phenomenon. Our treatment of causal relations as being relative to a settable system accords with the statement in Heckman (2005) that “causality is a property of a model of hypotheticals.”

## 4 Cause and Effect in Equilibrium

### 4.1 Comparable and Compatible Settable Systems

To relate individual best responses to jointly rational responses (Bayesian-Nash equilibrium), we use the concepts of comparable and compatible settable systems.

#### 4.1.1 Comparability

**Definition 4.1 (Comparability, Nesting, Finer and Coarser Partitions)** *Suppose topological settable systems  $\mathcal{S}^f$  and  $\mathcal{S}^c$  have identical  $(\mathbf{A}, \mathbf{a})$ . Then they are **comparable**.*

*Let comparable topological settable systems  $\mathcal{S}^f$  and  $\mathcal{S}^c$  have partitions  $\Pi^f = \{\Pi_1^f, \dots, \Pi_{B^f}^f\}$  and  $\Pi^c = \{\Pi_1^c, \dots, \Pi_{B^c}^c\}$ , respectively. If each element of  $\Pi^c$  is a union of elements of  $\Pi^f$ , then  $\Pi^c$  **neests**  $\Pi^f$ ,  $\Pi^c$  is **coarser than**  $\Pi^f$ ,  $\Pi^f$  is **finer than**  $\Pi^c$ , and we write  $\Pi^f \lesssim \Pi^c$ .*

When  $\Pi^f \lesssim \Pi^c$  and  $\Pi^c \lesssim \Pi^f$ , then  $\Pi^f$  and  $\Pi^c$  are the same,  $\Pi^f = \Pi^c$ .<sup>20</sup> When we discuss just two partitions  $\Pi^f$  and  $\Pi^c$  such that  $\Pi^f \neq \Pi^c$  and  $\Pi^f$  is finer than  $\Pi^c$ , we write  $\Pi^f \prec \Pi^c$ , and we call  $\Pi^f$  the *fine partition* and  $\Pi^c$  the *coarse partition*. The coarsest partition is the *global partition*,  $\Pi^g \equiv \{\{1, \dots, n\}\}$ . This nests every partition; in particular, it nests the finest partition, which is the elementary partition,  $\Pi^e \equiv \{\{1\}, \dots, \{n\}\}$ . We use the same nesting terminology for comparable  $\mathcal{S}^f$  and  $\mathcal{S}^c$ . For example, if  $\Pi^c$  nests  $\Pi^f$ , then we also say  $\mathcal{S}^c$  nests  $\mathcal{S}^f$  or that  $\mathcal{S}^f$  is finer than  $\mathcal{S}^c$ , and we write  $\mathcal{S}^f \lesssim \mathcal{S}^c$ .

In our incomplete information game, the elementary partition  $\Pi^e$  and the agent partition  $\Pi^a$  are associated with comparable topological settable systems  $\mathcal{S}^e$  and  $\mathcal{S}^a$ , say, with the coarser system  $\mathcal{S}^a$  nesting the finer system  $\mathcal{S}^e$ . This follows because  $\Pi^a = \{\Pi_i^a\}$ , with  $\Pi_b^a = \{b, b + N\} = \{b\} \cup \{b + N\}$ ; each element of the agent partition is the union of elementary partition elements.

<sup>20</sup>Let  $\mathbf{\Pi}$  denote the collection of all partitions of  $\{1, \dots, n\}$ . Then it is easily checked that  $(\mathbf{\Pi}, \lesssim)$  is a partially ordered set. As every partition has a greatest lower bound (inf),  $(\mathbf{\Pi}, \lesssim)$  is a complete lattice (e.g., Burris and Sankappanavar, 1981, theorem 4.2, p.17).

### 4.1.2 Compatibility

To motivate the definition of compatibility, first recall that for  $\Pi_b^c \in \Pi^c$ , we have

$$Y_{[b]}^c = r_{[b]}^c(Z_{(b)}^c; \mathbf{a}). \quad (5)$$

Next, define  $r_{[b]}^f \equiv (r_i^f, i \in \Pi_b^c)$ . This is the vector of response functions for the *fine* partition corresponding to the *coarse* partition element  $\Pi_b^c$ . Let  $\mathbb{S}_{(b)}^c$  denote the set of coarse partition admissible values for the settings  $Z_{(b)}^c$ , and suppose that for each  $z_{(b)}^c \in \mathbb{S}_{(b)}^c$  there is an admissible  $y_{[b]}^c (\in \mathbb{S}_{[b]}^c)$  such that<sup>21</sup>

$$y_{[b]}^c = r_{[b]}^f(y_{[b]}^c, z_{(b)}^c; \mathbf{a}). \quad (6)$$

When (6) holds, we see that  $y_{[b]}^c$  is a fixed point of what we call the *fine partition  $\Pi_b^c$ -system* (6). Significantly, and unlike the PCM, this fixed point need not be unique. As noted above,  $\mathbf{a}$  can embody equilibrium selection mechanisms, so it can specify which of possibly many fixed points for (6) is represented by (5).

We refer to (6) as a *mutual consistency* condition, as it ensures that the coarse partition is consistent with the fine partition in this specific way. Reinforcing the message of Strotz and Wold (1960), we emphasize that although these equations are precisely a system of simultaneous equations, they are *not* structural equations carrying causal meaning.<sup>22</sup> That meaning is carried by the structural equations of  $\mathcal{S}^f$  and  $\mathcal{S}^c$ . Causal discourse is valid in both  $\mathcal{S}^f$  and  $\mathcal{S}^c$  regardless of whether (6) holds.<sup>23</sup> Instead, the role of the simultaneous equations system (6) is to provide functional links between finer and coarser systems. These connections make causal discourse and economic explanations coherent between partitions in a precise sense.

To relate these ideas to our auction game, let the fine partition be the agent partition,  $\Pi^a$ , and let the coarse partition be the global partition,  $\Pi^g$ . Clearly,  $\Pi^g$  nests  $\Pi^a$ . The settable system  $\mathcal{S}^g$  corresponding to  $\Pi^g$  describes how all bidders' strategies and bids jointly respond to all bidders' types. The fixed point condition (6) has the form

$$\begin{aligned} \beta_i^{g,*} &= r_i^a(\beta_{-i}^{g,*}, b_{-i}^{g,*}, x; \mathbf{a}) = \mathbf{r}_i(\beta_{-i}^{g,*}) \\ b_i^{g,*} &= r_{i+N}^a(\beta_{-i}^{g,*}, b_{-i}^{g,*}, x; \mathbf{a}) = \mathbf{b}_i(\beta_{-i}^{g,*}, x_i), \quad i = 1, \dots, N, \end{aligned}$$

where we modify the arguments of  $r_i^a$  and  $r_{i+N}^a$  to make the dependencies clear.<sup>24</sup> These simul-

<sup>21</sup>We write  $r_{[b]}^f$  in this form for notational convenience. Although  $r_{[b]}^f$  as a whole depends on all elements of  $y_{[b]}^c$ , for given  $i \in \Pi_b^c$ ,  $r_i^f$  depends only on the subvector of  $y_{[b]}^c$  that omits the variables determined within the fine partition block containing  $i$ . Note also that  $y_{[b]}^c$  must take values admissible to the components of  $r_{[b]}^f$ .

<sup>22</sup>In a related context, Strotz and Wold (1960, p.423) state, "what are simultaneous equilibrium conditions ought not to be confused with causal relations."

<sup>23</sup>In sharp contrast, causal discourse in the PCM is defined *only* when the analog of (6) has a unique solution.

<sup>24</sup>To ensure that the global partition strategies are admissible to the agent partition response functions, we take  $\mathbb{S}_{[1]}^{\Pi^g} = \mathcal{M}_s^N \times \mathbb{I}^N$ . That is, only strictly monotone global partition strategies are allowed.



taneous equations correspond to the Bayesian-Nash equilibrium conditions (3). Drawing on the discussion above, we can now resolve another of the questions posed at the outset: the Bayesian-Nash equilibrium conditions (3) are mutual consistency conditions, *not* structural equations.

Given an admissible fixed point,  $\beta^\dagger$ , and assuming that  $\mathbf{a}$  specifies that responses for the global partition are governed by Bayesian-Nash equilibrium, it follows that  $\beta^{g,*} = \beta^\dagger$ . Then<sup>25</sup>

$$\begin{aligned}\beta_i^{g,*} &= r_i^g(x; \mathbf{a}) = \beta_i^\dagger \\ b_i^{g,*} &= r_{i+N}^g(x; \mathbf{a}) = \beta_i^\dagger(x_i), \quad i = 1, \dots, N.\end{aligned}\tag{7}$$

The responses of the agent partition and those of the global partition are then mutually consistent.

