

Identification with Conditioning Instruments in Causal Systems

Karim Chalak¹
Department of Economics
Boston College

Halbert White
Department of Economics
University of California, San Diego

May 17, 2007

Abstract: We study the structural identification of causal effects with conditioning instruments within the setttable system framework. In particular, we provide causal and predictive conditions sufficient for conditional exogeneity to hold. We provide two procedures based on “exclusive of A ” ($\sim A$)-causality matrices and the direct causality matrix for inferring conditional causal isolation among vectors of setttable variables and consequently conditional independence among corresponding vectors of random variables. Similarly, we provide sufficient conditions for conditional stochastic isolation in terms of the σ -algebras generated by the conditioning variables. We build on these results to study the structural identification of average effects and average marginal effects with conditioning instruments. We distinguish between structural proxies and predictive proxies.

PRELIMINARY AND INCOMPLETE

Acknowledgment: The authors thank Julian Betts, Graham Elliott, Clive Granger, Mark Machina, Dimitris Politis, and Ruth Williams for their helpful comments and suggestions. All errors and omissions are the authors’ responsibility.

¹ Karim Chalak, Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467 (email: chalak@bc.edu).

Halbert White, Department of Economics 0508, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0508 (email: hwhite@ucsd.edu).

1. Introduction

Researchers are often interested in measuring causal relationships and evaluating the effects of policies and interventions. In observational studies, this is a particularly challenging task for several reasons, including, among others, the inability of the researcher to control relevant variables, the fact that some relevant causes may be unobserved, and the unknown forms of response functions. White and Chalak (2006) (WC) discuss the role that *conditional exogeneity* can play in permitting informative causal inference in such situations. We employ WC's *settable system* framework to study the specification of *proxies* that can function as *conditioning instruments*, as discussed in WC and Chalak and White (2007a), to support the *structural identification* of causal effects of interest. In particular, we build on results of Chalak and White (2007b) (CW) to provide causal and predictive conditions sufficient for conditional exogeneity to hold.

In the nomenclature of the treatment effect literature, conditional exogeneity implies the property of “ignorability” or “unconfoundedness” (see e.g. Rubin, 1974; Rosenbaum and Rubin, 1983). Nevertheless, the treatment effect literature is typically silent as to how to construct valid covariates that ensure conditional exogeneity in observational studies. In standard experimental settings, the researcher may be able to control the assignment mechanism of the treatment, collect pre-treatment measurements, and perform diagnostic tests to ensure that conditional exogeneity holds (see e.g. Rosenbaum, 2002). For example, randomization or *conditional randomization* may be sufficient for useful causal inference (see WC). In such circumstances, the determination of suitable covariates is typically straightforward. On the other hand, in observational studies and certain experimental settings, such privileges are typically not available, and the researcher has to resort to alternative means for identification of effects of interest. In particular, economic theory can play a central role in providing guidance for constructing conditioning instruments that ensure conditional exogeneity and thus support the structural identification of causal effects of interest. A main goal of the present analysis is to examine in detail how information provided by underlying economic (or other) theory can be used to generate valid conditioning instruments, thereby identifying causal effects of interest in observational settings.

We proceed with our analysis of structural identification with conditioning instruments under minimal assumptions. In econometrics, assumptions that may permit structural inference are typically concerned with the distribution of the unobserved variables or the functional form of the response functions or both (see e.g. Heckman and Robb, 1985; Heckman and Honore, 1990; Blundell and Powell, 2003; Matzkin, 2003; Heckman and Vytlacil, 2005). Here we impose only the weakest possible assumptions on the distribution of the unobserved variables and the functional form of the response functions. This extends the analysis of Chalak and White (2007a), who study in detail the identification of causal effects within the structural equations framework under the classic assumption of linearity. Their analysis reveals several extension of the concept of instrumental variables.

In brief, the main contribution of this paper is to provide formal and general causal and predictive conditions ensuring that conditional exogeneity holds, thus permitting informative causal inference. Our hope is that once these conditions have been properly justified by economic theory and expert knowledge, they can serve as helpful templates to guide empirical researchers in forming causal inferences.

