Linking Granger Causality and the Pearl Causal Model with Settable Systems

Halbert White HWHITE@UCSD.EDU
Department of Economics
University of California, San Diego
La Jolla, CA 92093

Karim Chalak CHALAK@BC.EDU
Department of Economics
Boston College
140 Commonwealth Avenue
Chestnut Hill, MA 02467

Xun Lu XUNLU@UST.HK
Department of Economics
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

Editor: Florin Popescu and Isabelle Guyon

Abstract

The causal notions embodied in the concept of Granger causality have been argued to belong to a different category than those of Judea Pearl's Causal Model, and so far their relation has remained obscure. Here, we demonstrate that these concepts are in fact closely linked by showing how each relates to straightforward notions of direct causality embodied in settable systems, an extension and refinement of the Pearl Causal Model designed to accommodate optimization, equilibrium, and learning. We then provide straightforward practical methods to test for direct causality using tests for Granger causality.

Keywords: Causal Models, Conditional Exogeneity, Conditional Independence, Granger Non-causality

1. Introduction

The causal notions embodied in the concept of Granger causality ("$G-$causality") (e.g., Granger, 1969; Granger and Newbold, 1986) are probabilistic, relating to the ability of one time series to predict another, conditional on a given information set. On the other hand, the causal notions of the Pearl Causal Model ("PCM") (e.g., Pearl, 2000) involve specific notions of interventions and of functional rather than probabilistic dependence. The relation between these causal concepts has so far remained obscure. For his part, Granger (1969) acknowledged that $G-$causality was not "true" causality, whatever that might be, but that it seemed likely to be an important part of the full story. On the other hand, Pearl (2000, p.39) states that "econometric concepts such as 'Granger causality' (Granger 1969) and 'strong exogeneity' (Engle et al. 1983) will be classified as statistical rather than causal." In

practice, especially in economics, numerous studies have used $G-$causality either explicitly or implicitly to draw structural or policy conclusions, but without any firm foundation.

Recently, White and Lu (2010a, "WL") have provided conditions under which $G-$causality is equivalent to a form of direct causality arising naturally in dynamic structural systems, defined in the context of *settable systems*. The settable systems framework, introduced by White and Chalak (2009, "WC"), extends and refines the PCM to accommodate optimization, equilibrium, and learning. In this paper, we explore the relations between direct structural causality in the settable systems framework and notions of direct causality in the PCM for both recursive and non-recursive systems. The close correspondence between these concepts in the recursive systems relevant to $G-$causality then enables us to show that there is in fact a close linkage between $G-$causality and PCM notions of direct causality. This enables us to provide straightforward practical methods to test for direct causality using tests for Granger causality.

In a related paper, Eichler and Didelez (2009) also study the relation between $G-$causality and interventional notions of causality. They give conditions under which Granger non-causality implies that an intervention has no effect. In particular, Eichler and Didelez (2009) use graphical representations as in Eichler (2007) of given $G-$causality relations satisfying the "global Granger causal Markov property" to provide graphical conditions for the identification of effects of interventions in "stable" systems. Here, we pursue a different route for studying the interrelations between $G-$causality and interventional notions of causality. Specifically, we see that $G-$causality and certain settable systems notions of direct causality based on functional dependence are *equivalent* under a conditional form of exogeneity. Our conditions are alternative to "stability" and the "global Granger causal Markov property," although particular aspects of our conditions have a similar flavor.

As a referee notes, the present work also provides a rigorous complement, in discrete time, to work by other authors in this volume (for example Roebroeck, Seth, and Valdés-Sosa, 2011) on combining structural and dynamic concepts of causality.

The plan of the paper is as follows. In Section 2, we briefly review the PCM. In Section 3, we motivate settable systems by discussing certain limitations of the PCM using a series of examples involving optimization, equilibrium, and learning. We then specify a formal version of settable systems that readily accommodates the challenges to causal discourse presented by the examples of Section 3. In Section 4, we define direct structural causality for settable systems and relate this to corresponding notions in the PCM. The close correspondence between these concepts in recursive systems establishes the first step in linking $G-$causality and the PCM. In Section 5, we discuss how the results of WL complete the chain by linking direct structural causality and $G-$causality. This also involves a conditional form of exogeneity. Section 6 constructs convenient practical tests for structural causality based on proposals of WL, using tests for $G-$causality and conditional exogeneity. Section 7 contains a summary and concluding remarks.

2. Pearl's Causal Model

Pearl's definition of a *causal model* (Pearl, 2000, def. 7.1.1, p. 203) provides a formal statement of elements supporting causal reasoning. The PCM is a triple $M := (u, v, f)$, where $u := \{u_1, ..., u_m\}$ contains "background" variables determined outside the model,

$v := \{v_1, ..., v_n\}$ contains "endogenous" variables determined within the model, and $f := \{f_1, ..., f_n\}$ contains "structural" functions specifying how each endogenous variable is determined by the other variables of the model, so that $v_i = f_i(v_{(i)}, u),\ \ i = 1, ..., n$. Here, $v_{(i)}$ is the vector containing every element of v but v_i. The integers m and n are finite. The elements of u and v are system "units."

Finally, the PCM requires that for each u, f yields a unique fixed point. Thus, there must be a unique collection $g := \{g_1, ..., g_n\}$ such that for each u,

$$v_i = g_i(u) = f_i(g_{(i)}(u), u), \quad i = 1, ..., n. \tag{1}$$

The unique fixed point requirement is crucial to the PCM, as this is necessary for defining the *potential response function* (Pearl, 2000, def. 7.1.4). This provides the foundation for discourse about causal relations between endogenous variables; without the potential response function, causal discourse is not possible in the PCM. A variant of the PCM (Halpern, 2000) does not require a fixed point, but if any exist, there may be multiple collections of functions g yielding a fixed point. We call this a Generalized Pearl Causal Model (GPCM). As GPCMs also do not possess an analog of the potential response function in the absence of a unique fixed point, causal discourse in the GPCM is similarly restricted.

In presenting the PCM, we have adapted Pearl's notation somewhat to facilitate subsequent discussion, but all essential elements are present and complete.

Pearl (2000) gives numerous examples for which the PCM is ideally suited for supporting causal discourse. As a simple game-theoretic example, consider a market in which there are exactly two firms producing similar but not identical products (e.g., Coke and Pepsi in the cola soft-drink market). Price determination in this market is a two-player game known as "Bertrand duopoly."

In deciding its price, each firm maximizes its profit, taking into account the prevailing cost and demand conditions it faces, as well as the price of its rival. A simple system representing price determination in this market is

$$\begin{aligned} p_1 &= a_1 + b_1\, p_2 \\ p_2 &= a_2 + b_2\, p_1. \end{aligned}$$

Here, p_1 and p_2 represent the prices chosen by firms 1 and 2 respectively, and a_1, b_1, a_2, and b_2 embody the prevailing cost and demand conditions.

We see that this maps directly to the PCM with $n = 2$, endogenous variables $v = (p_1, p_2)$, background variables $u = (a_1, b_1, a_2, b_2)$, and structural functions

$$\begin{aligned} f_1(v_2, u) &= a_1 + b_1\, p_2 \\ f_2(v_1, u) &= a_2 + b_2\, p_1. \end{aligned}$$

These functions are the Bertrand "best response" or "reaction" functions.

Further, provided $b_1 b_2 \neq 1$, this system has a unique fixed point,

$$\begin{aligned} p_1 &= g_1(u) = (a_1 + b_1 a_2)/(1 - b_1 b_2) \\ p_2 &= g_2(u) = (a_2 + b_2 a_1)/(1 - b_1 b_2). \end{aligned}$$

This fixed point represents the Nash equilibrium for this two-player game.

Clearly, the PCM applies perfectly, supporting causal discourse for this Bertrand duopoly game. Specifically, we see that p_1 causes p_2 and vice-versa, and that the effect of p_2 on p_1 is b_1, whereas that of p_1 on p_2 is b_2.

In fact, the PCM applies directly to a wide variety of games, provided that the game has a unique equilibrium. But there are many important cases where there may be no equilibrium or multiple equilibria. This limits the applicability of the PCM. We explore examples of this below, as well as other features of the PCM that limit its applicability.

3. Settable Systems

3.1 Why Settable Systems?

WC motivate the development of the settable system (SS) framework as an extension of the PCM that accommodates optimization, equilibrium, and learning, which are central features of the explanatory structures of interest in economics. But these features are of interest more broadly, especially in machine learning, as optimization corresponds to any intelligent or rational behavior, whether artificial or natural; equilibrium (e.g., Nash equilibrium) or transitions toward equilibrium characterize stable interactions between multiple interacting systems; and learning corresponds to adaptation and evolution within and between interacting systems. Given the prevalence of these features in natural and artificial systems, it is clearly desirable to provide means for explicit and rigorous causal discourse relating to systems with these features.

To see why an extension of the PCM is needed to handle optimization, equilibrium, and learning, we consider a series of examples that highlight certain limiting features of the PCM: (*i*) in the absence of a unique fixed point, causal discourse is undefined; (*ii*) background variables play no causal role; (*iii*) the role of attributes is restricted; and (*iv*) only a finite rather than a countable number of units is permitted. WC discuss further relevant aspects of the PCM, but these suffice for present purposes.

Example 3.1 (Equilibria in Game Theory) Our first example concerns general two-player games, extending the discussion that we began above in considering Bertrand duopoly.

Let two players $i = 1, 2$ have strategy sets S_i and utility functions u_i, such that $\pi_i = u_i(z_1, z_2)$ gives player i's payoff when player 1 plays $z_1 \in S_1$ and player 2 plays $z_2 \in S_2$. Each player solves the optimization problem

$$\max_{z_i \in S_i} \; u_i(z_1, z_2).$$

The solution to this problem, when it exists, is player i's *best response*, denoted

$$y_i = r_i^e(z_{(i)}; \mathbf{a}),$$

where r_i^e is player i's *best response function* (the superscript "e" stands for "elementary," conforming to notation formally introduced below); $z_{(i)}$ denotes the strategy played by the player other than i; and $\mathbf{a} := (S_1, u_1, S_2, u_2)$ denotes given attributes defining the game. For simplicity here, we focus on "pure strategy" games; see Gibbons (1992) for an accessible introduction to game theory.

Different configurations for **a** correspond to different games. For example, one of the most widely known games is *prisoner's dilemma*, where two suspects in a crime are separated and offered a deal: if one confesses and the other does not, the confessor is released and the other goes to jail. If both confess, both receive a mild punishment. If neither confesses, both are released. The strategies are whether to confess or not. Each player's utility is determined by both players' strategies and the punishment structure.

Another well known game is *hide and seek*. Here, player 1 wins by matching player 2's strategy and player 2 wins by mismatching player 1's strategy. A familiar example is a penalty kick in soccer: the goalie wins by matching the direction (right or left) of the kicker's kick; the kicker wins by mismatching the direction of the goalie's lunge. The same structure applies to baseball (hitter vs. pitcher) or troop deployment in battle (aggressor vs. defender).

A third famous game is *battle of the sexes*. In this game, Ralph and Alice are trying to decide how to spend their weekly night out. Alice prefers the opera, and Ralph prefers boxing; but both would rather be together than apart.

Now consider whether the PCM permits causal discourse in these games, e.g., about the effect of one player's action on that of the other. We begin by mapping the elements of the game to the elements of the PCM. First, we see that **a** corresponds to PCM background variables u, as these are specified outside the system. The variables determined within the system, i.e., the PCM endogenous variables are $z := (z_1, z_2)$ corresponding to v, provided that (for now) we drop the distinction between y_i and z_i. Finally, we see that the best response functions r_i^e correspond to the PCM structural functions f_i.

To determine whether the PCM permits causal discourse in these games, we can check whether there is a unique fixed point for the best responses. In prisoner's dilemma, there is indeed a unique fixed point (both confess), provided the punishments are suitably chosen. The PCM therefore applies to this game to support causal discourse. But there is no fixed point for hide and seek, so the PCM cannot support causal discourse there. On the other hand, there are two fixed points for battle of the sexes: both Ralph and Alice choose opera or both choose boxing. The PCM does not support causal discourse there either. Nor does the GPCM apply to the latter games, because even though it does not require a unique fixed point, the potential response functions required for causal discourse are not defined.

The importance of game theory generally in describing the outcomes of interactions of goal-seeking agents and the fact that the unique fixed point requirement prohibits the PCM from supporting causal discourse in important cases strongly motivates formulating a causal framework that drops this requirement. As we discuss below, the SS framework does not require a unique fixed point, and it applies readily to games generally. Moreover, recognizing and enforcing the distinction between y_i (i's best response strategy) and z_i (an arbitrary setting of i's strategy) turns out to be an important component to eliminating this requirement.

Another noteworthy aspect of this example is that **a** is a *fixed* list of elements that define the game. Although elements of **a** may differ across players, they do not vary for a given player. This distinction should be kept in mind when referring to the elements of **a** as background "variables."