We formalize these ideas with the following definition.

**Definition 4.2 (Mutual Consistency/Compatibility)** *Let  $\mathcal{S}^f$  and  $\mathcal{S}^c$  be comparable topological settable systems with  $\mathcal{S}^f \lesssim \mathcal{S}^c$ . If (6) holds for all  $\Pi_b^c$ ,  $b = 1, \dots, B^c$ , then  $\mathcal{S}^f$  and  $\mathcal{S}^c$  are **mutually consistent** or **compatible** settable systems, and we write  $\mathcal{S}^f \lesssim_c \mathcal{S}^c$ .*

An important consequence of the cross-partition coherence afforded by compatibility is that it makes possible valid statements about cause and effect in unobservable fine partition (e.g., individual best response) structures using knowledge gained solely from observable coarse partition (e.g. Bayesian-Nash equilibrium) structures. Specifically, by linking response functions across partitions, compatibility ensures functional dependence between fine and coarse partition effects, generalizing the effect linkages studied in the classical identification problem for linear structural systems (e.g., Fisher, 1966). The Appendix contains an example providing further insight.

To examine the causal content of the equilibrium response function, we inspect eq. (7) and apply Definition 2.2. This shows that only the bidder’s own type (directly) causes the bidder’s action (bid) in the global partition. But what about the bidder’s equilibrium strategy  $\beta_i^{g,*}$ ? What causes it? The answer illustrates a subtle insight elucidated by the settable systems framework: *nothing* causes  $\beta_i^{g,*}$ . Nevertheless, the equilibrium strategy  $\beta_i^{g,*}$  is formally explained by the principles of Bayesian-Nash equilibrium, in which a bidder’s rational (optimizing) behavior and rational beliefs about others’ behavior, operating on the specifics of the game (the primitives of the system, as embodied in the attributes  $\mathbf{a}$ ), determine the bidder’s equilibrium strategy. In the deductive-nomological theory of explanation (e.g., Hempel and Oppenheim, 1948; Popper, 1959), to explain a phenomenon is to show that it is the logical consequence of the application of a specific set of governing principles to the primitives of a given system. In this sense, the explanatory answer to the question, “Why is  $\beta_i^{g,*} = \beta_i^\dagger$ ?” is that this is the logical consequence of the principles of

<sup>25</sup>For the equilibrium bid representation,  $b_i^{g,*} = \mathbf{b}_i(\beta_{-i}^{g,*}, x_i) = \mathbf{e}(\mathbf{r}_i(\beta_{-i}^{g,*}), x_i) = \mathbf{e}(\beta_i^{g,*}, x_i) = \mathbf{e}(\beta_i^\dagger, x_i) = \beta_i^\dagger(x_i)$ .

rational action and belief underlying Bayesian-Nash equilibrium applied to the system primitives, including  $N$  and  $F$ . Interestingly, this is an example of a non-causal explanation.<sup>26</sup>

Equation (7), obtained under the global partition, formalizes the intuitive joint determination of equilibrium strategy responses. Correspondingly, the equation for  $\beta_i^{g,*}$  is an example of a structural equation in which causality is absent. This shows that, within the settable systems paradigm, structural relations are more basic than causal relations, as causal relations are necessarily structural, but structural relations need not be causal. With this understanding, we can resolve our questions at the outset about the structural and causal content of the “reduced form”: The reduced form is structural. In particular, both  $\beta_i^{g,*} = \beta_i^\dagger$  and  $b_i^{g,*} = b_i^\dagger$  have structural meaning as responses in a structural system; however, only  $b_i^{g,*}$  is causally determined.

## 4.2 Recursive and Canonical Topological Settable Systems

Causal discourse is especially straightforward in recursive systems. To describe these systems formally, for  $b > 0$ , we let  $\Pi_{[1:b]} \equiv \cup_{a=1}^b \Pi_a$  and  $\Pi_{[0:b]} \equiv \Pi_0 \cup \Pi_{[1:b]}$ , where  $\Pi_0 \equiv \{(0, 1), \dots, (0, m)\}$ . By convention,  $\Pi_{[0:0]} = \Pi_0$  and  $Z_{[0:0]} = Z_0$ .

**Definition 4.3 (Recursivity)** *Let  $\mathcal{S}$  be a partitioned topological settable system. For  $b = 0, 1, \dots, B$ , let  $Z_{[0:b]}^\Pi$  denote the vector containing the settings  $Z_i^\Pi$  for  $i \in \Pi_{[0:b]}$  and taking values in  $\mathbb{S}_{[0:b]} \subseteq \mathbb{S}_0 \times_{i \in \Pi_{[1:b]}} \mathbb{S}_i$ ,  $\mathbb{S}_{[0:b]} \neq \emptyset$ . For  $b = 1, \dots, B$ , suppose that  $r^\Pi \equiv \{r_{[b]}^\Pi\}$  is such that the responses  $Y_{[b]}^\Pi$  are jointly determined as*

$$Y_{[b]}^\Pi = r_{[b]}^\Pi(Z_{[0:b-1]}^\Pi; \mathbf{a}).$$

*Then we say that  $\Pi$  is a **recursive partition**,  $r^\Pi$  is **recursive**, and that  $\mathcal{S}$  is a **recursive topological settable system** or simply that  $\mathcal{S}$  is **recursive**.*

Recursive systems are also called *triangular* or *acyclical*. The global partition is always recursive, as  $B = 1$  and global partition responses are jointly determined as  $Y^g = r_{[1]}^{\Pi^g}(Z_{[0:0]}^g; \mathbf{a})$ . In the global partition, we see that only the fundamental settable variables  $\mathcal{X}_0$  are potential direct causes of the settable variables  $\mathcal{X}^g$ . This matches what we saw in Bayesian-Nash equilibrium for our auction game, where only types could potentially cause equilibrium strategies and bids.

A somewhat richer example involves the *strategy-action* partition,  $\Pi^s = \{\Pi_1^s, \Pi_2^s\}$ , where  $\Pi_1^s = \{1, \dots, N\}$  and  $\Pi_2^s = \{N + 1, \dots, 2N\}$ . The joint responses for  $\Pi_1^s$  are the jointly determined strategies  $\beta_i^{s,*}$ ,  $i = 1, \dots, N$ , and those for  $\Pi_2^s$  are the jointly determined bids  $b_i^{s,*}$ ,  $i = 1, \dots, N$ .

<sup>26</sup>There is an on-going debate in the philosophy of science about whether all explanations must be causal. Together, the deductive-nomological theory of explanation and the settable systems causal framework provide a particular context in which this issue can be resolved: non-causal explanations are possible. Further, the nomological-deductive theory provides a foundation for causal explanation.

Observe that the agent partition  $\Pi^a$  is not nested in  $\Pi^s$ , as there is no way to form  $\Pi_1^s$  as a union of elements of  $\Pi^a$ . But the elementary partition  $\Pi^e = \{\Pi_1^e, \dots, \Pi_{2N}^e\}$ ,  $\Pi_i^e = \{i\}$ , is nested in  $\Pi^s$ , as is easily checked; thus  $\mathcal{S}^e \preceq \mathcal{S}^s$ . We determine the responses for the strategy-action partition by requiring that these be mutually consistent with the elementary partition responses.

From eq.(2) and related discussion, the elementary partition strategy and bid responses are

$$\begin{aligned}\beta_i^{e,*} &= r_i^e(\beta_{-i}, b, x; a) = \mathbf{r}_i(\beta_{-i}) \\ b_i^{e,*} &= r_{i+N}^e(\beta, b_{-i}, x; a) = \mathbf{e}(\beta_i, x_i), \quad i = 1, \dots, N,\end{aligned}$$

where we have adapted the notation to make the elementary partition settings clear. As we have seen throughout, rationality in behavior and the underlying information constraints imply specific restrictions on the response functions in the form of exclusion restrictions.