The paper is organized as follows. In Section 2, we build on results of CW to study conditioning instruments that ensure conditional exogeneity in a canonical recursive settable system, \mathcal{S} . In Section 2.1, we define the *direct causality matrix* $C^{d(\mathcal{S})}$ and, for a given set of causal variables whose indexes belong to a set A , the “exclusive of A ” ($\sim A$)-*causality matrix* $C_{\sim A}^{\mathcal{S}}$. We then provide two procedures for verifying the presence of *structural proxies* using $C_{\sim A}^{\mathcal{S}}$ and $C^{d(\mathcal{S})}$. These procedures rely on the fact that conditional independence holds by conditioning on the common causes of the variables of interest or on variables that fully mediate the effects of these common causes. This includes as a special case the notion of d -separation introduced in the artificial intelligence literature (see e.g. Pearl, 2000, p.16-17). Section 2.2 studies another special case of conditioning instruments that we refer to as *predictive proxies*, following WC and Chalak and White (2007a). In this case, conditional independence can hold due to a particular predictive relationship that holds among the variables of interest. Section 3 builds on results of WC to provide fully structural conditions ensuring that conditioning instruments permit the structural identification of *covariate-conditioned average effects* and *covariate-conditioned average marginal ceteris paribus effect*. The structural identification results in Section 3 that employ structural proxies include as a special case methods based on the “back door criterion” introduced in the artificial intelligence literature by Pearl (1995, 2000). We are not aware of results that study structural identification with *predictive proxies* other than the results of WC and Chalak and White (2007a). Section 4 concludes and discusses directions for future research. Formal mathematical proofs are collected into the Mathematical Appendix.

2. Conditioning Instruments in Settable Systems

We work with the version of settable systems presented in definition 2.1 of CW. Specifically, we consider *canonical recursive* settable systems $\mathcal{S} \equiv \{(\Omega, \mathbb{F}), (Z, \Pi, r^{\Pi}, \mathcal{X})\}$ (CW definition 2.6). A settable system has a stochastic component, (Ω, \mathbb{F}) , and a structural component, $(Z, \Pi, r^{\Pi}, \mathcal{X})$. WC provide a detailed discussion. A settable system \mathcal{S} represents agents indexed by $h = 1, 2, \dots$, each governing settable variables indexed by $j = 1, 2, \dots$, represented as $\mathcal{X}_h \equiv \{\mathcal{X}_{h,j}\}$, where $\mathcal{X}_{h,j}: \{0,1\} \times \Omega \rightarrow \mathbb{R}$. A special settable variable, the *fundamental* settable variable, is denoted \mathcal{X}_0 . Settable variables $\mathcal{X} \equiv \{\mathcal{X}_h\}$ generate settings $Z = \{Z_{h,j} = \mathcal{X}_{h,j}(1, \cdot)\}$ and, for the given partition $\Pi \equiv \{\Pi_b\}$, responses $Y = \{Y_{h,j} = \mathcal{X}_{h,j}(0, \cdot)\}$, given by $Y_{h,j} = r_{h,j}^{\Pi}(Z_{(b)})$ for (h, j) in Π_b , $h \neq 0$, where $r_{h,j}^{\Pi}$ is the *response function*, and $Z_{(b)}$ denotes settings of variables not belonging to partition block

Π_b . By convention, $\Pi_0 = \{0\}$. In recursive settable systems, the blocks are ordered such that responses in higher-level blocks depend only on settings in lower-level blocks. In canonical recursive settable systems, we further have $Z_{h,j} = Y_{h,j}$ for $h \neq 0$. By convention, we also have $Z_0 = Y_0$. For simplicity, we sometimes consider *finite* settable systems, that is, systems with a finite number of agents and responses so that $j = 1, \dots, J_h$, with $J_h < \infty$, $h = 1, \dots, H < \infty$.

2.1 Causally Isolating Conditioning Instruments

Next, we provide definitions of certain matrices that summarize aspects of interest associated with settable systems. We make use of these definitions shortly for inspecting conditional independence relationships.

Following CW, we write $\mathcal{X}_{h,j} \xRightarrow{d} \mathcal{X}_{i,k}$ when $\mathcal{X}_{h,j}$ *directly causes* $\mathcal{X}_{i,k}$ in \mathcal{S} and we write $\mathcal{X}_{h,j} \not\xRightarrow{d} \mathcal{X}_{i,k}$ otherwise (see CW, definition 2.3). For a settable system \mathcal{S} , we define the direct causality matrix $C^{d(\mathcal{S})}$ associated with \mathcal{S} .