Example 3.2 (Optimization in Consumer Demand) The neoclassical theory of consumer demand posits that consumers determine their optimal goods consumption by maximizing utility subject to a budget constraint (see, e.g., Varian, 2009). Suppose for simplicity that there are just two goods, say beer and pizza. Then a typical consumer solves the problem

$$\max_{z_1,z_2} \ \mathcal{U}(z_1, z_2) \quad s.t. \quad m = z_1 + pz_2,$$

where z_1 and z_2 represent quantities consumed of goods 1 (beer) and 2 (pizza) respectively and $\mathcal{U}$ is the utility function that embodies the consumer's preferences for the two goods. For simplicity, let the price of a beer be \$1, and let p represent the price of pizza; m represents funds available for expenditure, "income" for short[1]. The budget constraint $m = z_1 + pz_2$ ensures that total expenditure on beer and pizza does not exceed income (no borrowing) and also that total expenditure on beer and pizza is not less than m. (As long as utility is increasing in consumption of the goods, it is never optimal to expend less than the funds available.)

Solving the consumer's demand problem leads to the optimal consumer demands for beer and pizza, y_1 and y_2. It is easy to show that these can be represented as

$$y_1 = r_1^a(p, m; \mathbf{a}) \qquad \text{and} \qquad y_2 = r_2^a(p, m; \mathbf{a}),$$

where r_1^a and r_2^a are known as the consumer's *market demand functions* for beer and pizza. The "a" superscript stands for "agent," corresponding to notation formally introduced below. The attributes $\mathbf{a}$ include the consumer's utility function $\mathcal{U}$ (preferences) and the admissible values for z_1, z_2, p, and m, e.g., $\mathbb{R}^+ := [0, \infty)$.

Now consider how this problem maps to the PCM. First, we see that $\mathbf{a}$ and (p, m) correspond to the background variables u, as these are not determined within the system. Next, we see that $y := (y_1, y_2)$ corresponds to PCM endogenous variables v. Finally, we see that the consumer demand functions r_i^a correspond to the PCM structural functions f_i. Also, because the demand for beer, y_1, does not enter the demand function for pizza, r_2^a, and vice versa, there is a unique fixed point for this system of equations. Thus, the PCM supports causal discourse in this system.

Nevertheless, this system is one where, in the PCM, the causal discourse natural to economists is unavailable. Specifically, economists find it natural to refer to "price effects" and "income effects" on demand, implicitly or explicitly viewing price p and income m as causal drivers of demand. For example, the pizza demand price effect is $(\partial/\partial p) r_2^a(p, m; \mathbf{a})$. This represents how much optimal pizza consumption (demand) will change as a result of a small (marginal) increase in the price of pizza. Similarly, the pizza demand income effect is $(\partial/\partial m) r_2^a(p, m; \mathbf{a})$, representing how much optimal pizza consumption will change as a result of a small increase in income. But in the PCM, causal discourse is reserved only for endogenous variables y_1 and y_2. The fact that background variables p and m do not have causal status prohibits speaking about their effects.

Observe that the "endogenous" status of y and "exogenous" status of p and m is determined in SS by utility maximization, the "governing principle" here. In contrast, there

1. Since a beer costs a dollar, it is the "numeraire," implying that income is measured in units of beer. This is a convenient convention ensuring that we only need to keep track of the price *ratio* between pizza and beer, p, rather than their two separate prices.

is no formal mechanism in the PCM that permits making these distinctions. Although causal discourse in the PCM can be rescued for such systems by "endogenizing" p and m, that is, by positing additional structure that explains the genesis of p and m in terms of further background variables, this is unduly cumbersome. It is much more natural simply to permit p and m to have causal status from the outset, so that price and income effects are immediately meaningful, without having to specify their determining processes. The SS framework embodies this direct approach. Those familiar with theories of price and income determination will appreciate the considerable complications avoided in this way. The same simplifications occur with respect to the primitive variables appearing in any responses determined by optimizing behavior.

Also noteworthy here is the important distinction between **a,** which represents fixed attributes of the system, and p and m, which are true variables that can each take a range of different possible values. As WC (p.1774) note, restricting the role of attributes by "lumping together" attributes and structurally exogenous variables as background objects without causal status creates difficulties for causal discourse in the PCM:

> [this] misses the opportunity to make an important distinction between invariant aspects of the system units on the one hand and counterfactual variation admissible for the system unit values on the other. Among other things, assigning attributes to u interferes with assigning natural causal roles to structurally exogenous variables.

By distinguishing between attributes and structurally exogenous variables, settable systems permit causal status for variables determined outside a given system, such as when price and income drive consumer demand.

Example 3.3 (Learning in Structural Vector Autoregressions) Structural vector autoregressions (VARs) are widely used to analyze time-series data. For example, consider the structural VAR

$$\begin{aligned} y_{1,t} &= a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + u_{1,t} \\ y_{2,t} &= a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + u_{2,t}, \qquad t = 1, 2, ..., \end{aligned}$$

where $y_{1,0}$ and $y_{2,0}$ are given scalars, $\mathbf{a} := (a_{11}, a_{12}, a_{21}, a_{22})'$ is a given real "coefficient" vector, and $\{u_t := (u_{1,t}, u_{2,t}) : t = 1, 2, ...\}$ is a given sequence. This system describes the evolution of $\{y_t := (y_{1,t}, y_{2,t}) : t = 1, 2, ...\}$ through time.

Now consider how this maps to the PCM. We see that $y_0 := (y_{1,0}, y_{2,0})$, $\{u_t\}$, and $\mathbf{a}$ correspond to the PCM background variables u, as these are not determined within the system. Further, we see that the sequence $\{y_t\}$ corresponds to the endogenous variables v, and that the PCM structural functions f_i correspond to

$$\begin{aligned} r_{1,t}(y^{t-1}, u^t; \mathbf{a}) &= a_{11}y_{1,t-1} + a_{12}y_{2,t-1} + u_{1,t} \\ r_{2,t}(y^{t-1}, u^t; \mathbf{a}) &= a_{21}y_{1,t-1} + a_{22}y_{2,t-1} + u_{2,t}, \qquad t = 1, 2, ..., \end{aligned}$$

where $y^{t-1} := (y_0, ..., y_{t-1})$ and $u^t := (u_1, ..., u_t)$ represent finite "histories" of the indicated variables. We also see that this system is recursive, and therefore has a unique fixed point.

The challenge to the PCM here is that it permits only a finite rather than a countable number of units: both the number of background variables (m) and endogenous variables (n) must be finite in the PCM, whereas the structural VAR requires a countable infinity of background and endogenous variables. In contrast, settable systems permit (but do not require) a countable infinity of units, readily accommodating structural VARs.

In line with our previous discussion, settable systems distinguish between system attributes $\mathbf{a}$ (a fixed vector) and structurally exogenous causal variables y_0 and $\{u_t\}$. The difference in the roles of y_0 and $\{u_t\}$ on the one hand and $\mathbf{a}$ on the other are particularly clear in this example. In the PCM, these are lumped together as background variables devoid of causal status. Since $\mathbf{a}$ is fixed, its lack of causal status is appropriate; indeed, $\mathbf{a}$ represents *effects* here[2], not causes. But the lack of causal status is problematic for the variables y_0 and $\{u_t\}$; for example, this prohibits discussing the effects of structural "shocks" u_t.

Observe that the structural VAR represents $u_{1,t}$ as a causal driver of $y_{1,t}$, as is standard. Nevertheless, settable systems do not admit "instantaneous" causation, so even though $u_{1,t}$ has the same time index as $y_{1,t}$, i.e. t, we adopt the convention that $u_{1,t}$ is realized prior to $y_{1,t}$. That is, there must be some positive time interval $\delta > 0$, no matter how small, separating these realizations. For example, δ can represent the amount of time it takes to compute $y_{1,t}$ once all its determinants are in place. Strictly speaking, then, we could write $u_{1,t-\delta}$ in place of $u_{1,t}$, but for notational convenience, we leave this implicit. We refer to this as "contemporaneous" causation to distinguish it from instantaneous causation.

A common focus of interest when applying structural VARs is to learn the coefficient vector $\mathbf{a}$. In applications, it is typically assumed that the realizations $\{y_t\}$ are observed, whereas $\{u_t\}$ is unobserved. The least squares estimator for a sample of size T, say $\hat{\mathbf{a}}_T$, is commonly used to learn (estimate) $\mathbf{a}$ in such cases. This estimator is a straightforward function of y^T, say $\hat{\mathbf{a}}_T = r_{\mathbf{a},T}(y^T)$. If $\{u_t\}$ is generated as a realization of a sequence of mean zero finite variance independent identically distributed (IID) random variables, then $\hat{\mathbf{a}}_T$ generally converges to $\mathbf{a}$ with probability one as $T \to \infty$, implying that $\mathbf{a}$ can be fully learned in the limit. Viewing $\hat{\mathbf{a}}_T$ as causally determined by y^T, we see that we require a countable number of units to treat this learning problem.

As these examples demonstrate, the PCM exhibits a number of features that limit its applicability to systems involving optimization, equilibrium, and learning. These limitations motivate a variety of features of settable systems, extending the PCM in ways that permit straightforward treatment of such systems. We now turn to a more complete description of the SS framework.

3.2 Formal Settable Systems

We now provide a formal description of settable systems that readily accommodates causal discourse in the foregoing examples and that also suffices to establish the desired linkage between Granger causality and causal notions in the PCM. The material that follows is adapted from Chalak and White (2010). For additional details, see WC.

2. For example, $(\partial/\partial y_{1,t-1}) r_{1,t}(y^{t-1}, e^t; a) = a_{11}$ can be interpreted as the marginal effect of $y_{1,t-1}$ on $y_{1,t}$.

A *stochastic settable system* is a mathematical framework in which a countable number of *units* i, $i = 1, ..., n$, interact under uncertainty. Here, $n \in \bar{\mathbb{N}}^+ := \mathbb{N}^+ \cup \{\infty\}$, where $\mathbb{N}^+$ denotes the positive integers. When $n = \infty$, we interpret $i = 1, ..., n$ as $i = 1, 2, ...$ Units have *attributes* $a_i \in A$; these are fixed for each unit, but may vary across units. Each unit also has associated random variables, defined on a measurable space $(\Omega, \mathcal{F})$. It is convenient to define a *principal space* Ω_0 and let $\Omega := \times_{i=0}^{n} \Omega_i$, with each Ω_i a copy of Ω_0. Often, $\Omega_0 = \mathbb{R}$ is convenient. A probability measure $P_{\mathbf{a}}$ on $(\Omega, \mathcal{F})$ assigns probabilities to events involving random variables. As the notation suggests, $P_{\mathbf{a}}$ can depend on the attribute vector $\mathbf{a} := (a_1, ..., a_n) \in \mathbf{A} := \times_{i=1}^{n} A$.

The random variables associated with unit i define a *settable variable* $\mathcal{X}_i$ for that unit. A settable variable $\mathcal{X}_i$ has a dual aspect. It can be *set* to a random variable denoted by Z_i (the *setting*), where $Z_i : \Omega_i \to \mathbb{S}_i$. $\mathbb{S}_i$ denotes the *admissible setting values* for Z_i, a multi-element subset of $\mathbb{R}$. Alternatively, the settable variable can be *free* to respond to settings of other settable variables. In the latter case, it is denoted by the *response* $Y_i : \Omega \to \mathbb{S}_i$. The response Y_i of a settable variable $\mathcal{X}_i$ to the settings of other settable variables is determined by a *response function*, r_i. For example, r_i can be determined by optimization, determining the response for unit i that is best in some sense, given the settings of other settable variables. The dual role of a settable variable $\mathcal{X}_i : \{0, 1\} \times \Omega \to \mathbb{S}_i$, distinguishing responses $\mathcal{X}_i(0, \omega) := Y_i(\omega)$ and settings $\mathcal{X}_i(1, \omega) := Z_i(\omega_i)$, $\omega \in \Omega$, permits formalizing the directional nature of causal relations, whereby settings of some variables (causes) determine responses of others.

The *principal unit* $i = 0$ also plays a key role. We let the *principal setting* Z_0 and *principal response* Y_0 of the *principal settable variable* $\mathcal{X}_0$ be such that $Z_0 : \Omega_0 \to \Omega_0$ is the identity map, $Z_0(\omega_0) := \omega_0$, and we define $Y_0(\omega) := Z_0(\omega_0)$. The setting Z_0 of the principal settable variable may directly influence all other responses in the system, whereas its response Y_0 is unaffected by other settings. Thus, $\mathcal{X}_0$ supports introducing an aspect of "pure randomness" to responses of settable variables.