Imposing compatibility between the strategy-action and elementary partitions gives

$$\begin{aligned}\beta_i^{s,*} &= r_i^e(\beta_{-i}^{s,*}, b^s, x; a) = \mathbf{r}_i(\beta_{-i}^{s,*}) \\ b_i^{s,*} &= r_{i+N}^e(\beta^s, b_{-i}^{s,*}, x; a) = \mathbf{e}(\beta_i^s, x_i), \quad i = 1, \dots, N,\end{aligned}$$

where we write  $b^s$  and  $\beta^s$  to denote arbitrary admissible strategy-action *settings* for bids and strategies, distinguishing these from the strategy-action *responses*  $b^{s,*}$  and  $\beta^{s,*}$  for bids and strategies. When  $\mathbf{a}$  specifies that responses for the strategy-action partition are governed by Bayesian-Nash equilibrium, the system above gives  $\beta^{s,*} = \beta^\dagger$ . The fixed point for the bid responses is trivial. Thus, the strategy-action settable system  $\mathcal{S}^s$  compatible with the elementary system  $\mathcal{S}^e$  is given by

$$\begin{aligned}\beta^{s,*} &= \beta^\dagger \\ b_i^{s,*} &= \mathbf{e}(\beta_i^s, x_i), \quad i = 1, \dots, N.\end{aligned}$$

Observe that although the strategy responses for the strategy-action partition are the same as those for the global partition, the bid response functions differ between the two. Only the own types can directly cause global partition bids. In contrast, both the own types and the own strategies can cause the strategy-action partition bids.

Also observe that the strategy-action partition is recursive, as responses in each block depend at most only on settings in “predecessor” blocks. That is, responses in block 1 depend at most on settings in block 0, and responses in block 2 depend at most on settings in blocks 0 and 1.

In recursive settable systems, mutual causality is absent. We can then represent the unimpeded evolution of the system by equating the settings for a given block with the responses for that block. We formalize this with the notion of *canonical* topological settable systems.

**Definition 4.4 (Canonical Topological Settable System)** *Let  $\mathcal{S}$  be a recursive topological settable system such that  $Z_{[b]}^{\Pi} = Y_{[b]}^{\Pi}$ ,  $b = 1, \dots, B$ . Then  $\mathcal{S}$  is a **canonical** topological settable system.*

The canonical version of the strategy-action partition is then

$$\begin{aligned}\beta^{s,*} &= \beta^{\dagger} \\ b_i^{s,*} &= e(\beta_i^{s,*}, x_i), \quad i = 1, \dots, N.\end{aligned}$$

Observe that the canonical strategy-action partition yields the same strategies and bids as the global partition, but there is still a subtle distinction: strategies can directly cause bids in the strategy-action partition, but not in the global partition.

Because mutual causality is absent and because settings and responses coincide in canonical settable systems, it is natural to simplify causal discourse by dropping explicit references to settable variables  $\mathcal{X}$  and, for  $a < b$ , instead simply speaking about the direct effects of, say,  $Y_{[a]}^{\Pi}$  on  $Y_{[b]}^{\Pi}$  or of  $y_{[a]}^{\Pi}$  on  $y_{[b]}^{\Pi}$ . For example, in the canonical strategy-action partition, we can say that  $\beta_i^{s,*}$  and  $x_i$  are direct causes of  $b_i^{s,*}$ , whereas in the global partition only  $x_i$  is a direct cause of  $b_i^{g,*}$ .

Standard canonical settable systems support natural definitions of indirect and total effects of  $Y_{[a]}^{\Pi}$  on  $Y_{[b]}^{\Pi}$ , as Chalak and White (2012) show. The patterns of total and indirect effects can then imply specific probabilistic conditional independence relations, supporting recovery of structural/causal information from observed data. The notions of indirect and total effects extend to canonical topological settable systems, but we leave this analysis aside here for brevity.

## 5 Discussion

As we have seen, neither the PCM nor the real settable systems framework of WC is able to support causal discourse in the  $N$ -bidder private-information auction game. In contrast, topological settable systems readily apply. Given the non-metrizable nature of the topological spaces required for this application, topological settable systems are not more general than necessary.

There is, however, one feature of topological settable systems that we did not exploit in mapping the private-value auction game to our causal framework. This is the possibility of having a countable number of structurally exogenous (fundamental) variables or endogenous variables. This feature, however, is not superfluous. It is required in a variety of economic contexts, not only in repeated games, time-series modeling, adaptive learning, and rational expectations modeling (see WC, section 7; Chen and White, 1998;<sup>27</sup> and White, Al-Sadoon, and Chalak, 2011), but also

---

<sup>27</sup>Among their examples of recursive learning in Banach spaces, Chen and White (1998) consider a learning agent

in Bayesian games of incomplete information with countably infinite-dimensional type and action spaces, as in Reny’s (2011) equilibrium existence theory for monotone pure-strategy Bayesian games. In fact, it may be possible to map topological setttable systems to Reny’s general setup, supporting causal discourse there. We forego this here, as the impressive generality of Reny’s framework would interfere with our desire to make our discussion relatively transparent.

Pure-strategy equilibrium exists when, among other things, types do not contain atoms – that is, types do not take some value with positive probability. As Reny (2011) notes, when type spaces have atoms, pure-strategy equilibrium need not exist; however, mixed-strategy equilibrium may exist under suitable conditions. Plausibly, topological setttable systems also apply to mixed-strategy games. The key to showing this is to specify appropriate topological spaces to which the mixed strategy settings and responses belong.

To see the basic idea, suppose that mixed-strategy bidders randomly choose among a countable number of pure strategies  $\beta_{\mathfrak{s}}$ ,  $\mathfrak{s} = 1, 2, \dots$ . We can represent this random choice as

$$\beta_S = \sum_{\mathfrak{s}=1}^{\infty} \beta_{\mathfrak{s}} 1\{S = \mathfrak{s}\} = \sum_{\mathfrak{s}=1}^{\infty} \beta_{\mathfrak{s}} 1\{\zeta(X, \cdot) = \mathfrak{s}\},$$

where  $1\{\cdot\}$  is the indicator function and  $S \equiv \zeta(X(\cdot), \cdot)$  is a random integer whose realization  $\mathfrak{s} = \zeta(X(\omega), \omega)$  determines which strategy is played. A new feature here is the “strategy choice” function  $\zeta : \mathbb{I} \times \Omega \rightarrow \mathbb{N}^+$ ; here, we permit agent type<sup>28</sup> to influence the probability of choosing a particular pure strategy. The mixed strategy settings  $\mathfrak{B} = \beta_S$  are thus random strategies, as defined in Section 2. The ex ante best response  $\mathfrak{B}^*$  can be written

$$\mathfrak{B}^* = \sum_{\mathfrak{s}=1}^{\infty} \beta_{\mathfrak{s}}^* 1\{\zeta^*(X, \cdot) = \mathfrak{s}\},$$

where  $\{\beta_1^*, \beta_2^*, \dots\}$  is the (selected) set of optimally chosen pure strategies and  $\zeta^*$  is the (selected) optimal strategy choice function.

From this, we see that mixed strategy settings can be viewed as elements of  $(\times_{\mathfrak{s}=1}^{\infty} \mathcal{M}_{\mathfrak{s}}) \times \mathbf{S}$ , where  $\mathbf{S}$  is a collection of measurable mappings  $\zeta$ , and that mixed strategy best responses take values in  $(\times_{\mathfrak{s}=1}^{\infty} \mathcal{M}) \times \mathbf{S}$ . To ensure that topological setttable systems apply, it suffices that  $\mathbf{S}$  has a topology, since we can then take the topology for  $(\times_{\mathfrak{s}=1}^{\infty} \mathcal{M}) \times \mathbf{S}$  to be the product topology. For example, if the elements of  $\mathbf{S}$  have finite expectation, then we can define the metric  $d_{\mathbf{S}}(S_1, S_2) \equiv E(|S_1 - S_2|) = \int |S_1(\omega) - S_2(\omega)| dP_{\mathbf{a}}(\omega)$  and let the topology for  $\mathbf{S}$  be the  $d_{\mathbf{S}}$ -metric topology.

Finally, we note that multiple equilibria are generic to realistic games. So far, we have only explicitly discussed one way of handling multiplicity in topological setttable systems, equilibrium

---

solving a stochastic dynamic programming problem and the game of fictitious play with continuum strategies, an infinitely repeated dynamic game of incomplete information.