Definition 2.1: Direct Causality Matrix Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a *finite canonical recursive settable system*. The *direct causality matrix* associated with \mathcal{S} , denoted by $C^{d(\mathcal{S})}$, has elements given by:

$$c_{(h,j),(i,k)}^{d(\mathcal{S})} = 1 \quad \text{if } \mathcal{X}_{h,j} \xRightarrow{d} \mathcal{X}_{i,k},$$

$$c_{(h,j),(i,k)}^{d(\mathcal{S})} = 0 \quad \text{if } \mathcal{X}_{h,j} \not\xRightarrow{d} \mathcal{X}_{i,k}. \quad \blacksquare$$

Thus, $C^{d(\mathcal{S})}$ has the form:

$$C^{d(\mathcal{S})} = \begin{array}{c} \mathcal{X}_0(1, \cdot) \\ \mathcal{X}_{1,1}(1, \cdot) \\ \vdots \\ \mathcal{X}_{H,J_H}(1, \cdot) \end{array} \left| \begin{array}{cccc} \mathcal{X}_0(0, \cdot) & \mathcal{X}_{1,1}(0, \cdot) & \dots & \mathcal{X}_{H,J_H}(0, \cdot) \\ 0 & & & \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{array} \right|$$

We note some properties of $C^{d(\mathcal{S})}$. The diagonal entries are $c_{(h,j),(h,j)}^{d(\mathcal{S})} = 0$ for $j = 1, \dots, J_h$, $h = 1, \dots, H$, since $\mathcal{X}_{h,j} \not\xRightarrow{d} \mathcal{X}_{h,j}$ by definition. Also, the entries of the first column $c_{(h,j),0}^{d(\mathcal{S})} = 0$ for $j = 1, \dots, J_h$, $h = 1, \dots, H$, since $\mathcal{X}_{h,j} \not\xRightarrow{d} \mathcal{X}_0$ by definition. Blank entries of $C^{d(\mathcal{S})}$

can take the values 0 or 1. Let $G = (V, E)$ be the directed graph associated with \mathcal{S} , where $V = \{\mathcal{X}_{h,j} : j = 1, \dots, J_h; h = 0, 1, \dots, H\}$ is a non-empty finite set of vertices and $E \subset V \times V$ is a set of arcs such that an arc $(\mathcal{X}_{h,j}, \mathcal{X}_{i,k})$ belongs to E if and only if $\mathcal{X}_{h,j} \xrightarrow{d} \mathcal{X}_{i,k}$. Since \mathcal{S} is recursive, G admits an acyclic ordering of its vertices. Bang-Jensen and Gutin's (2001 p. 175) DFSA algorithm outputs an acyclic ordering of the vertices of G . It follows that there exists a $\sum_{h=1}^H J_h \times \sum_{h=1}^H J_h$ permutation matrix $M^{d(\mathcal{S})}$ such that $M^{d(\mathcal{S})} \times C^{d(\mathcal{S})}$ is strictly upper triangular.

The recursiveness of \mathcal{S} further imposes the following constraints on $C^{d(\mathcal{S})}$:

Proposition 2.2: Acyclicity Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a finite canonical recursive settable system and let $C^{d(\mathcal{S})}$ be the direct causality matrix associated with \mathcal{S} . Then, for all sets of n distinct elements, say $\{(h_1, j_1), \dots, (h_n, j_n)\}$, of $\{0, (1,1), \dots, (H, J_H)\}$ with $1 \leq n \leq \sum_{h=1}^H J_h$ we have:

$$c_{(h_1, j_1), (h_2, j_2)}^{d(\mathcal{S})} \times c_{(h_2, j_2), (h_3, j_3)}^{d(\mathcal{S})} \dots \times c_{(h_n, j_n), (h_1, j_1)}^{d(\mathcal{S})} = 0. \quad \blacksquare$$