3.2.1 Elementary Settable Systems

In *elementary* settable systems, Y_i is determined (actually or potentially) by the settings of *all* other system variables, denoted $Z_{(i)}$. Thus, in elementary settable systems, $Y_i = r_i(Z_{(i)}; \mathbf{a})$. The settings $Z_{(i)}$ take values in $\mathbb{S}_{(i)} \subseteq \Omega_0 \times_{j \neq i} \mathbb{S}_j$. We have that $\mathbb{S}_{(i)}$ is a strict subset of $\Omega_0 \times_{j \neq i} \mathbb{S}_j$ if there are joint restrictions on the admissible settings values, for example, when certain elements of $\mathbb{S}_{(i)}$ represent probabilities that sum to one.

We now give a formal definition of elementary settable systems.

Definition 3.1 (Elementary Settable System) *Let* $\mathbf{A}$ *be a set and let* **attributes** $\mathbf{a} \in \mathbf{A}$ *be given. Let* $n \in \bar{\mathbb{N}}^+$ *be given, and let* $(\Omega, \mathcal{F}, P_{\mathbf{a}})$ *be a complete probability space such that* $\Omega := \times_{i=0}^{n} \Omega_i$*, with each* Ω_i *a copy of the* **principal space** Ω_0*, containing at least two elements.*

Let the **principal setting** $Z_0 : \Omega_0 \to \Omega_0$ *be the identity mapping. For* $i = 1, 2, ..., n$*, let* $\mathbb{S}_i$ *be a multi-element Borel-measurable subset of* $\mathbb{R}$ *and let* **settings** $Z_i : \Omega_i \to \mathbb{S}_i$ *be surjective measurable functions. Let* $Z_{(i)}$ *be the vector including every setting except* Z_i *and taking values in* $\mathbb{S}_{(i)} \subseteq \Omega_0 \times_{j \neq i} \mathbb{S}_j$*,* $\mathbb{S}_{(i)} \neq \varnothing$*. Let* **response functions** $r_i(\,\cdot\,; \mathbf{a}) : \mathbb{S}_{(i)} \to \mathbb{S}_i$ *be measurable functions and define* **responses** $Y_i(\omega) := r_i(Z_{(i)}(\omega); \mathbf{a})$*. Define* **settable**

variables $\mathcal{X}_i : \{0,1\} \times \Omega \to \mathbb{S}_i$ *as*

$$\mathcal{X}_i(0,\omega) := Y_i(\omega) \quad \textit{and} \quad \mathcal{X}_i(1,\omega) := Z_i(\omega_i), \quad \omega \in \Omega.$$

Define Y_0 *and* $\mathcal{X}_0$ *by* $Y_0(\omega) := \mathcal{X}_0(0,\omega) := \mathcal{X}_0(1,\omega) := Z_0(\omega_0)$, $\omega \in \Omega$.

Put $\mathcal{X} := \{\mathcal{X}_0, \mathcal{X}_1, ...\}$. *The triple* $\mathcal{S} := \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, P_\mathbf{a}), \mathcal{X}\}$ *is an* **elementary settable system**.

An elementary settable system thus comprises an attribute component, $(\mathbf{A}, \mathbf{a})$, a stochastic component, $(\Omega, \mathcal{F}, P_\mathbf{a})$, and a structural or causal component $\mathcal{X}$, consisting of settable variables whose properties are crucially determined by response functions $r := \{r_i\}$. It is formally correct to write $\mathcal{X}_\mathbf{a}$ instead of $\mathcal{X}$; we write $\mathcal{X}$ for simplicity.

Note the absence of any fixed point requirement, the distinct roles played by fixed attributes $\mathbf{a}$ and setting variables Z_i (including principal settings Z_0), and the countable number of units allowed.

Example 3.1 is covered by this definition. There, $n = 2$. Attributes $\mathbf{a} := (S_1, u_1, S_2, u_2)$ belong to a suitably chosen set $\mathbf{A}$. Here, $\mathbb{S}_i = S_i$. We take $z_i = Z_i(\omega_i)$, $\omega_i \in \Omega_i$ and $y_i = Y_i(\omega) = r_i^e(Z_{(i)}(\omega); \mathbf{a}) = r_i^e(z_{(i)}; \mathbf{a})$, $i = 1, 2$. The "e" superscript in r_i^e emphasizes that the response function is for an elementary settable system. In the example games, the responses y_i only depend on settings (z_1, z_2). In more elaborate games, dependence on $z_0 = \omega_0$ can accommodate random responses.

3.2.2 Partitioned Settable Systems

In elementary settable systems, each single response Y_i can freely respond to settings of all other system variables. We now consider systems where several settable variables jointly respond to settings of the remaining settable variables, as when responses represent the solution to a joint optimization problem. For this, *partitioned* settable systems group jointly responding variables into blocks. In elementary settable systems, every unit i forms a block by itself. We now define general partitioned settable systems.

Definition 3.2 (Partitioned Settable System) *Let* $(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}, P_\mathbf{a})$, $\mathcal{X}_0$, n, *and* $\mathbb{S}_i$, $i = 1, ..., n$, *be as in Definition 3.1. Let* $\Pi = \{\Pi_b\}$ *be a partition of* $\{1, ..., n\}$, *with cardinality* $B \in \bar{\mathbb{N}}^+$ $(B := \#\Pi)$.

For $i = 1, 2, ..., n$, *let* Z_i^Π *be settings and let* $Z_{(b)}^\Pi$ *be the vector containing* Z_0 *and* $Z_i^\Pi, i \notin \Pi_b$, *and taking values in* $\mathbb{S}_{(b)}^\Pi \subseteq \Omega_0 \times_{i \notin \Pi_b} \mathbb{S}_i$, $\mathbb{S}_{(b)}^\Pi \neq \varnothing$, $b = 1, ..., B$. *For* $b = 1, ..., B$ *and* $i \in \Pi_b$, *suppose there exist measurable functions* $r_i^\Pi(\,\cdot\,; \mathbf{a}) : \mathbb{S}_{(b)}^\Pi \to \mathbb{S}_i$, *specific to* Π *such that responses* $Y_i^\Pi(\omega)$ *are jointly determined as*

$$Y_i^\Pi := r_i^\Pi(Z_{(b)}^\Pi; \mathbf{a}).$$

Define the settable variables $\mathcal{X}_i^\Pi : \{0,1\} \times \Omega \to \mathbb{S}_i$ as

$$\mathcal{X}_i^\Pi(0,\omega) := Y_i^\Pi(\omega) \quad \textit{and} \quad \mathcal{X}_i^\Pi(1,\omega) := Z_i^\Pi(\omega_i) \quad \omega \in \Omega.$$

Put $\mathcal{X}^\Pi := \{\mathcal{X}_0, \mathcal{X}_1^\Pi, \mathcal{X}_2^\Pi ...\}$. *The triple* $\mathcal{S} := \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}), (\Pi, \mathcal{X}^\Pi)\}$ *is a* **partitioned settable system**.

The settings $Z^{\Pi}_{(b)}$ may be partition-specific; this is especially relevant when the admissible set $\mathbb{S}^{\Pi}_{(b)}$ imposes restrictions on the admissible values of $Z^{\Pi}_{(b)}$. Crucially, response functions and responses are partition-specific. In Definition 3.2, the joint response function $r^{\Pi}_{[b]} := (r^{\Pi}_i, i \in \Pi_b)$ specifies how the settings $Z^{\Pi}_{(b)}$ outside of block Π_b determine the joint response $Y^{\Pi}_{[b]} := (Y^{\Pi}_i, i \in \Pi_b)$, i.e., $Y^{\Pi}_{[b]} = r^{\Pi}_{[b]}(Z^{\Pi}_{(b)}; \mathbf{a})$. For convenience below, we let $\Pi_0 = \{0\}$ represent the block corresponding to $\mathcal{X}_0$.

Example 3.2 makes use of partitioning. Here, we have $n = 4$ settable variables with $B = 2$ blocks. Let settable variables 1 and 2 correspond to beer and pizza consumption, respectively, and let settable variables 3 and 4 correspond to price and income. The *agent* partition groups together all variables under the control of a given agent. Let the consumer be agent 2, so $\Pi_2 = \{1, 2\}$. Let the rest of the economy, determining price and income, be agent 1, so $\Pi_1 = \{3, 4\}$. The agent partition is $\Pi^a = \{\Pi_1, \Pi_2\}$. Then for block 2,

$$\begin{aligned} y_1 &= Y^a_1(\omega) = r^a_1(Z_0(\omega_0), Z^a_3(\omega_3), Z^a_4(\omega_4); \mathbf{a}) = r^a_1(p, m; \mathbf{a}) \\ y_2 &= Y^a_2(\omega) = r^a_2(Z_0(\omega_0), Z^a_3(\omega_3), Z^a_4(\omega_4); \mathbf{a}) = r^a_2(p, m; \mathbf{a}) \end{aligned}$$

represents the joint demand for beer and pizza (belonging to block 2) as a function of settings of price and income (belonging to block 1). This joint demand is unique under mild conditions. Observe that $z_0 = Z_0(\omega_0)$ formally appears as an allowed argument of r^a_i after the second equality, but when the consumer's optimization problem has a unique solution, there is no need for a random component to demand. We thus suppress this argument in writing $r^a_i(p, m; \mathbf{a})$, $i = 1, 2$. Nevertheless, when the solution to the consumer's optimization problem is not unique, a random component can act to ensure a unique consumer demand. We do not pursue this here; WC provide related discussion.

We write the block 1 responses for the price and income settable variables as

$$\begin{aligned} y_3 &= Y^a_3(\omega) = r^a_3(Z_0(\omega_0), Z^a_1(\omega_1), Z^a_2(\omega_2); \mathbf{a}) = r^a_3(z_0; \mathbf{a}) \\ y_4 &= Y^a_4(\omega) = r^a_4(Z_0(\omega_0), Z^a_1(\omega_1), Z^a_2(\omega_2); \mathbf{a}) = r^a_4(z_0; \mathbf{a}). \end{aligned}$$

In this example, price and income are not determined by the individual consumer's demands, so although $Z^a_1(\omega_1)$ and $Z^a_2(\omega_2)$ formally appear as allowed arguments of r^a_i after the second equality, we suppress these in writing $r^a_i(z_0; \mathbf{a})$, $i = 3, 4$. Here, price and income responses (belonging to block 1) are determined solely by block 0 settings $z_0 = Z_0(\omega_0) = \omega_0$. This permits price and income responses to be randomly distributed, under the control of $P_{\mathbf{a}}$.

It is especially instructive to consider the elementary partition for this example, $\Pi^e = \{\{1\}, \{2\}, \{3\}, \{4\}\}$, so that $\Pi_i = \{i\}$, $i = 1, ..., 4$. The elementary partition specifies how each system variable freely responds to settings of all other system variables. In particular, it is easy to verify that when consumption of pizza is set to a given level, the consumer's optimal response is to spend whatever income is left on beer, and vice versa. Thus,

$$\begin{aligned} y_1 &= r^e_1(Z_0(\omega_0), Z^e_2(\omega_2), Z^e_3(\omega_3), Z^e_4(\omega_4); \mathbf{a}) = r^e_1(z_2, p, m; \mathbf{a}) = m - pz_2 \\ y_2 &= r^e_2(Z_0(\omega_0), Z^e_1(\omega_2), Z^e_3(\omega_3), Z^e_4(\omega_4); \mathbf{a}) = r^e_2(z_1, p, m; \mathbf{a}) = (m - z_1)/p. \end{aligned}$$

Replacing (y_1, y_2) with (z_1, z_2), we see that this system does not have a unique fixed point, as any (z_1, z_2) such that $m = z_1 + pz_2$ satisfies both

$$z_1 = m - pz_2 \qquad \text{and} \qquad z_2 = (m - z_1)/p.$$

Causal discourse in the PCM is ruled out by the lack of a fixed point. Nevertheless, the settable systems framework supports the natural economic causal discourse here about effects of prices, income, and, e.g., pizza consumption on beer demand. Further, in settable systems, the governing principle of optimization (embedded in $\mathbf{a}$) ensures that the response functions for both the agent partition and the elementary partition are mutually consistent.

3.2.3 Recursive and Canonical Settable Systems

The link between Granger causality and the causal notions of the PCM emerges from a particular class of *recursive* partitioned settable systems that we call *canonical* settable systems, where the system evolves naturally without intervention. This corresponds to what are also called "idle regimes" in the literature (Pearl, 2000; Eichler and Didelez, 2009; Dawid, 2010).

To define recursive settable systems, for $b \geq 0$ define $\Pi_{[0:b]} := \Pi_0 \cup ... \cup \Pi_{b-1} \cup \Pi_b$.