<sup>28</sup>The type could be countably dimensioned, but we suppress this possibility here for simplicity.

selection mechanisms. An important complementary method is equilibrium refinement. This is concerned with eliminating “unreasonable” equilibria, leaving only equilibria that are “self-enforcing” in some sense. Leading equilibrium refinement approaches include sequential rationality (Kreps and Wilson 1982), perfectness (Selten, 1974), properness (Myerson, 1978), and strategic stability (Kohlberg and Mertens, 1986), among others. Topological settable systems can readily incorporate equilibrium refinement. We noted earlier that a system’s attributes,  $\mathbf{a}$ , can include a Bayesian game’s structure,  $\Gamma^b$ . Equilibrium refinement typically specifies certain necessary properties of the limit, suitably defined, of a sequence of games  $\{\Gamma_k^b\}$ , where  $k$  is a “refinement index”. Refinement amounts to the convention that players will play only those equilibria having the specified necessary limit properties. Both this convention and the sequence  $\{\Gamma_k^b\}$  can be included in the system attributes,  $\mathbf{a}$ .

As our discussion suggests, topological settable systems can support causal discourse in economics quite broadly. Nevertheless, there are important areas where topological settable systems, as formulated here, do not apply. For example, our framework does not apply to continuous-time structural models, such as those used to model asset price evolution in mathematical finance. The difficulty is that here our units are discretely indexed, but treating continuous time requires units to be continuously indexed. Extending the present *discrete* topological settable systems to the continuous case appears feasible and is an important direction for future research.

## 6 Conclusion

Despite the central role of causal discourse in explaining economic behavior, sufficiently general rigorous foundations for this discourse have so far been missing. We illustrate this lack in the familiar context of an  $N$ –bidder private-value auction, posing a variety of relevant causal questions that cannot be addressed within existing causal frameworks. We then introduce a new causal framework that delivers the missing answers, *topological settable systems*, an extension of the causal frameworks of Pearl (2000) and White and Chalak (2009). The examples in WC, White, Chalak, and Lu (2011), and White, Al-Sadoon, and Chalak (2011), together with our further discussion here, show the versatility of topological settable systems in supporting causal discourse in economics.

In the Introduction, we posed a variety of causal questions applicable to our auction game. We sum up by reviewing the answers emerging from our topological settable systems framework:

- *In what sense is a bidder’s strategy or action causally affected by other bidders’ strategies or actions?*

Without specifying which variables are jointly responding to the other variables of the system,

any discussion of causality is ambiguous – the sense in which causality may or may not be present requires specifying this. For example, answers can differ between the contexts of individual best response and Bayesian-Nash equilibrium. In the language of settable systems, causal discourse is partition specific. For individual best responses in our auction game (that is, in either the elementary or agent partition), other bidders’ suboptimal strategies directly affect a given agent’s strategy. On the other hand, in Bayesian-Nash equilibrium (i.e., the global partition) *nothing* causally affects agents’ strategies. There, strategies are non-causally explained by the game’s primitives, embedded in the system attributes,  $\mathbf{a}$ . Player actions have no effect on other bidders’ strategies or actions, whether for individually rational or equilibrium behavior, and the players know this.

- *How is the ceteris paribus “effect” of bidder  $j$ ’s strategy on bidder  $i$ ’s strategy defined?*

Because general topological spaces need not be vector spaces, we define an effect of a strategy intervention, for individually rational behavior, as a pair of corresponding best responses. Nevertheless, since the strategy space  $\mathcal{M}$  is a subset of the set of bounded Borel-measurable functions from  $\mathbb{I}$  to  $\mathbb{R}$ , the “magnitude” of this effect can be defined as the  $L_p$  norm distance between the pair of best responses.

- *How do rationality in behavior and belief matter for causal discourse?*

Rationality in behavior determines the individual best response functions for the agents. Because causal relations are properties of the response functions, rationality in behavior thus determines these causal relations. Requiring rationality in belief imposes the mutual consistency conditions of Bayesian-Nash equilibrium. This determines equilibrium structural relations and the causal relations there; it also links causal effects across partitions.

- *What is the causal role of types,  $X_i$ , if any?*

For individually rational behavior, own type directly causes an agent’s bid but not her strategy. In Bayesian-Nash equilibrium, the own type also causes an agent’s bid but not her strategy.

- *Do  $N$  and  $F$  have effects? If so, how? If not, why not?*

In the  $N$ –bidder private-information game considered here,  $N$  and  $F$  do not have causal effects. This is because  $N$  and  $F$  are fixed, not variable. They are therefore not subject to intervention, so effects cannot be defined for them. In other contexts where  $N$  and/or  $F$  are variable, then appropriate effects can be defined for  $N$  and  $F$ . Causal relations not only are partition specific, but also are specific to the given attributes.

- *What are the structural equations here?*

The structural equations relate partition-specific responses to partition-specific settings. As Strotz and Wold (1960) prescribe, the structural equations can embody mutual causality but not instantaneous causality or simultaneity. For example, elementary partition responses are given by

$$\begin{aligned}\beta_i^{e,*} &= \mathbf{r}_i(\beta_{-i}) \\ b_i^{e,*} &= \mathbf{e}(\beta_i, x_i), \quad i = 1, \dots, N.\end{aligned}$$

- *Specifically, are the simultaneous equations of Bayesian-Nash equilibrium (3) structural?*

No – these simultaneous equations are not structural equations. They are constructed from structural equations to enforce mutually consistency or compatibility conditions across partitions.

- *How about the equilibrium “reduced form”? Is it structural?*

Yes – the equations of the equilibrium reduced form are structural. In particular, the equilibrium equations for strategies are structural, but causality is absent.

- *Do equilibrium strategies and actions ( $\beta_i^\dagger$  and  $b_i^\dagger$ ) have structural meaning and/or causal content?*

Both  $\beta_i^\dagger$  and  $b_i^\dagger$  have structural meaning as response values for agent  $i$ 's equilibrium strategy and bid, respectively.  $\beta_i^\dagger$  also defines the equilibrium (constant) strategy response function. Its constancy implies that strategies have no causes in equilibrium. Equilibrium bids are caused by agents' own types.

We emphasize that, of necessity, causal discourse is meaningful only within a well-defined causal framework. Thus, our answers to these causal questions should be viewed as valid solely within the context of topological settable systems. If causal inquiries can be addressed using this framework then this acts as validation for the settable systems framework. If economic structures of interest fall outside of topological settable systems then this provides motivation for formulating a more sensible or comprehensive alternative framework.

## Appendix

### A.1 Monotone strategies weakly dominate non-monotone strategies

We verify the claim in Section 2 that monotone strategies are never worse than non-monotone strategies. Specifically, for any non-monotone strategy  $\beta$ , we construct an explicit monotone



strategy  $\beta^m$  that weakly dominates it. If  $\beta$  is monotone, then the identical construction gives  $\beta^m = \beta$  a.e. For brevity, we omit the straightforward demonstration of this last fact. The conditions and notation of Section 2 apply here unless explicitly noted otherwise.

**Proposition A.1 (Weak Domination):** *Let  $\beta_{-i} : \mathbb{I}^{N-1} \rightarrow \mathbb{I}^{N-1}$  be a given profile with measurable components that are not necessarily monotone, and let  $\beta : \mathbb{I} \rightarrow \mathbb{I}$  be a non-monotone measurable strategy. Then there exists a monotone strategy  $\beta^m : \mathbb{I} \rightarrow \mathbb{I}$  such that  $\psi_i(\beta^m; \beta_{-i}) \geq \psi_i(\beta; \beta_{-i})$ .*

**Proof:** For convenience, write the probability that bidder  $i$  wins by bidding  $b$  as  $G(b) = G_i(b; \beta_{-i})$ . Note that  $G$  is a (weakly) monotone function in  $b$ , although it need not be either left or right continuous. The objective function  $\psi_i(\beta; \beta_{-i}) \equiv \mathbb{E}[G(\beta(X_i)) \times (X_i - \beta(X_i))]$  is well defined for non-monotone  $\beta$  and  $\beta_{-i}$ , provided these are measurable, as assumed.