Similarly, for a set A of index pairs, we define the *causality matrix exclusive of A* (or $\sim A$ -causality matrix) $C_{\sim A}^{\mathcal{S}}$ associated with a recursive settable system \mathcal{S} . Following CW, we let $\mathcal{I}_{(h,j):(i,k)}$ denote the set of $(\mathcal{X}_{h,j}, \mathcal{X}_{i,k})$ $d(\mathcal{S})$ -intercessors and we let $ind(\mathcal{I}_{(h,j):(i,k)})$ denote the set of indexes of the elements of $\mathcal{I}_{(h,j):(i,k)}$. For $A \subset ind(\mathcal{I}_{(h,j):(i,k)})$, we write $\mathcal{X}_{h,j} \xrightarrow{\sim A} \mathcal{X}_{i,k}$ when $\mathcal{X}_{h,j}$ causes $\mathcal{X}_{i,k}$ exclusive of \mathcal{X}_A (or $\mathcal{X}_{h,j} \sim A$ -causes $\mathcal{X}_{i,k}$) with respect to \mathcal{S} . We write $\Pi_0 \equiv \{0\}$.

Definition 2.2: Causality Matrix Exclusive of A ($\sim A$ -Causality Matrix) Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a finite canonical recursive settable system. For given non-negative integers b_1 and b_2 , let $(h, j) \in \Pi_{b_1}$, and let $(i, k) \in \Pi_{b_2}$. Let $A \subset \Pi \cup \Pi_0$, and let $A_{(h,j):(i,k)} = A \cap ind(\mathcal{I}_{(h,j):(i,k)})$. The $\sim A$ -causality matrix associated with \mathcal{S} denoted by $C_{\sim A}^{\mathcal{S}}$ has elements given by:

$$c_{(h,j):(i,k)}^{\mathcal{S}, \sim A} = 1 \quad \text{if } b_1 < b_2, (h, j), (i, k) \notin A, \text{ and } \mathcal{X}_{h,j} \xrightarrow{\sim A_{(h,j):(i,k)}} \mathcal{X}_{i,k},$$

$$c_{(h,j):(i,k)}^{\mathcal{S}, \sim A} = 0 \quad \text{otherwise.} \quad \blacksquare$$

$C_{\sim A}^{\mathcal{S}}$ and $C^{d(\mathcal{S})}$ share similar properties. In particular, $c_{(h,j),(h,j)}^{\mathcal{S},\sim A} = c_{(h,j),0}^{\mathcal{S},\sim A} = 0$ for $j = 1, \dots, J_h$, $h = 1, \dots, H$ by definition. Similarly, since there exists a non-recursive ordering of the sequence $\{\Pi_b\}$, it follows that there exists a $\sum_{h=1}^H J_h \times \sum_{h=1}^H J_h$ permutation matrix $M^{\mathcal{S}}$ such that $M^{\mathcal{S}} \times C_{\sim A}^{\mathcal{S}}$ is strictly upper triangular. When $A = \emptyset$, we refer to the $\sim \emptyset$ -causality matrix simply as the *causality matrix*, denoted by $C^{\mathcal{S}}$.

Let A and B be nonempty disjoint subsets of $\Pi \cup \Pi_0$, and denote by \mathcal{X}_A and \mathcal{X}_B the corresponding vectors of settable variables. Following CW, we denote by $\mathcal{I}_{A:B} \equiv \bigcup_{(h,j) \in A} \bigcup_{(i,k) \in B} \mathcal{I}_{(h,j):(i,k)} \setminus (\mathcal{X}_A \cup \mathcal{X}_B)$ the set of $(\mathcal{X}_A, \mathcal{X}_B)$ $d(\mathcal{S})$ -intercessors and we let $ind(\mathcal{I}_{A:B})$ denote the set of indexes of the elements of $\mathcal{I}_{A:B}$. Also, we adopt CW's definitions of *direct*, *indirect*, and *A-causality* for vectors of settable variables. In particular, for $C \subset ind(\mathcal{I}_{A:B})$, we write $\mathcal{X}_A \xrightarrow{\sim C} \mathcal{X}_B$ when \mathcal{X}_A causes \mathcal{X}_B exclusive of \mathcal{X}_C (i.e., $\mathcal{X}_A \sim C$ -causes \mathcal{X}_B) with respect to \mathcal{S} .