Definition 3.3 (Recursive Partitioned Settable System) *Let $\mathcal{S}$ be a partitioned settable system. For $b = 0, 1, ..., B$, let $Z^{\Pi}_{[0:b]}$ denote the vector containing the settings Z^{Π}_i for $i \in \Pi_{[0:b]}$ and taking values in $\mathbb{S}_{[0:b]} \subseteq \Omega_0 \times_{i \in \Pi_{[1:b]}} \mathbb{S}_i$, $\mathbb{S}_{[0:b]} \neq \varnothing$. For $b = 1, ..., B$ and $i \in \Pi_b$, suppose that $r^{\Pi} := \{r^{\Pi}_i\}$ is such that the responses $Y^{\Pi}_i = \mathcal{X}^{\Pi}_i(1, \cdot)$ are determined as*

$$Y^{\Pi}_i := r^{\Pi}_i(Z^{\Pi}_{[0:b-1]}; \mathbf{a}).$$

Then we say that Π is a **recursive partition**, *that r^{Π} is* **recursive**, *and that $\mathcal{S} := \{(\mathbf{A}, \mathbf{a}), (\Omega, \mathcal{F}), (\Pi, \mathcal{X}^{\Pi})\}$ is a* **recursive partitioned settable system** *or simply that $\mathcal{S}$ is* **recursive**.

Example 3.2 is a recursive settable system, as the responses of block 1 depend on the settings of block 0, and the responses of block 2 depend on the settings of block 1.

Canonical settable systems are recursive settable systems in which the settings for a given block equal the responses for that block, i.e.,

$$Z^{\Pi}_{[b]} = Y^{\Pi}_{[b]} := r^{\Pi}_{[b]}(Z^{\Pi}_{[0:b-1]}; \mathbf{a}), \qquad b = 1, ..., B.$$

Without loss of generality, we can represent canonical responses and settings solely as a function of ω_0, so that

$$Z^{\Pi}_{[b]}(\omega_0) = Y^{\Pi}_{[b]}(\omega_0) := r^{\Pi}_{[b]}(Z^{\Pi}_{[0:b-1]}(\omega_0); \mathbf{a}), \qquad b = 1, ..., B.$$

The *canonical representation* drops the distinction between settings and responses; we write

$$Y^{\Pi}_{[b]} = r^{\Pi}_{[b]}(Y^{\Pi}_{[0:b-1]}; \mathbf{a}), \qquad b = 1, ..., B.$$

It is easy to see that the structural VAR of Example 3.3 corresponds to the canonical representation of a canonical settable system. The canonical responses y_0 and $\{u_t\}$ belong to the first block, and canonical responses $y_t = (y_{1,t}, y_{2,t})$ belong to block $t + 1$, $t = 1, 2, ...$ Example 3.3 implements the *time* partition, where joint responses for a given time period depend on previous settings.

4. Causality in Settable Systems and in the PCM

In this section we examine the relations between concepts of direct causality in settable systems and in the PCM, specifically the PCM notions of direct cause and controlled direct effect (Pearl, 2000, p.222; Pearl, 2001, definition 1). The close correspondence between these notions for the recursive systems relevant to Granger causality enables us to take the first step in linking Granger causality and causal notions in the PCM. Section 5 completes the chain by linking direct structural causality and Granger causality.

4.1 Direct Structural Causality in Settable Systems

Direct structural causality is defined for both recursive and non-recursive partitioned settable systems. For notational simplicity in what follows, we may drop the explicit partition superscript Π when the specific partition is clearly understood. Thus, we may write Y, Z, and $\mathcal{X}$ in place of the more explicit Y^{Π}, Z^{Π}, and $\mathcal{X}^{\Pi}$ when there is no possibility of confusion.

Let $\mathcal{X}_j$ belong to block b ($j \in \Pi_b$). Heuristically, we say that a settable variable $\mathcal{X}_i$, outside of block b, *directly causes* $\mathcal{X}_j$ *in* $\mathcal{S}$ when the response for $\mathcal{X}_j$ differs for different settings of $\mathcal{X}_i$, while holding all other variables outside of block b to the same setting values. There are two main ingredients to this notion. The first ingredient is an admissible intervention. To define this, let $z^*_{(b);i}$ denote the vector otherwise identical to $z_{(b)}$, but replacing elements z_i with z^*_i. An *admissible intervention* $z_{(b)} \to z^*_{(b);i} := (z_{(b)}, z^*_{(b);i})$ is a pair of distinct elements of $\mathbb{S}_{(b)}$. The second ingredient is the behavior of the response under this intervention.

We formalize this notion of direct causality as follows.

Definition 4.1 (Direct Causality) *Let $\mathcal{S}$ be a partitioned settable system. For given positive integer b, let $j \in \Pi_b$. (i) For given $i \notin \Pi_b$, $\mathcal{X}_i$* **directly causes** *$\mathcal{X}_j$* **in** *$\mathcal{S}$ if there exists an admissible intervention $z_{(b)} \to z^*_{(b);i}$ such that*

$$r_j(z^*_{(b);i}; \mathbf{a}) - r_j(z_{(b)}; \mathbf{a}) \neq 0,$$

and we write $\mathcal{X}_i \overset{D}{\Rightarrow}_{\mathcal{S}} \mathcal{X}_j$. Otherwise, we say $\mathcal{X}_i$ **does not directly cause** *$\mathcal{X}_j$* **in** *$\mathcal{S}$ and write $\mathcal{X}_i \overset{D}{\not\Rightarrow}_{\mathcal{S}} \mathcal{X}_j$. (ii) For $i, j \in \Pi_b$, $\mathcal{X}_i \overset{D}{\not\Rightarrow}_{\mathcal{S}} \mathcal{X}_j$.*

We emphasize that although we follow the literature in referring to "interventions," with their mechanistic or manipulative connotations, the formal concept only involves the properties of a response function on its domain.

By definition, variables within the same block do not directly cause each other. In particular $\mathcal{X}_i \overset{D}{\not\Rightarrow}_{\mathcal{S}} \mathcal{X}_i$. Also, Definition 4.1 permits *mutual causality*, so that $\mathcal{X}_i \overset{D}{\Rightarrow}_{\mathcal{S}} \mathcal{X}_j$ and $\mathcal{X}_j \overset{D}{\Rightarrow}_{\mathcal{S}} \mathcal{X}_i$ without contradiction for i and j in different blocks. Nevertheless, in recursive systems, mutual causality is ruled out: if $\mathcal{X}_i \overset{D}{\Rightarrow}_{\mathcal{S}} \mathcal{X}_j$ then $\mathcal{X}_j \overset{D}{\not\Rightarrow}_{\mathcal{S}} \mathcal{X}_i$.

We call the response value difference in Definition 4.1 the *direct effect of $\mathcal{X}_i$ on $\mathcal{X}_j$ in $\mathcal{S}$* of the specified intervention. Chalak and White (2010) also study various notions of indirect and total causality.

These notions of direct cause and direct effect are well defined regardless of whether or not the system possesses a unique fixed point. Further, all settable variables, including $\mathcal{X}_0$, can act as causes and have effects. On the other hand, attributes $\mathbf{a}$, being fixed, do not play a causal role. These definitions apply regardless of whether there is a finite or countable number of units. It is readily verified that this definition rigorously supports causal discourse in each of the examples of Section 3.

As we discuss next, in the recursive systems relevant for $G-$causality, these concepts correspond closely to notions of direct cause and "controlled" direct effect in Pearl (2000, 2001). To distinguish the settable system direct causality concept from Pearl's notion and later from Granger causality, we follow WL and refer to direct causality in settable systems as *direct structural causality.*

4.2 Direct Causes and Effects in the PCM

Pearl (2000, p.222), drawing on Galles and Pearl (1997), gives a succinct statement of the notion of direct cause, coherent with the PCM as specified in Section 2:

> *X is a direct cause of Y* if there exist two values x and x' of X and a value u of U such that $Y_{xr}(u) \neq Y_{x'r}(u)$, where r is some realization of $V\backslash\{X, Y\}$.

To make this statement fully meaningful requires applying Pearl's (2000) definitions 7.1.2 (Submodel) and 7.1.4 (Potential Response) to arrive at the *potential response,* $Y_{xr}(u)$. For brevity, we do not reproduce Pearl's definitions here. Instead, it suffices to map $Y_{xr}(u)$ and its elements to their settable system counterparts. Specifically, u corresponds to $(\mathbf{a}, z_0)$; x corresponds to z_i; r corresponds to the elements of $z_{(b)}$ other than z_0 and z_i, say $z_{(b)(i,0)}$; and, provided it exists, $Y_{xr}(u)$ corresponds to $r_j(z_{(b)}; \mathbf{a})$.

The caveat about the existence of $Y_{xr}(u)$ is significant, as $Y_{xr}(u)$ is not defined in the absence of a unique fixed point for the system. Further, even with a unique fixed point, the potential response $Y_{xr}(u)$ must also uniquely solve a set of equations denoted F_x (see Pearl, 2000, eq.(7.1)) for a submodel, and there is no general guarantee of such a solution. Fortunately, however, this caveat matters only for non-recursive PCMs. In the recursive case relevant for $G-$causality, the potential response is generally well defined.

Making a final identification between x' and z_i^*, and given the existence of potential responses $Y_{x'r}(u)$ and $Y_{xr}(u)$, we see that $Y_{x'r}(u) \neq Y_{xr}(u)$ corresponds to the settable systems requirement $r_j(z^*_{(b);i}; \mathbf{a}) - r_j(z_{(b)}; \mathbf{a}) \neq 0$.

Pearl (2001, definition 1) gives a formal statement of the notion stated above, saying that if for given u and some r, x, and x' we have $Y_{xr}(u) \neq Y_{x'r}(u)$ then X *has a* ***controlled direct effect*** *on Y in model M and situation $U = u$.* In definition 2, Pearl (2001) labels $Y_{x'r}(u) - Y_{xr}(u)$ the *controlled direct effect*, corresponding to the direct structural effect $r_j(z^*_{(b);i}; \mathbf{a}) - r_j(z_{(b)}; \mathbf{a})$ defined for settable systems.

Thus, although there are important differences, especially in non-recursive systems, the settable systems and PCM notions of direct causality and direct effects closely correspond in recursive systems. These differences are sufficiently modest that the results of WL linking direct structural causality to Granger causality, discussed next, also serve to closely link the PCM notion of direct cause to that of Granger causality.

5. $G-$Causality and Direct Structural Causality

In this section we examine the relation between direct structural causality and Granger causality, drawing on results of WL. See WL for additional discussion and proofs of all formal results given here and in Section 6.

5.1 Granger Causality

Granger (1969) defined $G-$causality in terms of conditional expectations. Granger and Newbold (1986) gave a definition using conditional distributions. We work with the latter, as this is what relates generally to structural causality. In what follows, we adapt Granger and Newbold's notation, but otherwise preserve the conceptual content.

For any sequence of random vectors $\{Y_t,\ t = 0, 1, ...\}$, let $Y^t := (Y_0, ..., Y_t)$ denote its "$t-$history," and let $\sigma(Y^t)$ denote the sigma-field ("information set") generated by Y^t. Let $\{Q_t, S_t, Y_t\}$ be a sequence of random vectors. Granger and Newbold (1986) say that Q_{t-1} does not G-cause Y_{t+k} with respect to $\sigma(Q^{t-1}, S^{t-1}, Y^{t-1})$ if for all $t = 0, 1, ...,$

$$F_{t+k}(\ \cdot\ |\ Q^{t-1}, S^{t-1}, Y^{t-1}) = F_{t+k}(\ \cdot\ |\ S^{t-1}, Y^{t-1}), \quad k = 0, 1, ..., \tag{2}$$

where $F_{t+k}(\ \cdot\ |\ Q^{t-1}, S^{t-1}, Y^{t-1})$ denotes the conditional distribution function of Y_{t+k} given $Q^{t-1}, S^{t-1}, Y^{t-1}$, and $F_{t+k}(\ \cdot\ |\ S^{t-1}, Y^{t-1})$ denotes that of Y_{t+k} given S^{t-1}, Y^{t-1}. Here, we focus only on the $k = 0$ case, as this is what relates generally to structural causality.

As Florens and Mouchart (1982) and Florens and Fougère (1996) note, G non-causality is a form of conditional independence. Following Dawid (1979), we write $X \perp Y \mid Z$ when X and Y are independent given Z. Translating (2) gives the following version of the classical definition of Granger causality:

Definition 5.1 (Granger Causality) *Let $\{Q_t, S_t, Y_t\}$ be a sequence of random vectors. Suppose that*

$$Y_t \perp Q^{t-1} \mid Y^{t-1}, S^{t-1} \quad t = 1, 2, ... \ . \tag{3}$$

Then Q **does not $G-$cause Y with respect to** *S. Otherwise, Q $G-$***causes** *Y* **with respect to** *S.*

As it stands, this definition has no necessary structural content, as Q_t, S_t, and Y_t can be any random variables whatsoever. This definition relates solely to the ability of Q^{t-1} to help in predicting Y_t given Y^{t-1} and S^{t-1}.