Let  $Z \equiv \beta(X_i)$ , and let  $H$  denote the CDF of  $Z$ , so that  $\Pr(Z \leq z) = H(z)$  for all  $z \in [0, 1]$ . Now we construct  $\beta^m : \mathbb{I} \rightarrow \mathbb{I}$  satisfying: (i)  $\beta^m$  is monotone; and (ii)  $Z$  and  $Z^m \equiv \beta^m(X_i)$  have the same distribution. For all  $p \in (0, 1)$ , define the quantile function

$$H^{-1}(p) \equiv \inf\{z \in [0, 1] : p \leq H(z)\}.$$

Observe that this is well defined for all CDFs  $H$ ;  $Z$  need not be continuous. We then define

$$\beta^m(x) \equiv H^{-1}(F(x)),$$

where  $F$  is the (strictly increasing) distribution of  $X_i$ . As  $\beta^m$  is the composition of monotone functions, it is also monotone. With  $Z^m \equiv \beta^m(X_i)$ , we have  $H^m(z) \equiv P[Z^m \leq z] = P[H^{-1}(F(X_i)) \leq z] = P[H^{-1}(U) \leq z]$ , where  $U \equiv F(X_i)$  has the uniform distribution on  $(0, 1)$  by the probability integral transform theorem. It follows from the quantile function theorem of Angus (1994, theorem 2) that  $H^{-1}(U)$  has the CDF  $H$ . Thus  $H^m = H$ .

Because  $Z^m$  and  $Z$  have the same distribution, then

$$\mathbb{E}[G(Z^m) Z^m] = \mathbb{E}[G(Z) Z].$$

It thus suffices to show

$$\mathbb{E}[G(Z^m) X_i] \geq \mathbb{E}[G(Z) X_i].$$

To show this, we write

$$\mathbb{E}([G(Z^m) - G(Z)] X_i) = \int [G(\beta^m(x)) - G(\beta(x))] x dF(x).$$

Let  $u = F(x)$ ,  $\gamma^m(u) = \beta^m(F^{-1}(u))$ ,  $\gamma(u) = \beta(F^{-1}(u))$ . Then

$$\mathbb{E}([G(Z^m) - G(Z)] X_i) = \int_0^1 [G(\gamma^m(u)) - G(\gamma(u))] F^{-1}(u) du.$$

For any Borel subset  $B$  of  $[0, 1]$ , define the measures  $\mu^m(B) \equiv \int_B G(\gamma^m(u)) du$  and  $\mu(B) \equiv \int_B G(\gamma(u)) du$ . By construction,  $\mu^m$  and  $\mu$  are absolutely continuous with respect to Lebesgue measure, with respective Radon-Nikodym derivatives  $d\mu^m/du = G(\gamma^m(\cdot))$  and  $d\mu/du = G(\gamma(\cdot))$ , uniquely defined *a.e.* (e.g., Bartle (1966, theorem 8.9)). This permits us to write

$$\mathbb{E}([G(Z^m) - G(Z)] X_i) = \int_0^1 F^{-1}(u) [d\mu^m(u) - d\mu(u)].$$

Letting  $M^m(u) \equiv \mu^m([0, u]) = \int_0^u G(\gamma^m(v)) dv$  and  $M(u) \equiv \mu([0, u]) = \int_0^u G(\gamma(v)) dv$  and applying Lebesgue-Stieltjes integration by parts (e.g., Hewitt, 1960, theorem A), we obtain

$$\begin{aligned} \int_0^1 F^{-1}(u) [d\mu^m(u) - d\mu(u)] &= \{\mu^m(\mathbb{I}) - \mu(\mathbb{I})\} F^{-1}(1) - \int_0^1 \{M^m(u) - M(u)\} dF^{-1}(u) \\ &= - \int_0^1 \{M^m(u) - M(u)\} dF^{-1}(u), \end{aligned}$$

as  $\mu^m(\mathbb{I}) - \mu(\mathbb{I}) = \mathbb{E}[G(Z^m)] - \mathbb{E}[G(Z)] = 0$  by the identical distribution of  $Z^m$  and  $Z$ .

We now show that if we condition on  $U \in [0, u]$ , for any  $u \in [0, 1]$ , then  $G(\gamma^m(U))$  is first-order stochastically dominated by  $G(\gamma(U))$ . (i) For any  $t \in [0, G(\gamma^m(u))]$ ,

$$\Pr(G(\gamma^m(U)) \leq t \mid U \leq u) = \frac{\Pr(G(\gamma^m(U)) \leq t)}{u} = \frac{\Pr(G(\gamma(U)) \leq t)}{u},$$

where the first equality holds by the monotonicity of  $G$  and  $\gamma^m$ , and the second holds as  $G(\gamma^m(U))$  and  $G(\gamma(U))$  have the same distribution. Now  $\Pr(G(\gamma(U)) \leq t) \geq \Pr(G(\gamma(U)) \leq t \cap U \leq u) = \Pr(G(\gamma(U)) \leq t \mid U \leq u) \Pr(U \leq u) = \Pr(G(\gamma(U)) \leq t \mid U \leq u) u$ , which implies

$$\Pr\{G(\gamma^m(U)) \leq t \mid U \leq u\} \geq \Pr\{G(\gamma(U)) \leq t \mid U \leq u\}.$$

(ii) For  $t \in (G(\gamma^m(u)), 1]$ ,

$$1 = \Pr\{G(\gamma^m(U)) \leq t \mid U \leq u\} \geq \Pr\{G(\gamma(U)) \leq t \mid U \leq u\}.$$

Because  $u$  is arbitrary, then conditional on  $U \leq u$  for any  $u \in [0, 1]$ ,  $G(\gamma^m(U))$  is first-order stochastically dominated by  $G(\gamma(U))$ .

Hence,

$$\mathbb{E}[G(\gamma(U)) \mid U \leq u] \geq \mathbb{E}[G(\gamma^m(U)) \mid U \leq u],$$

which implies that  $M(u) \geq M^m(u)$  for all  $u \in [0, 1]$ . As  $F^{-1}(\cdot)$  is a strictly increasing function, it follows that  $dF^{-1}(u) > 0$ , so that

$$\int_0^1 \{M^m(u) - M(u)\} dF^{-1}(u) \leq 0.$$

That is,  $\mathbb{E}([G(Z^m) - G(Z)] X_i) \geq 0$ , and the proof is complete. ■

## A.2 Measurable selection for non-metric spaces

Results establishing the existence of a measurable selection when optimizing over a metric space are well known (see, e.g., Aliprantis and Border (2006, theorem 18.19), and the references cited there). Here, (2) is optimized over  $\mathcal{M}$ , which is not metrizable, so those results do not apply. Lemma 7 of van Zandt and Vives (2007) or the Monotonicity Theorem of Milgrom and Shannon (1994) ensure that the solution to (2) is a non-empty compact-valued upper hemi-continuous (uhc) correspondence. A useful measurable selection result for such correspondences is that of Leese (1978, theorem 4.2). As Leese notes, this result makes use of the axiom of choice. We state a convenient direct corollary of Leese's result, which immediately delivers the measurable selection  $r_i$ . For simplicity, we use Leese's notation. Recall that  $\aleph_1$  denotes the first uncountable cardinal.

**Proposition A.2 (Measurable Selection)** *Let  $S$  be a space on which is defined a  $\sigma$ -algebra  $\mathfrak{M}$  of subsets and  $X$  any topological space that is either (a) first countable, with  $X$  having cardinality  $\aleph_1$ ; or (b) second countable. Let  $\Gamma$  be a uhc correspondence from  $S$  into the space of non-empty compact subsets of  $X$ . Then  $\Gamma$  has a selector  $\gamma$  such that  $\gamma^{-1}(F) \in \mathfrak{M}$  for every closed set  $F$  in  $X$ . That is,  $\gamma$  is a measurable selector.*

**Proof:** We verify that  $X$  satisfies Leese's Condition (B) and that  $\Gamma$  satisfies Leese's regularity condition on  $\Gamma^-$ , defined below. The result then follows.