Theorem 4.6 of CW provides necessary and sufficient conditions for stochastic conditional dependence among certain random vectors in settable systems. In particular, it follows from this result that if \mathcal{X}_A and \mathcal{X}_B are *conditionally causally isolated given \mathcal{X}_C* , then their corresponding random vectors Y_A and Y_B are conditionally independent given Y_C . Our next proposition demonstrates that \mathcal{X}_C -conditional causal isolation and thus certain conditional independence relationships can be immediately verified by inspecting the $\sim C$ -causality matrix, $C_{\sim C}^{\mathcal{S}}$. For this, we let $\bar{a} = \max\{b: \text{there exists } (h,j) \in \Pi_b \cap A\}$ and $\bar{b} = \max\{b: \text{there exists } (i,k) \in \Pi_b \cap B\}$. As in CW, for any blocks a, b , $0 \leq a \leq b$, we let $\Pi_{[a:b]} \equiv \Pi_a \cup \dots \cup \Pi_{b-1} \cup \Pi_b$.

Proposition 2.3 Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}, P), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a finite canonical settable system. Let A and B be nonempty disjoint subsets of $\Pi \cup \Pi_0$, and let \mathcal{X}_A and \mathcal{X}_B be corresponding vectors of settable variables. Let $C \subset \Pi_{1[\max(\bar{a}, \bar{b})-1]} \setminus (A \cup B)$, let \mathcal{X}_C be the corresponding vector of settable variables, and let $C_{\sim C}^{\mathcal{S}}$ be the $\sim C$ -causality matrix associated with \mathcal{S} . Let $Y_A = \mathcal{X}_A(0, \cdot)$, $Y_B = \mathcal{X}_B(0, \cdot)$, and $Y_C = \mathcal{X}_C(0, \cdot)$. Then (a) \mathcal{X}_A and \mathcal{X}_B are conditionally causally isolated given \mathcal{X}_C if and only if

- (i) $0 \in A$ and $c_{0,(i,k)}^{\mathcal{S},\sim C} = 0$ for all $(i,k) \in B$; or
- (ii) $0 \in B$ and $c_{0,(h,j)}^{\mathcal{S},\sim C} = 0$ for all $(h,j) \in A$; or
- (iii) $0 \notin A \cup B$ and $c_{0,(h,j)}^{\mathcal{S},\sim C} = 0$ for all $(h,j) \in A$ or $c_{0,(i,k)}^{\mathcal{S},\sim C} = 0$ for all $(i,k) \in B$.

(b) Suppose that either (a.i), (a.ii), or (a.iii) hold; then for every probability measure P on (Ω, \mathbb{F}) , $Y_A \perp Y_B \mid Y_C$. ■

A practical benefit of this result is that it can be straightforwardly converted into a computer algorithm that can be used to verify conditional causal isolation, and thus conditional independence, for any triple Y_A, Y_B, Y_C . Note, however, that the failure to verify conditional causal isolation does not imply conditional dependence. We discuss this further below.

As discussed in CW, determining $\sim A$ -causality relationships among settable variables generally requires knowledge of the functional form of the response functions associated with the settable variables under study. In economics and other fields where observational studies predominate, it is often true that the researcher does not have detailed information about the functional form of the response functions. Frequently, the researcher may know only whether or not a given variable is a direct cause of another. Is it possible to deduce conditional independence from knowledge only of a direct causality matrix? Our next results demonstrate that the answer to this question is indeed positive.

For brevity, we may drop explicit reference to $d(\mathcal{S})$ in what follows when referring to $c_{(h,j),(i,k)}^{d(\mathcal{S})}$, as well as the matrices that obtain from $C^{d(\mathcal{S})}$ and their entries. For example, we write $c_{(h,j),(i,k)}$ to denote an element of $C^{d(\mathcal{S})}$.

Definition 2.4: Path Conditional Matrix given \mathcal{X}_A Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a finite canonical settable system. Let $A \subset \Pi \cup \Pi_0$ and let \mathcal{X}_A be the corresponding vector of settable variables. The *path conditional matrix given \mathcal{X}_A* , denoted P_A , is given by $P_A = p_A (C^{d(\mathcal{S})})$ where the elements of $p_A (\cdot)$ are defined as:

$$\begin{aligned}
 p_{(h,j),(i,k)}^A &= 1 && \text{if } (h,j), (i,k) \notin A, \text{ and } c_{(h,j),(i,k)} = 1 \text{ or there exists a subset} \\
 &&& \{(h_1, j_1), \dots, (h_n, j_n)\} \text{ of } \Pi \setminus A \text{ with } 1 \leq n \leq \sum_{h=1}^H J_h \text{ such that} \\
 &&& c_{(h,j),(h_1,j_1)} \times c_{(h_1,j_1),(h_2,j_2)} \times \dots \times c_{(h_n,j_n),(i,k)} = 1, \\
 p_{(h,j),(i,k)}^A &= 0 && \text{otherwise. } \blacksquare
 \end{aligned}$$

Definition 2.4 can be conveniently visualized using the graph G associated with \mathcal{S} . In fact, $p_{(h,j),(i,k)}^A = 1$ if and only if there exists in G an $(\mathcal{X}_{h,j}, \mathcal{X}_{i,k})$ -path of positive length that does not contain elements of A .