In practice, researchers do not test classical $G-$causality, as this involves data histories of arbitrary length. Instead, researchers test a version of $G-$causality involving only a finite number of lags of Y_t, Q_t, and S_t. This does not test classical $G-$causality, but rather a related property, *finite-order* $G-$causality, that is neither necessary nor sufficient for classical $G-$causality.

Because of its predominant practical relevance, we focus here on finite-order rather than classical $G-$causality. (See WL for discussion of classical $G-$causality.) To define the finite-order concept, we define the finite histories $\boldsymbol{Y}_{t-1} := (Y_{t-\ell}, ..., Y_{t-1})$ and $\boldsymbol{Q}_t := (Q_{t-k}, ..., Q_t)$.

Definition 5.2 (Finite-Order Granger Causality) *Let $\{Q_t, S_t, Y_t\}$ be a sequence of random variables, and let $k \geq 0$ and $\ell \geq 1$ be given finite integers. Suppose that*

$$Y_t \perp \boldsymbol{Q}_t \mid \boldsymbol{Y}_{t-1}, S_t, \quad t = 1, 2, ... \ .$$

Then we say Q **does not finite-order** $G-$**cause** Y **with respect to** S. *Otherwise, we say* Q **finite-order** $G-$**causes** Y **with respect to** S.

We call $\max(k, \ell - 1)$ the "order" of the finite-order G non-causality.

Observe that $\boldsymbol{Q}_t$ replaces Q^{t-1} in the classical definition, that $\boldsymbol{Y}_{t-1}$ replaces Y^{t-1}, and that S_t replaces S^{t-1}. Thus, in addition to dropping all but a finite number of lags in Q^{t-1} and Y^{t-1}, this version includes Q_t. As WL discuss, however, the appearance of Q_t need not involve instantaneous causation. It suffices that realizations of Q_t precede those of Y_t, as in the case of contemporaneous causation discussed above. The replacement of S^{t-1} with S_t entails first viewing S_t as representing a finite history, and second the recognition that since S_t plays purely a conditioning role, there need be no restriction whatever on its timing. We thus call S_t "covariates." As WL discuss, the covariates can even include leads relative to time t. When covariate leads appear, we call this the "retrospective" case.

In what follows, when we refer to $G-$causality, it will be understood that we are referring to finite-order $G-$causality, as just defined. We will always refer to the concept of Definition 5.1 as *classical* $G-$causality to avoid confusion.

5.2 A Dynamic Structural System

We now specify a canonical settable system that will enable us to examine the relation between $G-$causality and direct structural causality. As described above, in such systems "predecessors" structurally determine "successors," but not vice versa. In particular, future variables cannot precede present or past variables, enforcing the causal direction of time. We write $Y \Leftarrow X$ to denote that Y succeeds X (X precedes Y). When Y and X have identical time indexes, $Y \Leftarrow X$ rules out instantaneous causation but allows contemporaneous causation.

We now specify a version of the causal data generating structures analyzed by WL and White and Kennedy (2009). We let $\mathbb{N}$ denote the integers $\{0, 1, ...\}$ and define $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$. For given $\ell, m, \in \mathbb{N}$, $\ell \geq 1$, we let $\boldsymbol{Y}_{t-1} := (Y_{t-\ell}, ..., Y_{t-1})$ as above; we also define $\boldsymbol{Z}_t := (Z_{t-m}, ..., Z_t)$. For simplicity, we keep attributes implicit in what follows.

Assumption A.1 Let $\{U_t, W_t, Y_t, Z_t;\ t = 0, 1, ...\}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$, a complete probability space, with U_t, W_t, Y_t, and Z_t taking values in $\mathbb{R}^{k_u}, \mathbb{R}^{k_w}, \mathbb{R}^{k_y}$, and $\mathbb{R}^{k_z}$ respectively, where $k_u \in \bar{\mathbb{N}}$ and $k_w, k_y, k_z \in \mathbb{N}$, with $k_y > 0$. Further, suppose that $Y_t \Leftarrow (Y^{t-1}, U^t, W^t, Z^t)$, where, for an unknown measurable $k_y \times 1$ function q_t, and for given $\ell, m, \in \mathbb{N}, \ell \geq 1$, $\{Y_t\}$ is structurally generated as

$$Y_t = q_t(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t, U_t), \quad t = 1, 2, ..., \tag{4}$$

such that, with $Y_t := (Y_{1,t}', Y_{2,t}')'$ and $U_t := (U_{1,t}', U_{2,t}')'$,

$$Y_{1,t} = q_{1,t}(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t, U_{1,t}) \qquad Y_{2,t} = q_{2,t}(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t, U_{2,t}).$$

Such structures are well suited to representing the structural evolution of time-series data in economic, biological, or other systems. Because Y_t is a vector, this covers the case of panel data, where one has a cross-section of time-series observations, as in fMRI or EEG data sets. For practical relevance, we explicitly impose the Markov assumption that Y_t is

determined by only a finite number of its own lags and those of Z_t and U_t. WL discuss the general case.

Throughout, we suppose that realizations of W_t, Y_t, and Z_t are observed, whereas realizations of U_t are not. Because U_t, W_t, or Z_t may have dimension zero, their presence is optional. Usually, however, some or all will be present. Since there may be a countable infinity of unobservables, there is no loss of generality in specifying that Y_t depends only on U_t rather than on a finite history of lags of U_t.

This structure is general: the structural relations may be nonlinear and non-monotonic in their arguments and non-separable between observables and unobservables. This system may generate stationary processes, non-stationary processes, or both. Assumption A.1 is therefore a general structural VAR; Example 3.3 is a special case.

The vector Y_t represents responses of interest. Consistent with a main application of $G-$causality, our interest here attaches to the effects on $Y_{1,t}$ of the lags of $Y_{2,t}$. We thus call $Y_{2,t-1}$ and its further lags "causes of interest." Note that A.1 specifies that $Y_{1,t}$ and $Y_{2,t}$ each have their own unobserved drivers, $U_{1,t}$ and $U_{2,t}$, as is standard.

The vectors U_t and Z_t contain causal drivers of Y_t whose effects are not of primary interest; we thus call U_t and Z_t "ancillary causes." The vector W_t may contain responses to U_t. Observe that W_t does not appear in the argument list for q_t, so it explicitly does not directly determine Y_t. Note also that $Y_t \Leftarrow (Y^{t-1}, U^t, W^t, Z^t)$ ensures that W_t is not determined by Y_t or its lags. A useful convention is that $W_t \Leftarrow (W^{t-1}, U^t, Z^t)$, so that W_t does not drive unobservables. If a structure does not have this property, then suitable substitutions can usually yield a derived structure satisfying this convention. Nevertheless, we do not require this, so W_t may also contain drivers of unobservable causes of Y_t.

For concreteness, we now specialize the settable systems definition of direct structural causality (Definition 4.1) to the specific system given in A.1. For this, let $\boldsymbol{y}_{\boldsymbol{s},t-1}$ be the sub-vector of $\boldsymbol{y}_{t-1}$ with elements indexed by the non-empty set $\boldsymbol{s} \subseteq \{1, ..., k_y\} \times \{t-\ell, ..., t-1\}$, and let $\boldsymbol{y}_{(\boldsymbol{s}),t-1}$ be the sub-vector of $\boldsymbol{y}_{t-1}$ with elements of $\boldsymbol{s}$ excluded.

Definition 5.3 (Direct Structural Causality) *Given A.1, for given $t > 0, j \in \{1, ..., k_y\}$, and $\boldsymbol{s}$, suppose that for all admissible values of $\boldsymbol{y}_{(\boldsymbol{s}),t-1}$, $\boldsymbol{z}_t$, and u_t, the function $\boldsymbol{y}_{\boldsymbol{s},t-1} \to q_{j,t}(\boldsymbol{y}_{t-1}, \boldsymbol{z}_t, u_t)$ is constant in $\boldsymbol{y}_{\boldsymbol{s},t-1}$. Then we say $\boldsymbol{Y}_{\boldsymbol{s},t-1}$* **does not directly structurally cause** *$Y_{j,t}$ and write $\boldsymbol{Y}_{\boldsymbol{s},t-1} \overset{D}{\not\Rightarrow}_{\mathcal{S}} Y_{j,t}$. Otherwise, we say $\boldsymbol{Y}_{\boldsymbol{s},t-1}$* **directly structurally causes** *$Y_{j,t}$ and write $\boldsymbol{Y}_{\boldsymbol{s},t-1} \overset{D}{\Rightarrow}_{\mathcal{S}} Y_{j,t}$.*

We can similarly define direct causality or non-causality of $Z_{\boldsymbol{s},t}$ or $U_{\boldsymbol{s},t}$ for $Y_{j,t}$, but we leave this implicit. We write, e.g., $\boldsymbol{Y}_{\boldsymbol{s},t-1} \overset{D}{\Rightarrow}_{\mathcal{S}} Y_t$ when $\boldsymbol{Y}_{\boldsymbol{s},t-1} \overset{D}{\Rightarrow}_{\mathcal{S}} Y_{j,t}$ for some $j \in \{1, ..., k_y\}$.

Building on work of White (2006a) and White and Kennedy (2009), WL discuss how certain exogeneity restrictions permit identification of expected causal effects in dynamic structures. Our next result shows that a specific form of exogeneity enables us to link direct structural causality and finite order $G-$causality. To state this exogeneity condition, we write $\boldsymbol{Y}_{1,t-1} := (Y_{1,t-\ell}, ..., Y_{1,t-1})$, $\boldsymbol{Y}_{2,t-1} := (Y_{2,t-\ell}, ..., Y_{2,t-1})$, and, for given $\tau_1, \tau_2 \geq 0$, $\boldsymbol{X}_t := (X_{t-\tau_1}, ..., X_{t+\tau_2})$, where $X_t := (W_t', Z_t')'$.

Assumption A.2 For ℓ and m as in A.1 and for $\tau_1 \geq m, \tau_2 \geq 0$, suppose that $\boldsymbol{Y}_{2,t-1} \perp U_{1,t} \mid (\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)$, $t = 1, ..., T - \tau_2$.

The classical *strict exogeneity* condition specifies that $(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t) \perp U_{1,t}$, which implies $\boldsymbol{Y}_{2,t-1} \perp U_{1,t} \mid (\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t)$. (Here, W_t can be omitted.) Assumption A.2 is a weaker requirement, as it may hold when strict exogeneity fails. Because of the conditioning involved, we call this *conditional exogeneity*. Chalak and White (2010) discuss structural restrictions for canonical settable systems that deliver conditional exogeneity. Below, we also discuss practical tests for this assumption.

Because of the finite numbers of lags involved in A.2, this is a *finite-order* conditional exogeneity assumption. For convenience and because no confusion will arise here, we simply refer to this as "conditional exogeneity."

Assumption A.2 ensures that expected direct effects of $\boldsymbol{Y}_{2,t-1}$ on $Y_{1,t}$ are identified. As WL note, it suffices for A.2 that $U^{t-1} \perp U_{1,t} \mid (Y_0, Z^{t-1}, \boldsymbol{X}_t)$ and $\boldsymbol{Y}_{2,t-1} \perp (Y_0, Z^{t-\tau_1-1}) \mid (\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)$. Imposing $U^{t-1} \perp U_{1,t} \mid (Y_0, Z^{t-1}, \boldsymbol{X}_t)$ is the analog of requiring that serial correlation is absent when lagged dependent variables are present. Imposing $\boldsymbol{Y}_{2,t-1} \perp (Y_0, Z^{t-\tau_1-1}) \mid (\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)$ ensures that ignoring Y_0 and omitting distant lags of Z_t from $\boldsymbol{X}_t$ doesn't matter.

Our first result linking direct structural causality and $G-$causality shows that, given A.1 and A.2 and with proper choice of Q_t and S_t, $G-$causality implies direct structural causality.

Proposition 5.4 *Let A.1 and A.2 hold. If* $\boldsymbol{Y}_{2,t-1} \overset{D}{\not\Rightarrow}_{\mathcal{S}} Y_{1,t}$, $t = 1, 2, ...,$ *then* $\boldsymbol{Y}_2$ *does not finite order* $G-$*cause* Y_1 *with respect to* $\boldsymbol{X}$*, i.e.,*

$$Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t, \quad t = 1, ..., T - \tau_2.$$

In stating G non-causality, we make the explicit identifications $\boldsymbol{Q}_t = \boldsymbol{Y}_{2,t-1}$ and $S_t = \boldsymbol{X}_t$.

This result leads one to ask whether the converse relation also holds: does direct structural causality imply $G-$causality? Strictly speaking, the answer is no. WL discuss several examples. The main issue is that with suitably chosen causal and probabilistic relationships, $\boldsymbol{Y}_{2,t-1}$ can cause $Y_{1,t}$, but $\boldsymbol{Y}_{2,t-1}$ and $Y_{1,t}$ can be independent, conditionally or unconditionally, i.e. Granger non-causal.