Condition (B) requires that  $X$  has a family of closed sets  $\{B_\alpha\}$ , whose cardinality is at most  $\aleph_1$ , which generates the Borel  $\sigma$ -algebra on  $X$ . We verify this only under assumption (a); the result is immediate given (b). Since  $X$  satisfies the first axiom of countability, the neighborhood system of each point  $x$  in  $X$  has a countable base, say  $\{\mathcal{U}_{x,n}\}$ . Without loss of generality, we can let  $\{\mathcal{U}_{x,n}\}$  be a decreasing nested family of open neighborhoods of  $x$ . Let  $\mathbf{B} \equiv \{\{\mathcal{U}_{x,n}\}, x \in X\}$ . We verify that  $\mathbf{B}$  is a base for  $\tau$ , the topology of  $X$ . Given  $x$ , let  $\mathcal{U}$  be a neighborhood of  $x$ . Since  $\mathcal{U}$  is a neighborhood, it contains an open set  $\mathcal{V}$  that is also a neighborhood of  $x$ . By the properties of our neighborhood system, we can take  $n_x$  sufficiently large that  $\mathcal{U}_{x,n_x} \subseteq \mathcal{V}$ . Since  $x$  and  $\mathcal{U}$  are arbitrary,  $\mathbf{B}$  is a base for  $\tau$ . Consequently, each member of  $\tau$  is the union of members of  $\mathbf{B}$ . The members of  $\tau$  generate the Borel  $\sigma$ -algebra on  $X$ , and so therefore does  $\mathbf{B}$ . Let  $\{B_\alpha\} \equiv \{\{\mathcal{U}_{x,n}^c\}, x \in X\}$ , where  $\mathcal{U}_{x,n}^c$  is the (closed) complement of  $\mathcal{U}_{x,n}$ . Thus,  $\{B_\alpha\}$  also generates the Borel  $\sigma$ -algebra on  $X$ . By assumption,  $X$  has cardinality  $\aleph_1$ . The cardinality of  $\{B_\alpha\}$  is then  $\aleph_0 \times \aleph_1 = \max(\aleph_0, \aleph_1) = \aleph_1$  under the axiom of choice. Condition (B) therefore holds.

Next, for any subset  $F$  of  $X$ , let  $\Gamma^-(F) \equiv \{t \in S : \Gamma(t) \cap F \neq \emptyset\}$ . We must show that  $\Gamma^-(F)$  belongs to  $\mathfrak{M}$  for all closed  $F$  in  $\mathcal{X}$ . Let  $F$  be any closed set in  $X$ ; we will show that  $\Gamma^-(F)$  is also a closed set and thus a Borel set. Pick an arbitrary convergent sequence  $f_n \in \Gamma^-(F)$ ,  $n = 1, 2, \dots$ , such that  $f_n \rightarrow f$  with  $f \in \mathcal{S}$ . This implies that  $\Gamma(f_n) \cap F \neq \emptyset$  for all  $n$ . It then suffices to show

that  $\Gamma(f) \cap F \neq \emptyset$ , so that  $\Gamma^-(F)$  is closed. We prove this by contradiction.

Thus, suppose that  $\Gamma(f) \cap F = \emptyset$ . Because  $\Gamma(\cdot)$  is uhc, for any open neighborhood  $\mathcal{N}$  of  $\Gamma(f)$ , there exists a neighborhood  $\mathcal{U}$  of  $f$  such that  $\Gamma(g)$  is a subset of  $\mathcal{N}$  for all  $g \in \mathcal{U}$ . In other words, for  $n$  sufficiently large,  $\Gamma(f_n)$  is a subset of  $\mathcal{N}$ . Hence,  $F \cap \mathcal{N}$  is non-empty, since  $F$  contains a subset in  $\Gamma(f_n)$ . Hence, for any open neighborhood  $\mathcal{N}$  of  $\Gamma(f)$ ,  $F \cap \mathcal{N} \neq \emptyset$ . Let  $\tilde{\mathcal{N}} = \mathcal{N} \cap F^c$ . Since  $\Gamma(f) \cap F = \emptyset$ , we have  $\Gamma(f) \subseteq F^c$ , so that  $\Gamma(f) \subseteq \tilde{\mathcal{N}}$ . Because  $\tilde{\mathcal{N}}$  is an open set, it is still an open neighborhood of  $\Gamma(f)$ . But  $F \cap \tilde{\mathcal{N}} = \emptyset$  by the definition of  $\tilde{\mathcal{N}}$ , a contradiction.

This verifies the conditions of Leese (1978, theorem 4.2), so the result follows. ■

This result immediately implies that the measurable selection  $\mathbf{r}_i$  exists, given the continuum hypothesis, that is, that the cardinality of the continuum,  $\mathfrak{c}$ , is  $\aleph_1$ . We apply the result above with  $S = \times_{j \neq i} \mathcal{M}$  and  $X = \mathcal{M}$ . As noted earlier,  $\mathcal{M}$  is first countable; it is also readily verified that it has cardinality  $\mathfrak{c}$ , which is  $\aleph_1$  under the continuum hypothesis.  $\Gamma$  is the solution to (2), which is non-empty, compact-valued, and uhc.  $\mathbf{r}_i$  then corresponds to  $\gamma$ , restricted to  $\times_{j \neq i} \mathcal{M}_s$ .

A convenient complementary result ensuring the existence of a non-empty, compact-valued, uhc solution to a general optimization problem is theorem 2 of Ausubel and Deneckere (1993). This result also ensures the existence of a solution to (2).

As the above proposition suggests, there may be cases where a measurable selection does not exist. But such cases may still be brought within topological settable systems by viewing correspondences taking set values in a topological space  $\mathfrak{X}$ , say, as (single-valued) mappings taking values in the power set,  $\mathcal{P}(\mathfrak{X})$ , the set of all subsets of  $\mathfrak{X}$ . This requires specifying a suitable topology for  $\mathcal{P}(\mathfrak{X})$ , known as a hypertopology. We leave aside the details here.

### A.3 $\mathcal{M}$ is a bounded Banach class

**Proposition A.3 (Bounded Banach Class)** *Let  $\mathcal{M}$  be the class of monotone functions from  $\mathbb{I}$  to  $\mathbb{I}$ . Then  $\mathcal{M}$  is a bounded Banach class.*

**Proof:** Following Aumann (1960), it suffices to show that there exist two classes of subsets of  $\mathbb{I}$ ,  $\mathcal{B}$  and  $\mathcal{U}$ , which are countable and generate  $\mathbb{I}$ , such that  $\mathcal{M} \subseteq L_\alpha(\mathcal{B}, \mathcal{U})$  for some  $\alpha \in \mathbb{N}$ , where  $L_\alpha(\mathcal{B}, \mathcal{U})$  is the Banach class of order  $\alpha$  for  $(\mathcal{U}, \mathcal{B})$ , defined as the set of all functions  $f : \mathbb{I} \rightarrow \mathbb{I}$ , such that for all  $B \in Q_1(\mathcal{B})$ ,  $f^{-1}(B) \in Q_{\alpha+1}(\mathcal{U})$ . For  $\alpha \in \mathbb{N}$  and a class of subsets  $\mathcal{A}$ ,  $Q_\alpha(\mathcal{A})$  is defined recursively in Aumann(1960):  $Q_1(\mathcal{A})$  consists of all countable unions of members of  $\mathcal{A}$ ;  $P_1(\mathcal{A})$  consists of all complements of members of  $Q_1(\mathcal{A})$ ; for  $\alpha \geq 2$ ,  $Q_\alpha(\mathcal{A}) = Q_1(\cup_{\beta < \alpha} P_\beta(\mathcal{A}))$  and  $P_\alpha(\mathcal{A}) = P_1(\cup_{\beta < \alpha} P_\beta(\mathcal{A}))$ .

Here, let  $\mathcal{B} = \mathcal{U} = \{[0, r) : r \text{ is a rational number in } \mathbb{I}\}$ . By definition,  $Q_1(\mathcal{B}) = \{[0, a) : a \in \mathbb{I}\}$ , and  $P_1(\mathcal{B})$  consists of all complements of members of  $Q_1(\mathcal{B})$ :  $P_1(\mathcal{B}) = \{[a, 1] : a \in \mathbb{I}\}$ .

Moreover,  $Q_2(\mathcal{B}) = Q_1(P_1(\mathcal{B})) = \{(a, 1] : a \in \mathbb{R}\} \cup \{[a, 1] : a \in \mathbb{R}\}$ , and  $P_2(\mathcal{B}) = P_1(P_1(\mathcal{B})) = \{[0, a] : a \in \mathbb{R}\} \cup \{[0, a] : a \in \mathbb{R}\}$ .

Because  $Q_3(\mathcal{B}) = Q_1(P_1(\mathcal{B}) \cup P_2(\mathcal{B}))$ , we have  $P_2(\mathcal{B}) \subseteq Q_3(\mathcal{B})$ . Thus, for any monotone function  $f$ ,  $f^{-1}(B) \in P_2(\mathcal{B}) \subseteq Q_3(\mathcal{B})$  for any  $B \in Q_1(\mathcal{B})$ . Therefore  $f \in L_2(\mathcal{B}, \mathcal{B})$ . Since  $f$  is an arbitrary element of  $\mathcal{M}$ ,  $\mathcal{M} \subseteq L_2(\mathcal{B}, \mathcal{B})$ . It follows that  $\mathcal{M}$  is a bounded Banach class. ■

#### A.4 Mutual consistency and cross-partition effect linkage

We illustrate the role of compatibility in linking fine partition and coarse partition effects. This also elucidates the relation between settable systems and classical simultaneous equation systems. (See also Heckman, 2005, section 2.5.)