Our next result demonstrates that it is possible to verify certain $\sim A$ -causality relationships from knowledge of only $C^{d(S)}$, and thus P_A , without further information on the functional form of the response functions.

Lemma 2.5 Let \mathcal{S} , \mathcal{X}_A , \mathcal{X}_B and \mathcal{X}_C be as in Proposition 2.3. Let P_C be the direct path conditional matrix given \mathcal{X}_C . Let $C_{A:B} = C \cap \text{ind}(\mathcal{I}_{A:B})$. Suppose that $p_{(h,j),(i,k)}^C = 0$ for all $(h,j) \in A$ and $(i,k) \in B$. Then $\mathcal{X}_A \xrightarrow[\mathcal{S}]{\sim C_{AB}} \mathcal{X}_B$. ■

We now provide sufficient conditions for the conditional independence of certain random vectors in settable systems expressed in terms of $C^{d(S)}$.

Corollary 2.6 Let \mathcal{S} , \mathcal{X}_A , \mathcal{X}_B , \mathcal{X}_C be as Proposition 2.3, and let P_C be the path conditional matrix given \mathcal{X}_C . Let $Y_A = \mathcal{X}_A(0, \cdot)$, $Y_B = \mathcal{X}_B(0, \cdot)$, and $Y_C = \mathcal{X}_C(0, \cdot)$. Suppose that either

- (i) $0 \in A$ and $p_{0,(i,k)}^C = 0$ for all $(i,k) \in B$; or
- (ii) $0 \in B$ and $p_{0,(h,j)}^C = 0$ for all $(h,j) \in A$; or
- (iii) $0 \notin A \cup B$ and $p_{0,(h,j)}^C = 0$ for all $(h,j) \in A$ or $p_{0,(i,k)}^C = 0$ for all $(i,k) \in B$.

Then \mathcal{X}_A and \mathcal{X}_B are conditionally causally isolated given \mathcal{X}_C and for every probability measure P on (Ω, \mathbb{F}) , $Y_A \perp Y_B | Y_C$. ■

The graph G associated with a system \mathcal{S} can play a particularly helpful role in verifying conditional independence via Corollary 2.6. In fact, when $0 \notin A$, if, for all $(h,j) \in A$, there does not exist a $(\mathcal{X}_0, \mathcal{X}_{h,j})$ -path that does not contain elements of C then $Y_A \perp Y_B | Y_C$. A parallel result holds when $0 \notin B$. When further assumptions are imposed on the probability measure P , Corollary 2.6 include as a special case the notion of d -separation discussed in the artificial intelligence literature (see Pearl 2000, p. 16-17, 68-69; and CW section 5).

As is true for Proposition 2.3, Corollary 2.6 provides the basis for a straightforward computer algorithm that can be used to verify conditional causal isolation, and thus conditional independence.

Corollary 2.6 provides sufficient but not necessary conditions for conditional independence. In fact, it is possible to have $\mathcal{X}_A \xrightarrow[\mathcal{S}]{\sim C_{AB}} \mathcal{X}_B$ even when $p_{(h,j),(i,k)}^C = 1$ for all $(h,j) \in A$ and $(i,k) \in B$, due to a cancellation of direct and indirect effects. Nevertheless, such inference requires detailed knowledge of the response functions. Thus, it is possible for \mathcal{X}_A and \mathcal{X}_B to be causally isolated given \mathcal{X}_C even when conditions (i), (ii), and (iii) of

Corollary 2.6 fail. For example, when $A = \{(h, j)\}$ and $B = \{(i, k)\}$ we may have $P_{0,(h,j)}^C \times P_{0,(i,k)}^C = 1$ and $Y_{h,j} \perp Y_{i,k} \mid Y_C$. Pearl (2000, p. 48-49) and Spirtes et al (1993, p. 35, 56) introduce the assumptions of “stability” or “faithfulness” on P to ensure that the conditions in Proposition 2.3 and Corollary 2.6 deliver the same conclusion. Here, however, we see that it is the behavior of the response functions and not the properties of the probability measure that determine whether or not the conclusions of Proposition 2.3 and Corollary 2.6 coincide.