As WL further discuss, however, these examples are exceptional, in the sense that mild perturbations to their structure destroy the Granger non-causality. WL introduce a refinement of the notion of direct structural causality that accommodates these special cases and that does yield a converse result, permitting a characterization of structural and Granger causality. Let $supp(Y_{1,t})$ denote the support of $Y_{1,t}$, i.e., the smallest set containing $Y_{1,t}$ with probability 1, and let $F_{1,t}(\cdot \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)$ denote the conditional distribution function of $U_{1,t}$ given $\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$. WL introduce the following definition:

Definition 5.5 *Suppose A.1 holds and that for given* $\tau_1 \geq m, \tau_2 \geq 0$ *and for each* $y \in supp(Y_{1,t})$ *there exists a* $\sigma(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)-$*measurable version of the random variable*

$$\int 1\{q_{1,t}(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t, u_{1,t}) < y\} \, dF_{1,t}(u_{1,t} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t).$$

Then $\boldsymbol{Y}_{2,t-1} \overset{D}{\not\Rightarrow}_{\mathcal{S}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)} Y_{1,t}$ (**direct non-causality**$-\sigma(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)$ *a.s.*)*. If not,* $\boldsymbol{Y}_{2,t-1} \overset{D}{\Rightarrow}_{\mathcal{S}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)} Y_{1,t}$.

For simplicity, we refer to this as *direct non-causality a.s.* The requirement that the integral in this definition is $\sigma(\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t)-$measurable means that the integral does not depend on $\boldsymbol{Y}_{2,t-1}$, despite its appearance inside the integral as an argument of $q_{1,t}$. For this, it suffices that $\boldsymbol{Y}_{2,t-1}$ does not directly cause $Y_{1,t}$; but it is also possible that $q_{1,t}$ and the conditional distribution of $U_{1,t}$ given $\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ are in just the right relation to hide the structural causality. Without the ability to manipulate this distribution, the structural causality will not be detectable. One possible avenue to manipulating this distribution is to modify the choice of $\boldsymbol{X}_t$, as there are often multiple choices for $\boldsymbol{X}_t$ that can satisfy A.2 (see White and Lu, 2010b). For brevity and because hidden structural causality is an exceptional circumstance, we leave aside further discussion of this possibility here. The key fact to bear in mind is that the causal concept of Definition 5.5 distinguishes between those direct causal relations that are empirically detectable and those that are not, for a given set of covariates $\boldsymbol{X}_t$.

We now give a structural characterization of $G-$causality for structural VARs:

Theorem 5.6 *Let A.1 and A.2 hold. Then* $\boldsymbol{Y}_{2,t-1} \overset{D}{\not\Rightarrow}_{\mathcal{S}(\boldsymbol{Y}_{1,t-1},\boldsymbol{X}_t)} Y_{1,t}$, $t = 1, ..., T - \tau_2$, *if and only if*

$$Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t, \quad t = 1, ..., T - \tau_2,$$

i.e., $\boldsymbol{Y}_2$ *does not finite-order* $G-$*cause* Y_1 *with respect to* $\boldsymbol{X}$.

Thus, given conditional exogeneity of $\boldsymbol{Y}_{2,t-1}$, *G non-causality implies direct non-causality a.s. and vice-versa*, justifying tests of direct non-causality *a.s.* in structural VARs using tests for $G-$causality.

This result completes the desired linkage between $G-$causality and direct causality in the PCM. Because direct causality in the recursive PCM corresponds essentially to direct structural causality in canonical settable systems, and because the latter is essentially equivalent to $G-$causality, as just shown, direct causality in the PCM is essentially equivalent to $G-$causality, provided A.1 and A.2 hold.

5.3 The Central Role of Conditional Exogeneity

To relate direct structural causality to $G-$causality, we maintain A.2, a specific conditional exogeneity assumption. Can this assumption be eliminated or weakened? We show that the answer is no: A.2 is in a precise sense a necessary condition. We also give a result supporting tests for conditional exogeneity.

First, we specify the sense in which conditional exogeneity is necessary for the equivalence of $G-$causality and direct structural causality.

Proposition 5.7 *Given A.1, suppose that* $\boldsymbol{Y}_{2,t-1} \overset{D}{\not\Rightarrow}_{\mathcal{S}} Y_{1,t}$, $t = 1, 2, ...$. *If A.2 does not hold, then for each t there exists $q_{1,t}$ such that* $Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ *does not hold.*

That is, if conditional exogeneity does not hold, then there are always structures that generate data exhibiting $G-$causality, despite the absence of direct structural causality. Because $q_{1,t}$ is unknown, this worst case scenario can never be discounted. Further, as WL show, the class of worst case structures includes precisely those usually assumed in

applications, namely separable structures (e.g., $Y_{1,t} = q_{1,t}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t) + U_{1,t}$), as well as the more general class of invertible structures. Thus, in the cases typically assumed in the literature, the failure of conditional exogeneity *guarantees* $G-$causality in the absence of structural causality. We state this formally as a corollary.

Corollary 5.8 *Given A.1 with* $\boldsymbol{Y}_{2,t-1} \overset{D}{\not\to}_{\mathcal{S}} Y_{1,t}$, $t = 1, 2, ...$, *suppose that* $q_{1,t}$ *is invertible in the sense that* $Y_{1,t} = q_{1,t}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t, U_{1,t})$ *implies the existence of* $\xi_{1,t}$ *such that* $U_{1,t} = \xi_{1,t}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t, Y_{1,t})$, $t = 1, 2, ...$. *If A.2 fails, then* $Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ *fails,* $t = 1, 2, ...$

Together with Theorem 5.6, this establishes that in the absence of direct causality and for the class of invertible structures predominant in applications, *conditional exogeneity is necessary and sufficient for G non-causality.*

Tests of conditional exogeneity for the general separable case follow from:

Proposition 5.9 *Given A.1, suppose that* $E(Y_{1,t}) < \infty$ *and that*

$$q_{1,t}(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t, U_{1,t}) = \zeta_t(\boldsymbol{Y}_{t-1}, \boldsymbol{Z}_t) + \upsilon_t(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t, U_{1,t}),$$

where ζ_t *and* υ_t *are unknown measurable functions. Let* $\varepsilon_t := Y_{1,t} - E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$. *If A.2 holds, then*

$$\begin{aligned} \varepsilon_t = \upsilon_t(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t, U_{1,t}) - E(\upsilon_t(\boldsymbol{Y}_{1,t-1}, \boldsymbol{Z}_t, U_{1,t}) \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t) \\ E(\varepsilon_t|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t) = E(\varepsilon_t|\boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t) = 0 \qquad \textit{and} \\ \boldsymbol{Y}_{2,t-1} \perp \varepsilon_t \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t. \end{aligned} \tag{5}$$

Tests based on this result detect the failure of A.2, given separability. Such tests are feasible because even though the regression error ε_t is unobserved, it can be consistently estimated, say as $\hat{\varepsilon}_t := Y_{1,t} - \hat{E}(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$, where $\hat{E}(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$ is a parametric or nonparametric estimator of $E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$. These estimated errors can then be used to test (5). If we reject (5), then we must reject A.2. We discuss a practical procedure in the next section. WL provide additional discussion.

WL also discuss dropping the separability assumption. For brevity, we maintain separability here. Observe that under the null of direct non-causality, $q_{1,t}$ is necessarily separable, as then ζ_t is the zero function.

6. Testing Direct Structural Causality

Here, we discuss methods for testing direct structural causality. First, we discuss a general approach that combines tests of G non-causality (GN) and conditional exogeneity (CE). Then we describe straightforward practical methods for implementing the general approach.

6.1 Combining Tests for GN and CE

Theorem 5.6 implies that if we test and reject GN, then we must reject either direct structural non-causality (SN) or CE, or both. If CE is maintained, then we can directly test SN by testing GN; otherwise, a direct test is not available.

Similarly, under the traditional separability assumption, Corollary 5.8 implies that if we test and reject CE, then we must reject either SN or GN (or both). If GN is maintained, then we can directly test SN by testing CE; otherwise, a direct test is not available.

When neither CE nor GN is maintained, no direct test of SN is possible. Nevertheless, we can test structural causality indirectly by combining the results of the GN and CE tests to isolate the source of any rejections. WL propose the following indirect test:

(1) *Reject* SN if either:

(i) the CE test *fails to reject* and the GN test *rejects*; or

(ii) the CE test *rejects* and the GN test *fails to reject.*

If these rejection conditions do not hold, however, we cannot just decide to "accept" (i.e., fail to reject) SN. As WL explain in detail, difficulties arise when CE and GN both fail, as failing to reject SN here runs the risk of Type II error, whereas rejecting SN runs the risk of Type I error. We resolve this dilemma by specifying the further rules:

(2) *Fail to reject* SN if the CE and GN tests both *fail to reject;*

(3) *Make no decision* as to SN if the CE and GN tests both *reject.*

In the latter case, we conclude only that CE and GN both fail, thereby obstructing structural inference. This sends a clear signal that the researcher needs to revisit the model specification, with particular attention to specifying covariates sufficient to ensure conditional exogeneity.

Because of the structure of this indirect test, it is not enough simply to consider its level and power. We must also account for the possibility of making no decision. For this, define

$$\begin{aligned} p &:= P[\text{ wrongly make a decision }] \\ &= P[\text{ fail to reject CE or GN } \mid \text{ CE is false and GN is false }] \\ q &:= P[\text{ wrongly make no decision }] \\ &= P[\text{ reject CE and GN } \mid \text{ CE is true or GN is true }]. \end{aligned}$$

These are the analogs of the probabilities of Type I and Type II errors for the "no decision" action. We would like these probabilities to be small. Next, we consider

$$\begin{aligned} \alpha^* &:= P[\text{ reject SN or make no decision } \mid \text{ CE is true and GN is true }] \\ \pi^* &:= P[\text{ reject SN } \mid \text{ exactly one of CE and GN is true }]. \end{aligned}$$

These quantities correspond to notions of level and power, but with the sample space restricted to the subset on which CE is true or GN is true, that is, the space where a decision can be made. Thus, α^* differs from the standard notion of level, but it does capture the probability of taking an incorrect action when SN (the null) holds in the restricted sample space, i.e., when CE and GN are both true. Similarly, π^* captures the probability of taking the correct action when SN does not hold in the restricted sample space. We would like the "restricted level" α^* to be small and the "restricted power" π^* to be close to one.

WL provide useful bounds on the asymptotic properties ($T \to \infty$) of the sample-size T values of the probabilities defined above, p_T, q_T, α_T^*, and π_T^*:

Proposition 6.1 *Suppose that for $T = 1, 2, ...$ the significance levels of the CE and GN tests are α_{1T} and α_{2T}, respectively, and that $\alpha_{1T} \to \alpha_1 < .5$ and $\alpha_{2T} \to \alpha_2 < .5$. Suppose the powers of the CE and GN tests are π_{1T} and π_{2T}, respectively, and that $\pi_{1T} \to 1$ and $\pi_{2T} \to 1$. Then*

$$\begin{aligned} p_T &\to 0, \qquad \limsup q_T \leq \max\{\alpha_1, \alpha_2\}, \\ |\alpha_1 - \alpha_2| &\leq \liminf \alpha_T^* \leq \limsup \alpha_T^* \leq \alpha_1 + \alpha_2 + \min\{\alpha_1, \alpha_2\}, \qquad \textit{and} \\ \min\{1-\alpha_1, 1-\alpha_2\} &\leq \liminf \pi_T^* \leq \limsup \pi_T^* \leq \max\{1-\alpha_1, 1-\alpha_2\}. \end{aligned}$$

When $\pi_{1T} \to 1$ and $\pi_{2T} \to 1$, one can also typically ensure $\alpha_1 = 0$ and $\alpha_2 = 0$ by suitable choice of an increasing sequence of critical values. In this case, $q_T \to 0$, $\alpha_T^* \to 0$, and $\pi_T^* \to 1$. Because GN and CE tests will not be consistent against every possible alternative, weaker asymptotic bounds on the level and power of the indirect test hold for these cases by Proposition 8.1 of WL. Thus, whenever possible, one should carefully design GN and CE tests to have power against particularly important or plausible alternatives. See WL for further discussion.

6.2 Practical Tests for GN and CE

To test GN and CE, we require tests for conditional independence. Nonparametric tests for conditional independence consistent against arbitrary alternatives are readily available (e.g., Linton and Gozalo, 1997; Fernandes and Flores, 2001; Delgado and Gonzalez-Manteiga, 2001; Su and White, 2007a, 2007b, 2008; Song, 2009; Huang and White, 2009). In principle, one can apply any of these to consistently test GN and CE.

But nonparametric tests are often not practical, due to the typically modest number of time-series observations available relative to the number of relevant observable variables. In practice, researchers typically use parametric methods. These are convenient, but they may lack power against important alternatives. To provide convenient procedures for testing GN and CE with power against a wider range of alternatives, WL propose augmenting standard tests with neural network terms, motivated by the "QuickNet" procedures introduced by White (2006b) or the extreme learning machine (ELM) methods of (Huang, Zhu, and Siew, 2006). We now provide explicit practical methods for testing GN and CE for a leading class of structures obeying A.1.