This illustration is easiest in a context simpler than our auction, so here we consider a different game, Bertrand duopoly. For maximum transparency, we first consider a linear system. We then consider a more generic case. Thus, suppose that the two firms have price reaction functions

$$p_1^{e,*} = b_1 p_2 + c'_1 s_1 \quad p_2^{e,*} = b_2 p_1 + c'_2 s_2,$$

where  $p_1^{e,*}$  and  $p_2^{e,*}$  are the (scalar) prices charged by firms 1 and 2 for their differentiated goods, respectively, given their rival's arbitrary prices  $p_2$  and  $p_1$  and vectors of possibly firm-specific cost and demand shifters,  $s_1$  and  $s_2$ . Viewed as a settable system, these are the elementary partition structural equations, with responses  $(p_1^{e,*}, p_2^{e,*})$ , price settings  $(p_1, p_2)$ , and fundamental settings  $(s_1, s_2)$ . The coefficients  $(b_1, c'_1)$  and  $(b_2, c'_2)$  embody elementary partition effects and are (functions of) attributes. For simplicity, we suppress the aggregate demand functions for price-taking consumers.

Next, suppose that equilibrium prices in this system are determined as

$$p_1^{g,*} = \pi'_1 s \quad p_2^{g,*} = \pi'_2 s,$$

where  $s \equiv (s'_1, s'_2)'$ . The  $\pi$ 's are coefficients embodying equilibrium price effects of the cost and demand shifters  $s$ . Viewed as a settable system, these are equilibrium structural equations, corresponding to the global partition, with equilibrium responses  $(p_1^{g,*}, p_2^{g,*})$  depending only on fundamental settings  $s$ . Like  $(b_1, c'_1)$  and  $(b_2, c'_2)$ , the  $\pi$ 's are (functions of) attributes.

Without further structure, the settable systems elementary and global partitions for the Bertrand duopoly game are comparable but need not be compatible. Even without compatibility, causal discourse is well defined for both the elementary (fine) and equilibrium (coarse) partitions. There is, however, no necessary relation between elementary partition effects and equilibrium effects. Observe that simultaneity is not present in these structural systems, although mutual causality is present in the elementary partition.

Compatibility between fine and coarse partitions requires

$$p_1^{g,*} = b_1 p_2^{g,*} + c'_1 s_1 \quad p_2^{g,*} = b_2 p_1^{g,*} + c'_2 s_2,$$

where equilibrium responses appear in the elementary partition response functions. Here, this enforces Nash equilibrium. When  $b_1 b_2 \neq 1$ , this simultaneous system has a unique solution:

$$\begin{aligned} p_1^{g,*} &= (1 - b_1 b_2)^{-1} [c'_1 s_1 + b_1 c'_2 s_2] \\ p_2^{g,*} &= (1 - b_1 b_2)^{-1} [b_2 c'_1 s_1 + c'_2 s_2]. \end{aligned}$$

Nash equilibrium thus ensures that the effects of  $s$  on equilibrium prices  $(p_1^{g,*}, p_2^{g,*})$  are

$$\begin{aligned} \pi_1 &= (\pi'_{11}, \pi'_{12})' = (1 - b_1 b_2)^{-1} [c'_1, b_1 c'_2]' \\ \pi_2 &= (\pi'_{21}, \pi'_{22})' = (1 - b_1 b_2)^{-1} [b_2 c'_1, c'_2]'. \end{aligned}$$

Compatibility thus implies functional dependence between equilibrium and elementary partition effects.

This functional dependence is a main feature of compatibility: it allows inference about fine partition effects based on coarse partition effects, whenever specific elements of  $(b', c')$  =  $((b_1, b_2), (c'_1, c'_2))$  can be recovered from  $\pi = (\pi'_1, \pi'_2)'$ . That is, one can make valid statements about cause and effect in unobservable fine partition structures using knowledge gained solely from observable coarse partition (e.g., equilibrium) structures. Recovering elements of  $(b', c')$  from  $\pi = (\pi'_1, \pi'_2)'$  is, of course, the classical identification problem in systems of structural equations (Fisher, 1966). This example shows, however, that in settable systems, instantaneous or simultaneous causality has no role to play. It is necessary and sufficient that each agent has sufficient information to compute their equilibrium response.<sup>29</sup>

The implications of compatibility for linking effects between partitions extend beyond linear structures. Suppose firms 1 and 2 have  $\ell_1$  and  $\ell_2$  products, respectively. Let  $s = (s'_0, s'_1, s'_2)'$ , where  $s_j$  has dimension  $k_j$ , and  $s_1$  and  $s_2$  do not have common elements. Consider the generic reaction functions for the firms,

$$p_1^{e,*} = r_1^e(p_2, s_0, s_1; \mathbf{a}) \quad p_2^{e,*} = r_2^e(p_1, s_0, s_2; \mathbf{a}),$$

together with the generic equilibrium responses,

$$p_1^{g,*} = r_1^g(s; \mathbf{a}) \quad p_2^{g,*} = r_2^g(s; \mathbf{a}).$$

---

<sup>29</sup>These informational requirements have important implications for modeling. With less information available to players, dynamic learning structures (e.g., Chen and White (1998)) become salient.

Suppose also that mutual consistency conditions hold:

$$p_1^{g,*} = r_1^e(p_2^{g,*}, s_0, s_1; \mathbf{a}) \quad p_2^{g,*} = r_2^e(p_1^{g,*}, s_0, s_2; \mathbf{a}).$$

To illustrate the implied effect linkages for firm 1, we substitute to obtain

$$r_1^g(s; \mathbf{a}) = r_1^e(r_2^g(s; \mathbf{a}), s_0, s_1; \mathbf{a}).$$

Assuming differentiability and differentiating with respect to  $s_2$  gives

$$\nabla_{s_2} r_1^g(s; \mathbf{a}) = \nabla_{s_2} r_2^g(s; \mathbf{a}) \nabla_{p_2} r_1^e(r_2^g(s; \mathbf{a}), s_0, s_1; \mathbf{a}),$$

$k_2 \times \ell_1 \qquad k_2 \times \ell_2 \qquad \ell_2 \times \ell_1$

where  $\nabla_{s_2}$  and  $\nabla_{p_2}$  are the gradient operators with respect to  $s_2$  and  $p_2$ , respectively. Now pre-multiply both sides of this equation by  $\nabla_{s_2} r_2^g(s; \mathbf{a})'$  to get

$$\nabla_{s_2} r_2^g(s; \mathbf{a})' \nabla_{s_2} r_1^g(s; \mathbf{a}) = [\nabla_{s_2} r_2^g(s; \mathbf{a})' \nabla_{s_2} r_2^g(s; \mathbf{a})] \nabla_{p_2} r_1^e(r_2^g(s; \mathbf{a}), s_0, s_1; \mathbf{a}).$$

This equation can be solved for  $\nabla_{p_2} r_1^e(r_2^g(s; \mathbf{a}), s_0, s_1; \mathbf{a})$ , provided  $\nabla_{s_2} r_2^g(s; \mathbf{a})' \nabla_{s_2} r_2^g(s; \mathbf{a})$  has full rank (the *generalized rank condition at s*), for which a necessary condition is that  $k_2 \geq \ell_2$ , the order condition. With full rank,

$$\nabla_{p_2} r_1^e(p_2^g, s_0, s_1; \mathbf{a}) = [\nabla_{s_2} r_2^g(s; \mathbf{a})' \nabla_{s_2} r_2^g(s; \mathbf{a})]^{-1} \nabla_{s_2} r_2^g(s; \mathbf{a})' \nabla_{s_1} r_1^g(s; \mathbf{a}).$$

These are the reaction function marginal effects of firm 2's prices on firm 1's prices at  $p_2^{g,*} = r_2^g(s; \mathbf{a})$  and  $(s_0, s_1)$ . These are expressed solely in terms of equilibrium effects, due to compatibility. White and Chalak (2011) provide further discussion of identification and estimation of causal effects using "derivative ratio" measures of this sort.