When $C = \emptyset$, we simply call P_S the path matrix associated with S and Corollary 2.6 then provides conditions ensuring the stochastic independence of Y_A and Y_B . This result is, however, less interesting since the conditions of Corollary 2.6 imply that either Y_A or Y_B (or both) are constants (see CW, lemma 3.1).

For given nonempty disjoint subsets A and B of $\Pi \cup \Pi_0$, one can ask what is the smallest set $C \subset \prod_{1 \leq i \leq \max(\bar{a}, \bar{b})-1} (A \cup B)$ (if any) such that the realizations of $Y_C = \mathcal{X}_C(0, \cdot)$ are observed and \mathcal{X}_C conditionally causally isolates \mathcal{X}_A and \mathcal{X}_B . In addition to supporting computer algorithms that can verify conditional causal isolation, Proposition 2.3 and Corollary 2.6 provide the basis for constructing algorithms that may output such a set C . We leave development of these algorithms to other work.

Thus, Proposition 2.3 and Corollary 2.6 provide causal conditions sufficient to ensure that vectors of settable variables \mathcal{X}_A and \mathcal{X}_B are causally isolated given \mathcal{X}_C and thus that $Y_A \perp Y_B \mid Y_C$. This conditional independence holds because the common causes of \mathcal{X}_A and \mathcal{X}_B , or variables that fully mediate the effects of these common causes on either (or both) \mathcal{X}_A and \mathcal{X}_B , are elements of \mathcal{X}_C (see CW, section 5). In Section 3, we discuss how such a “causally isolating” \mathcal{X}_C can play an *instrumental* role in ensuring conditional exogeneity. In this case, we refer to \mathcal{X}_C and its settings $Z_C = Y_C$ as *causally isolating conditioning instruments* or *structural proxies*.

2.2 Stochastically Isolating Conditioning Instruments

CW demonstrate that $Y_A \perp Y_B \mid Y_C$ can also hold when \mathcal{X}_A and \mathcal{X}_B are not conditionally causally isolated given \mathcal{X}_C , provided that P *stochastically isolates* \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_C . We now study special cases in which this can happen.

We build on results of Van Putten and Van Schuppen (1985), who study certain operations that leave conditional independence relationships invariant. In particular, we focus on operations that enlarge or reduce the conditioning σ -algebra and that preserve the conditional independence relationships of interest.

First, we adapt definitions from Van Putten and Van Schuppen (1985) to our context.

Definition 2.7 Projection Operator for σ -Algebras Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}, P), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a canonical recursive settable system.

- (i) Let $\mathbf{F} \equiv \{\mathbb{G} \subset \mathbb{F} \mid \mathbb{G} \text{ is a } \sigma\text{-algebra that contains all the } P\text{-null sets of } \mathbb{F}\}$.
- (ii) If $\mathbb{H}, \mathbb{G} \in \mathbf{F}$, then $\mathbb{H} \vee \mathbb{G}$ is the smallest σ -algebra in \mathbf{F} that contains \mathbb{H} and \mathbb{G} .
- (iii) For $\mathbb{G} \in \mathbf{F}$, let $L^+(\mathbb{G}) = \{g : \Omega \rightarrow \mathbb{R}^+ \mid g \text{ is } \mathbb{G}\text{-measurable}\}$.
- (iv) For $\mathbb{H}, \mathbb{G} \in \mathbf{F}$ let the *projection of \mathbb{H} on \mathbb{G}* be the σ -algebra

$$\sigma(\mathbb{H} \mid \mathbb{G}) \equiv \sigma(\{E[h \mid \mathbb{G}] \mid \text{for all } h \in L^+(\mathbb{H})\}) \in \mathbf{F},$$

with the understanding that the P -null sets of \mathbb{F} are adjoined to it. ■

The next result helps to characterize situations where P is stochastically isolating.