6.2.1 Testing Granger Non-Causality

Standard tests for finite-order $G-$causality (e.g., Stock and Watson, 2007, p. 547) typically assume a linear regression, such as[3]

$$E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t) = \alpha_0 + \boldsymbol{Y}'_{1,t-1}\rho_0 + \boldsymbol{Y}'_{2,t-1}\beta_0 + \boldsymbol{X}'_t\beta_1.$$

3. For notational convenience, we undertand that all regressors have been recast as vectors containing the referenced elements.

For simplicity, we let $Y_{1,t}$ be a scalar here. The extension to the case of vector $Y_{1,t}$ is completely straightforward. Under the null of GN, i.e., $Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$, we have $\beta_0 = 0$. The standard procedure therefore tests $\beta_0 = 0$ in the regression equation

$$Y_{1,t} = \alpha_0 + \boldsymbol{Y}'_{1,t-1}\rho_0 + \boldsymbol{Y}'_{2,t-1}\beta_0 + \boldsymbol{X}'_t\beta_1 + \varepsilon_t. \qquad \text{(GN Test Regression 1)}$$

If we reject $\beta_0 = 0$, then we also reject GN. But if we don't reject $\beta_0 = 0$, care is needed, as not all failures of GN will be indicated by $\beta_0 \neq 0$.

Observe that when CE holds and if GN Test Regression 1 is correctly specified, i.e., the conditional expectation $E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$ is indeed linear in the conditioning variables, then β_0 represents precisely the direct structural effect of $\boldsymbol{Y}_{2,t-1}$ on $Y_{1,t}$. Thus, GN Test Regression 1 may not only permit a test of GN, but it may also provide a consistent estimate of the direct structural effect of interest.

To mitigate specification error and gain power against a wider range of alternatives, WL propose augmenting GN Test Regression 1 with neural network terms, as in White's (2006b, p. 476) QuickNet procedure. This involves testing $\beta_0 = 0$ in

$$Y_{1,t} = \alpha_0 + \boldsymbol{Y}'_{1,t-1}\rho_0 + \boldsymbol{Y}'_{2,t-1}\beta_0 + \boldsymbol{X}'_t\beta_1 + \sum_{j=1}^{r} \psi(\boldsymbol{Y}'_{1,t-1}\gamma_{1,j} + \boldsymbol{X}'_t\gamma_j)\beta_{j+1} + \varepsilon_t. \qquad \text{(GN Test Regression 2)}$$

Here, the activation function ψ is a generically comprehensively revealing (GCR) function (see Stinchcombe and White, 1998). For example, ψ can be the logistic cdf $\psi(z) = 1/(1 + \exp(-z))$ or a ridgelet function, e.g., $\psi(z) = (-z^5 + 10z^3 - 15z)\exp(-.5z^2)$ (see, for example, Candès, 1999). The integer r lies between 1 and $\bar{r}$, the maximum number of hidden units. We randomly choose (γ_{0j}, γ_j) as in White (2006b, p. 477).

Parallel to our comment above about estimating direct structural effects of interest, we note that given A.1, A.2, and some further mild regularity conditions, such effects can be identified and estimated from a neural network regression of the form

$$\begin{aligned} Y_{1,t} = {} & \alpha_0 + \boldsymbol{Y}'_{1,t-1}\rho_0 + \boldsymbol{Y}'_{2,t-1}\beta_0 + \boldsymbol{X}'_t\beta_1 \\ & + \sum_{j=1}^{r} \psi(\boldsymbol{Y}'_{1,t-1}\gamma_{1,j} + \boldsymbol{Y}'_{2,t-1}\gamma_{2,j} + \boldsymbol{X}'_t\gamma_{3,j})\beta_{j+1} + \varepsilon_t. \end{aligned}$$

Observe that this regression includes $\boldsymbol{Y}_{2,t-1}$ inside the hidden units. With r chosen sufficiently large, this permits the regression to achieve a sufficiently close approximation to $E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$ and its derivatives (see Hornik, Stinchcombe, and White (1990) and Gallant and White (1992)) that regression misspecification is not such an issue. In this case, the derivative of the estimated regression with respect to $\boldsymbol{Y}_{2,t-1}$ well approximates

$$\begin{aligned} & (\partial/\partial \boldsymbol{y}_2) E(Y_{1,t} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{Y}_{2,t-1} = \boldsymbol{y}_2, \boldsymbol{X}_t) \\ & \qquad = E[\ (\partial/\partial \boldsymbol{y}_2) q_{1,t}(\boldsymbol{Y}_{1,t-1}, \boldsymbol{y}_2, \boldsymbol{Z}_t, U_{1,t}) \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t]. \end{aligned}$$

This quantity is the *covariate conditioned expected marginal direct effect of* $\boldsymbol{Y}_{2,t-1}$ *on* $Y_{1,t}$.

Although it is possible to base a test for GN on these estimated effects, we do not propose this here, as the required analysis is much more involved than that associated with GN Test Regression 2.

Finally, to gain additional power WL propose tests using transformations of $Y_{1,t}$, $\boldsymbol{Y}_{1,t-1}$, and $\boldsymbol{Y}_{2,t-1}$, as $Y_{1,t} \perp \boldsymbol{Y}_{2,t-1} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ implies $f(Y_{1,t}) \perp g(\boldsymbol{Y}_{2,t-1}) \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ for all measurable f and g. One then tests $\beta_{1,0} = 0$ in

$$\psi_1(Y_{1,t}) = \alpha_{1,0} + \psi_2(\boldsymbol{Y}_{1,t-1})'\rho_{1,0} + \psi_3(\boldsymbol{Y}_{2,t-1})'\beta_{1,0} + \boldsymbol{X}_t'\beta_{1,1} + \sum_{j=1}^{r} \psi(\boldsymbol{Y}_{1,t-1}'\gamma_{1,1,j} + \boldsymbol{X}_t'\gamma_{1,j})\beta_{1,j+1} + \eta_t. \quad \text{(GN Test Regression 3)}$$

We take ψ_1 and the elements of the vector ψ_3 to be GCR, e.g., ridgelets or the logistic cdf. The choices of γ, r, and ψ are as described above. Here, ψ_2 can be the identity ($\psi_2(\boldsymbol{Y}_{1,t-1}) = \boldsymbol{Y}_{1,t-1}$), its elements can coincide with ψ_1, or it can be a different GCR function.

6.2.2 Testing Conditional Exogeneity

Testing conditional exogeneity requires testing A.2, i.e., $\boldsymbol{Y}_{2,t-1} \perp U_{1,t} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$. Since $U_{1,t}$ is unobservable, we cannot test this directly. But with separability (which holds under the null of direct structural non-causality), Proposition 5.9 shows that $\boldsymbol{Y}_{2,t-1} \perp U_{1,t} \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$ implies $\boldsymbol{Y}_{2,t-1} \perp \varepsilon_t \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$, where $\varepsilon_t := Y_{1,t} - E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$. With correct specification of $E(Y_{1,t}|\boldsymbol{Y}_{t-1}, \boldsymbol{X}_t)$ in either GN Test Regression 1 or 2 (or some other appropriate regression), we can estimate ε_t and use these estimates to test $\boldsymbol{Y}_{2,t-1} \perp \varepsilon_t \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$. If we reject this, then we also must reject CE. We describe the procedure in detail below.

As WL discuss, such a procedure is not "watertight," as this method may miss certain alternatives to CE. But, as it turns out, there is no completely infallible method. By offering the opportunity of falsification, this method provides crucial insurance against being naively misled into inappropriate causal inferences. See WL for further discussion.

The first step in constructing a practical test for CE is to compute estimates of ε_t, say $\hat{\varepsilon}_t$. This can be done in the obvious way by taking $\hat{\varepsilon}_t$ to be the estimated residuals from a suitable regression. For concreteness, suppose this is either GN Test Regression 1 or 2.

The next step is to use $\hat{\varepsilon}_t$ to test $\boldsymbol{Y}_{2,t-1} \perp \varepsilon_t \mid \boldsymbol{Y}_{1,t-1}, \boldsymbol{X}_t$. WL recommend doing this by estimating the following analog of GN Test Regression 3:

$$\psi_1(\hat{\varepsilon}_t) = \alpha_{2,0} + \psi_2(\boldsymbol{Y}_{1,t-1})'\rho_{2,0} + \psi_3(\boldsymbol{Y}_{2,t-1})'\beta_{2,0} + \boldsymbol{X}_t'\beta_{2,1} + \sum_{j=1}^{r} \psi(\boldsymbol{Y}_{1,t-1}'\gamma_{2,1,j} + \boldsymbol{X}_t'\gamma_{2,j})\beta_{2,j+1} + \eta_t. \quad \text{(CE Test Regression)}$$

Note that the right-hand-side regressors are identical to those of GN Test Regression 3; we just replace the dependent variable $\psi_1(Y_{1,t})$ for GN with $\psi_1(\hat{\varepsilon}_t)$ for CE. Nevertheless, the transformations ψ_1, ψ_2, and ψ_3 here may differ from those of GN Test Regression 3. To keep the notation simple, we leave these possible differences implicit. To test CE using this regression, we test the null hypothesis $\beta_{2,0} = 0$: if we reject $\beta_{2,0} = 0$, then we reject CE.

As WL explain, the fact that $\hat{\varepsilon}_t$ is obtained from a "first-stage" estimation (GN) involving potentially the same regressors as those appearing in the CE regression means that choosing $\psi_1(\hat{\varepsilon}_t) = \hat{\varepsilon}_t$ can easily lead to a test with no power. For CE, WL thus recommend

choosing ψ_1 to be GCR. Alternatively, non-GCR choices may be informative, such as

$$\psi_1(\hat{\varepsilon}_t) = |\hat{\varepsilon}_t|, \qquad \psi_1(\hat{\varepsilon}_t) = \hat{\varepsilon}_t(\lambda - 1\{\hat{\varepsilon}_t < 0\}), \quad \lambda \in (0,1), \qquad \text{or} \qquad \psi_1(\hat{\varepsilon}_t) = \hat{\varepsilon}_t^2.$$

Significantly, the asymptotic sampling distributions needed to test $\beta_{2,0} = 0$ will generally be impacted by the first-stage estimation. Handling this properly is straightforward, but somewhat involved. To describe a practical method, we denote the first-stage (GN) estimator as $\hat{\theta}_{1,T} := (\hat{\alpha}_{1,T}, \hat{\rho}_{1,T}, \hat{\beta}'_{1,0,T}, \hat{\beta}'_{1,1,T}, ..., \hat{\beta}_{1,r+1,T})'$, computed from GN Test Regression 1 ($r = 0$) or 2 ($r > 0$). Let the second stage (CE) regression estimator be $\hat{\theta}_{2,T}$; this contains the estimated coefficients for $\boldsymbol{Y}_{2,t-1}$, say $\hat{\beta}_{2,0,T}$, which carry the information about CE. Under mild conditions, a central limit theorem ensures that

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \xrightarrow{d} N(0, C_0),$$

where $\hat{\theta}_T := (\hat{\theta}'_{1,T}, \hat{\theta}'_{2,T})'$, $\theta_0 := plim(\hat{\theta}_T)$, convergence in distribution as $T \to \infty$ is denoted $\xrightarrow{d}$, and $N(0, C_0)$ denotes the multivariate normal distribution with mean zero and covariance matrix $C_0 := A_0^{-1} B_0 A_0^{-1\prime}$, where

$$A_0 := \begin{bmatrix} A_{011} & \mathbf{0} \\ A_{021} & A_{022} \end{bmatrix}$$

is a two-stage analog of the log-likelihood Hessian and B_0 is an analog of the information matrix. See White (1994, pp. 103 - 108) for specifics.[4] This fact can then be use to construct a well behaved test for $\beta_{2,0} = 0$.

Constructing this test is especially straightforward when the regression errors of the GN and CE regressions, ε_t and η_t, are suitable martingale differences. Then B_0 has the form

$$B_0 := \begin{bmatrix} E[\, \mathcal{Z}_t \, \varepsilon_t \, \varepsilon'_t \, \mathcal{Z}'_t \,] & E[\, \mathcal{Z}_t \, \varepsilon_t \, \eta'_t \, \mathcal{Z}'_t \,] \\ E[\, \mathcal{Z}_t \, \eta_t \, \varepsilon'_t \, \mathcal{Z}'_t \,] & E[\, \mathcal{Z}_t \, \eta_t \, \eta'_t \, \mathcal{Z}'_t \,] \end{bmatrix},$$

where the CE regressors $\mathcal{Z}_t$ are measurable-$\sigma(\mathcal{X}_t)$, $\mathcal{X}_t := (vec[\boldsymbol{Y}_{t-1}]', vec[\boldsymbol{X}_t]')'$, $\varepsilon_t := Y_{1,t} - E(Y_{1,t} \mid \mathcal{X}_t)$, and $\eta_t := \psi_1(\varepsilon_t) - E[\psi_1(\varepsilon_t) \mid \mathcal{X}_t]$. For this, it suffices that $U_{1,t} \perp (Y^{t-\ell-1}, X^{t-\tau_1-1}) \mid \mathcal{X}_t$, as WL show. This memory condition is often plausible, as it says that the more distant history $(Y^{t-\ell-1}, X^{t-\tau_1-1})$ is not predictive for $U_{1,t}$, given the more recent history $\mathcal{X}_t$ of $(Y^{t-1}, X^{t+\tau_2})$. Note that separability is not needed here.