## References

- Aliprantis, C. and K. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. 3rd ed. Berlin: Springer Verlag.
- Anderson R. and R. Bing (1968), "A Complete Elementary Proof that Hilbert Space is Homeomorphic to the Countable Infinite Product of Real Lines," *Bulletin (New Series) of the American Mathematical Society*, 74, 771-792.
- Angus, J. (1994), "The Probability Integral Transform and Related Results," *SIAM Review*, 36, 652-654.
- Aumann, R. (1960), "Spaces of Measurable Transformations," *Bulletin (New Series) of the American Mathematical Society*, 66, 301-304.

- Ausubel, L. and R. Deneckere (1993), "A Generalized Theorem of the Maximum," *Economic Theory*, 3, 99-107.
- Bartle, R. (1966). *The Elements of Integration*. New York: Wiley.
- Burris, S. and H.P. Sankappanavar (1981). *A Course in Universal Algebra*. Berlin: Springer Verlag.
- Chalakov, K. and H. White (2012), "Causality, Conditional Independence, and Graphical Separation in Settable Systems," *Neural Computation*, 24, 1611–1668.
- Chen, X. and H. White (1998), "Nonparametric Adaptive Learning with Feedback," *Journal of Economic Theory*, 82, 190–222.
- Corbae, D., M. Stinchcombe, and J. Zeman (2009). *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton: Princeton University Press.
- Guerre, E., I. Perrigne, and Q. Vuong (2000), "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525-574.
- Fisher, F. (1966). *The Identification Problem in Econometrics*. New York: McGraw-Hill.
- Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–438.
- Halpern, J. (2000), "Axiomatizing Causal Reasoning," *Journal of Artificial Intelligence Research*, 12, 317–37.
- Halpern, J. and J. Pearl (2005a), "Causes and Explanations: A Structural-Model Approach. Part I: Causes," *British Journal of Philosophy of Science*, 843-887.
- Halpern, J. and J. Pearl (2005b), "Causes and Explanations: A Structural-Model Approach. Part II: Explanations," *British Journal of Philosophy of Science*, 889-911.
- Harsányi, J. (1967), "Games with Incomplete Information Played by 'Bayesian' Players, I-III," *Management Science*, 14, 159-182; 320–334; 486–502.
- Haavelmo, T. (1943), "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, 11, 1–12.
- Haavelmo, T. (1944), "The Probability Approach in Econometrics," *Supplement to Econometrica*, 12, 1-115.
- Heckman, J. (2005), "The Scientific Model of Causality," *Sociological Methodology*, 35, 1–97.
- Heckman, J. (2010), "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, 48, 356–98.
- Heckman, J. and E. Vytlacil (2005), "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73, 669-738.
- Hewitt, E. (1960), "Integration by Parts for Stieltjes Integrals," *The American Mathematical Monthly*, 67, 419-423.



- Hicks, J.R. (1939). *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford: Clarendon Press.
- Hempel, C. and P. Oppenheim (1948), “Studies in the Logic of Explanation,” *Philosophy of Science*, 15, 135-75.
- Holland, P.W. (1986), “Statistics and Causal Inference” (with Discussion), *Journal of the American Statistical Association*, 81, 945-970.
- Hoover, K. (2004), “Lost Causes,” *Journal of the History of Economic Thought*, 26, 149-164.
- Imbens, G. and J. Wooldridge (2009), “Recent Developments in the Econometrics of Program Evaluation,” *Journal of Economic Literature*, 47, 5-86.
- Jiang, A.X. and K. Leyton-Brown (2010), “Bayesian Action-Graph Games,” in J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta (eds.), *Advances in Neural Information Processing Systems* 23.
- Kelley, J. (1985). *General Topology*. Berlin: Springer Verlag.
- Kohlberg, E. and J.F. Mertens (1986), “On the Strategic Stability of Equilibria,” *Econometrica*, 54, 1003-1038.
- Koller, D. and R. Milch (2001), “Multi-Agent Influence Diagrams for Representing and Solving Games,” in *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01)*, pp. 1027–1034.
- Koller, D. and R. Milch (2003), “Multi-Agent Influence Diagrams for Representing and Solving Games,” *Games and Economic Behavior*, 45, 181-221.
- Kreps, D. and R. Wilson (1982), “Reputation and Imperfect Information,” *Journal of Economic Theory*, 27, 253-279.
- Krishna, V. (2010). *Auction Theory*. New York: Academic Press.
- Leese, S. (1978). “Measurable Selections and the Uniformization of Souslin Sets,” *American Journal of Mathematics*, 100, 19-41.
- Marschak, J. (1950), “Statistical Inference in Economics,” in T. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*, Cowles Commission for Research in Economics, Monograph 10. New York: Wiley, pp. 1–50.
- Marshall, A. (1890). *Principles of Economics*. London: Macmillan.
- Milgrom, P. and C. Shannon, C. (1994), “Monotone Comparative Statics,” *Econometrica*, 62, 157-180.
- Mill, J.S. (1848). *The Principles of Political Economy with Some of their Applications to Social Philosophy*. London: Longmans, Green and Co.
- Myerson, R. (1978), “Refinements of the Nash Equilibrium Concept,” *International Journal of Game Theory*, 7, 73-80.

- Nelsen, R.B. (1999). *An Introduction to Copulas*. New York: Springer Verlag.
- Pearl (2000). *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press.
- Penalva-Zuasti, J. and M. Ryall (2003), “Causal Assessment in Finite-Length Extensive-Form Games,” Universitat Pompeu Fabra Department of Economics Working Paper.
- Popper, K. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.
- Reny, P. (2011), “On the Existence of Monotone Pure-Strategy Equilibria in Bayesian Games,” *Econometrica*, 79, 499-553.
- Riley, J. G., and W. F. Samuelson (1981), “Optimal Auctions,” *American Economic Review*, 71, 381-392.
- Roebroek, A., A. Seth, and P. Valdés-Sosa (2011), “Causal Time Series Analysis of functional Magnetic Resonance Imaging Data,” *Journal of Machine Learning Research: Workshop and Conference Proceedings*, 12, 65-94.
- Samuelson, P. (1947). *Foundations of Economics*. Cambridge, MA: Harvard University Press.
- Selten, R. (1974), “A Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games,” *International Journal of Game Theory*, 4, 25-55.
- Simon, H. (1953), “Causal Ordering and Identifiability,” in *Studies in Econometric Method*, Cowles Commission for Research in Economics, Monograph 14. New York: Wiley, pp.49-74.
- Sklar, A. (1959), “Fonctions de Répartition à  $n$  Dimensions et leurs Marges,” *Publ. Inst. Statist. Univ. Paris* 8, 229–231.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan & T. Cadell.
- Strotz, R. and H. Wold (1960), “Recursive vs. Nonrecursive Systems: An Attempt at Synthesis,” *Econometrica*, 28, 417–427.
- Valdés-Sosa, P., J. Bornot-Sánchez, M. Vega-Hernández, L. Melie-García, A. Lage-Castellanos, and E. Canales-Rodríguez (2006), “Granger Causality on Spatial Manifolds: Applications to Neuroimaging.” B. Schelter, M. Winterhalder, and J. Timme (Eds.). *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*. Wiley-VCH Verlag GmbH & Co. KGaA.
- Valdés-Sosa, P., A. Roebroek, J. Daunizeau, and K. Friston (2011), “Effective Connectivity: Influence, Causality, and Biophysical Modeling,” *NeuroImage*, 58, 339–361.
- van Zandt, T., and X. Vives (2007), “Monotone Equilibria in Bayesian Games of Strategic Complementarities,” *Journal of Economic Theory*, 134, 339-360.
- Wellman, M., L. Hong, and S. Page (2011), “The Structure of Signals: Causal Interdependence Models for Games of Incomplete Information,” University of Michigan Department of Computer

Science & Engineering Working Paper.

White, H., M. Al-Sadoon, and K. Chalak (2011), “Rational Expectations and Causality: A Settable Systems View,” UCSD Department of Economics Discussion Paper.

White, H. and K. Chalak (2009), “Settable Systems: An Extension of Pearl’s Causal Model with Optimization, Equilibrium, and Learning,” *Journal of Machine Learning Research*, 10, 1759-1799.

White, H. and K. Chalak (2011), “Identification and Identification Failure for Treatment Effects using Structural Systems,” *Econometric Reviews*, forthcoming.

White, H., K. Chalak, and X. Lu (2011), “Linking Granger Causality and the Pearl Causal Model with Settable Systems,” *Journal of Machine Learning Research: Workshop and Conference Proceedings*, 12, 1 - 29.