The details of C_0 can be involved, especially with choices like $\psi_1(\hat{\varepsilon}_t) = |\hat{\varepsilon}_t|$. But this is a standard $m-$estimation setting, so we can avoid explicit estimation of C_0: suitable bootstrap methods deliver valid critical values, even without the martingale difference property (see, e.g., Gonçalves and White, 2004; Kiefer and Vogelsang, 2002, 2005; Politis, 2009).

An especially appealing method is the *weighted bootstrap* (Ma and Kosorok, 2005), which works under general conditions, given the martingale difference property. To implement this, for $i = 1, ..., n$ generate sequences $\{\mathcal{W}_{t,i}, t = 1, ..., T\}$ of IID positive scalar weights with $E(\mathcal{W}_{t,i}) = 1$ and $\sigma^2_{\mathcal{W}} := var(\mathcal{W}_{t,i}) = 1$. For example, take $\mathcal{W}_{t,i} \sim \chi^2_1/\sqrt{2} + (1 - 1/\sqrt{2})$, where χ^2_1 is chi-squared with one degree of freedom. The weights should be independent of

4. The regularity conditions include plausible memory and moment requirements, together with certain smoothness and other technical conditions.

the sample data and of each other. Then compute estimators $\hat{\theta}_{T,i}$ by weighted least squares applied to the GN and CE regressions using (the same) weights $\{\mathcal{W}_{t,i}, t = 1, ..., T\}$. By Ma and Kosorok (2005, theorem 2), the random variables

$$\sqrt{T}(\hat{\theta}_{T,i} - \hat{\theta}_T), \quad i = 1, ..., n$$

can then be used to form valid asymptotic critical values for testing hypotheses about θ_0.

To test CE, we test $\beta_{2,0} = 0$. This is a restriction of the form $\mathbb{S}_2\,\theta_0 = 0$, where $\mathbb{S}_2$ is the selection matrix that selects the elements $\beta_{2,0}$ from θ_0. Thus, to conduct an asymptotic level α test, we can first compute the test statistic, say

$$\mathcal{T}_T := T\,\hat{\theta}_T'\,\mathbb{S}_2'\,\mathbb{S}_2\,\hat{\theta}_T = T\,\hat{\beta}_{2,0,T}'\,\hat{\beta}_{2,0,T}.$$

We then reject CE if $\mathcal{T}_T > \hat{c}_{T,n,1-\alpha}$, where, with n chosen sufficiently large, $\hat{c}_{T,n,1-\alpha}$ is the $1-\alpha$ percentile of the weighted bootstrap statistics

$$\mathcal{T}_{T,i} := T\,(\hat{\theta}_{T,i} - \hat{\theta}_T)'\,\mathbb{S}_2'\,\mathbb{S}_2\,(\hat{\theta}_{T,i} - \hat{\theta}_T) = T\,(\hat{\beta}_{2,0,T,i} - \hat{\beta}_{2,0,T})'\,(\hat{\beta}_{2,0,T,i} - \hat{\beta}_{2,0,T}), \quad i = 1, ..., n.$$

This procedure is asymptotically valid, even though $\mathcal{T}_T$ is based on the "unstudentized" statistic $\mathbb{S}_2\,\hat{\theta}_T = \hat{\beta}_{2,0,T}$. Alternatively, one can construct a studentized statistic

$$\mathcal{T}_T^* := T\,\hat{\theta}_T'\,\mathbb{S}_2'\,[\mathbb{S}_2\,\hat{C}_{T,n}\,\mathbb{S}_2']^{-1}\,\mathbb{S}_2\,\hat{\theta}_T,$$

where $\hat{C}_{T,n}$ is an asymptotic covariance estimator constructed from $\sqrt{T}(\hat{\theta}_{T,i} - \hat{\theta}_T)$, $i = 1, ..., n$. The test rejects CE if $\mathcal{T}_T^* > c_{1-\alpha}$, where $c_{1-\alpha}$ is the $1-\alpha$ percentile of the chi-squared distribution with $\dim(\beta_{0,2})$ degrees of freedom. This method is more involved but may have better control over the level of the test. WL provide further discussion and methods.

Because the given asymptotic distribution is joint for $\hat{\theta}_{1,T}$ and $\hat{\theta}_{2,T}$, the same methods conveniently apply to testing GN, i.e., $\beta_{1,0} = \mathbb{S}_1\,\theta_0 = 0$, where $\mathbb{S}_1$ selects $\beta_{1,0}$ from θ_0. In this way, GN and CE test statistics can be constructed at the same time.

WL discuss three examples, illustrating tests for direct structural non-causality based on tests of Granger non-causality and conditional exogeneity. A matlab module, *testsn*, implementing the methods described here is available at http://ihome.ust.hk/~xunlu/code.htm.

7. Summary and Concluding Remarks

In this paper, we explore the relations between direct structural causality in the settable systems framework and direct causality in the PCM for both recursive and non-recursive systems. The close correspondence between these concepts in recursive systems and the equivalence between direct structural causality and $G-$causality established by WL enable us to show the close linkage between $G-$causality and PCM notions of direct causality. We apply WL's results to provide straightforward practical methods for testing direct causality using tests for Granger causality and conditional exogeneity.

The methods and results described here draw largely from work of WC and WL. These papers contain much additional relevant discussion and detail. WC provide further examples contrasting settable systems and the PCM. Chalak and White (2010) build on WC,

examining not only direct causality in settable systems, but also notions of indirect causality, which in turn yield implications for conditional independence relations, such as those embodied in conditional exogeneity, which plays a key role here. WL treat not only the structural VAR case analyzed here, but also the "time-series natural experiment" case, where causal effects of variables D_t, absorbed here into Z_t, are explicitly analyzed. The sequence $\{D_t\}$ represents external stimuli, not driven by $\{Y_t\}$, whose effects on $\{Y_t\}$ are of interest. For example, $\{D_t\}$ could represent passively observed visual or auditory stimuli, and $\{Y_t\}$ could represent measured neural activity. Interest may attach to which stimuli directly or indirectly affect which neurons or groups of neurons. WL also examine the structural content of classical Granger causality and a variety of related alternative versions that emerge naturally from different versions of Assumption A.1.

Acknowledgments

We express our deep appreciation to Sir Clive W.J. Granger for his encouragement of the research underlying the work presented here.

References

E. Candès. Ridgelets: Estimating with Ridge Functions. *Annals of Statistics*, 31:1561–1599, 1999.

K. Chalak and H. White. Causality, Conditional Independence, and Graphical Separation in Settable Systems. Technical report, Department of Economics, Boston College, 2010.

A.P. Dawid. Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society, Series B*, 41:1–31, 1979.

A.P. Dawid. Beware of the DAG! *Proceedings of the NIPS 2008 Workshop on Causality, Journal of Machine Learning Research Workshop and Conference Proceedings*, 6:59–86, 2010.

M.A. Delgado and W. Gonzalez-Manteiga. Significance Testing in Nonparametric Regression Based on the Bootstrap. *Annals of Statistics*, 29:1469–1507, 2001.

M. Eichler. Granger Causality and Path Diagrams for Multivariate Time Series. *Journal of Econometrics*, 137:334-353, 2007.

M. Eichler and V. Didelez. Granger-causality and the Effect of Interventions in Time Series. *Lifetime Data Analysis*, forthcoming.

R. Engle, D. Hendry, and J.F. Richard. Exogeneity. *Econometrica* 51:277–304, 1983.

M. Fernandes and R. G. Flores. Tests for Conditional Independence, Markovian Dynamics, and Noncausality. Technical report, European University Institute, 2001.

J.P. Florens and D. Fougère. Non-causality in Continuous Time. *Econometrica*, 64:1195–1212, 1996.

J.P. Florens and M. Mouchart. A Note on Non-causality. *Econometrica*, 50:583–591, 1982.

A.R. Gallant and H. White. On Learning the Derivatives of an Unknown Mapping with Multilayer Feedforward Networks. *Neural Networks*, 5:129–138, 1990.

D. Galles and J. Pearl. Axioms of Causal Relevance. *Artificial Intelligence*, 97:9–43, 1997.

R. Gibbons. Game Theory for Applied Economists. Princeton University Press, Princeton, 1992.

S. Gonçalves and H. White. Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models. *Journal of Econometrics*, 119:199–219, 2004.

C.W.J. Granger. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37:424–438, 1969.

C.W.J. Granger and P. Newbold. Forecasting Economic Time Series (2nd edition). Academic Press, New York, 1986.

J. Halpern. Axiomatizing Causal Reasoning. *Journal of Artificial Intelligence Research*, 12:317-337, 2000.

K. Hornik, M. Stinchcombe, and H. White. Universal Approximation of an Unknown Mapping and its Derivatives Using Multilayer Feedforward Networks. *Neural Networks*, 3:551–560, 1990.

G.B. Huang, Q.Y. Zhu, and C.K. Siew. Extreme Learning Machines: Theory and Applications. *Neurocomputing*, 70:489–501, 2006.

M. Huang and H. White. A Flexible Test for Conditional Independence. Technical report, Department of Economics, University of California, San Diego.

N. Kiefer and T. Vogelsang. Heteroskedasticity-autocorrelation Robust Testing Using Bandwidth Equal to Sample Size. *Econometric Theory*, 18:1350–1366, 2002.

N. Kiefer and T. Vogelsang. A New Asymptotic Theory for Heteroskedasticity-autocorrelation Robust Tests. *Econometric Theory*, 21:1130–1164, 2005.

O. Linton and P. Gozalo. Conditional Independence Restrictions: Testing and Estimation. Technical report, Cowles Foundation for Research, Yale University, 1997.

S. Ma and M. Kosorok. Robust Semiparametric M-estimation and the Weighted Bootstrap. *Journal of Multivariate Analysis,* 96:190-217, 2005.

J. Pearl. Causality. Cambridge University Press, New York, 2000.

J. Pearl. Direct and Indirect Effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, pages 411-420, 2001.

D.N. Politis. Higher-Order Accurate, Positive Semi-definite Estimation of Large-Sample Covariance and Spectral Density Matrices. Technical report, Department of Economics, University of California, San Diego, 2009.

A. Roebroeck, A.K. Seth, and P. Valdes-Sosa. Causality Analysis of Functional Magnetic Resonance Imaging Data. *Journal of Machine Learning Research*, (this issue), 2011.

K. Song. Testing Conditional Independence via Rosenblatt Transforms. *Annals of Statistics*, 37:4011-4015, 2009.

M. Stinchcombe and H. White. Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative. *Econometric Theory*, 14:295-324, 1998.

J. Stock and M. Watson. Introduction to Econometrics. Addison-Wesley, Boston, 2007.

L. Su and H. White. A Consistent Characteristic Function-Based Test for Conditional Independence. *Journal of Econometrics*, 141:807-834, 2007a.

L. Su and H. White. Testing Conditional Independence via Empirical Likelihood. Technical report, Department of Economics, University of California, San Diego, 2007b.

L. Su and H. White. A Nonparametric Hellinger Metric Test for Conditional Independence. *Econometric Theory*, 24:829–864, 2008.

H. Varian. Intermediate Microeconomics (8th edition). Norton, New York, 2009.

H. White. Estimation, Inference, and Specification Analysis. Cambridge University Press, New York, 1994.

H. White. Time-series Estimation of the Effects of Natural Experiments. *Journal of Econometrics*, 135:527-566, 2006a.

H. White. Approximate Nonlinear Forecasting Methods. In G. Elliott, C.W.J. Granger, and A. Timmermann, editors, *Handbook of Economic Forecasting*, pages 460–512, Elsevier, New York, 2006b.

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

H. White and P. Kennedy. Retrospective Estimation of Causal Effects Through Time. In J. Castle and N. Shephard editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, pages 59–87, Oxford University Press, Oxford, 2009.

H. White and X. Lu. Granger Causality and Dynamic Structural Systems. *Journal of Financial Econometrics*, 8:193-243, 2010a.

H. White and X. Lu. Causal Diagrams for Treatment Effect Estimation with Application to Selection of Efficient Covariates. Technical report, Department of Economics, University of California, San Diego, 2010b.