Theorem 2.8 Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}, P), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a canonical settable system. Let A and B be nonempty disjoint subsets of $\Pi \cup \Pi_0$, and let \mathcal{X}_A and \mathcal{X}_B be corresponding vectors of settable variables. Let C_1 and C_2 be subsets of $\prod_{t \in [\max(\bar{a}, \bar{b})-1]} \setminus (A \cup B)$ and let \mathcal{X}_{C_1} and \mathcal{X}_{C_2} denote the corresponding vectors of settable variables. Let $Y_A = \mathcal{X}_A(0, \cdot)$, $Y_B = \mathcal{X}_B(0, \cdot)$, $Y_{C_1} = \mathcal{X}_{C_1}(0, \cdot)$, and $Y_{C_2} = \mathcal{X}_{C_2}(0, \cdot)$ generate σ -algebras $\mathbb{A}, \mathbb{B}, \mathbb{C}_1$, and $\mathbb{C}_2 \in \mathbf{F}$ such that $\mathbb{C}_2 \subset \mathbb{C}_1$. Then (a.i) \mathcal{X}_{C_1} causally isolates \mathcal{X}_A and \mathcal{X}_B or (a.ii) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_1} and (b) $\sigma(\mathbb{A} \mid \mathbb{C}_1) \subset \mathbb{C}_2$ if and only if (c.i) \mathcal{X}_{C_2} causally isolates \mathcal{X}_A and \mathcal{X}_B or (c.ii) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_2} and (d) $\sigma(\mathbb{A} \mid \mathbb{B} \vee \mathbb{C}_1) \subset (\mathbb{B} \vee \mathbb{C}_2)$. ■

When Y_A, Y_B , and Y_{C_1} are as in Theorem 2.8, theorem 4.6 in CW gives that $Y_A \perp Y_B \mid Y_{C_1}$ if and only if \mathcal{X}_{C_1} causally isolates \mathcal{X}_A and \mathcal{X}_B or P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_1} . Heuristically, one can view a σ -algebra \mathbb{A} as representing information, $\mathbb{A} \vee \mathbb{B}$ as the smallest set containing the information in both \mathbb{A} and \mathbb{B} , and $\sigma(\mathbb{A} \mid \mathbb{C})$ as the information about \mathbb{A} inferred from the information in \mathbb{C} . Similarly, one can understand Y_A

$\perp Y_B | Y_C$ as the statement that knowledge of information \mathbb{C} renders information in \mathbb{A} irrelevant for information \mathbb{B} (and thus the information in \mathbb{B} irrelevant for information \mathbb{A}).

Now suppose that information \mathbb{C}_2 is a subset of information \mathbb{C}_1 . Theorem 2.8 states that (a) knowledge of \mathbb{C}_1 renders \mathbb{A} irrelevant for \mathbb{B} and (b) the information about \mathbb{A} inferred from \mathbb{C}_1 is a subset of \mathbb{C}_2 if and only if (c) knowledge of \mathbb{C}_2 renders \mathbb{A} irrelevant for \mathbb{B} and (d) the information about \mathbb{A} inferred from \mathbb{B} and \mathbb{C}_1 is contained in information \mathbb{B} and \mathbb{C}_2 . Thus, Theorem 2.8 provides necessary and sufficient conditions for preserving a conditional independence relationship among vectors of random variables Y_A and Y_B when the information that we condition on is either enlarged or reduced. Examples of this enlargement or reduction relevant here occur when we know that conditional independence holds when conditioning on unobservables, and we seek to find observables that we can condition on instead that will preserve conditional independence.

Next, we state a helpful Corollary of Theorem 2.8.

Corollary 2.9 Let \mathcal{S} , \mathcal{X}_A , \mathcal{X}_B , \mathcal{X}_{C_1} , and \mathcal{X}_{C_2} be as in Theorem 2.8. Let $Y_A = \mathcal{X}_A(0, \cdot)$, $Y_B = \mathcal{X}_B(0, \cdot)$, $Y_{C_1} = \mathcal{X}_{C_1}(0, \cdot)$, and $Y_{C_2} = \mathcal{X}_{C_2}(0, \cdot)$ generate σ -algebras \mathbb{A} , \mathbb{B} , \mathbb{C}_1 , and $\mathbb{C}_2 \in \mathbf{F}$. Suppose that (a.i) \mathcal{X}_{C_1} causally isolates \mathcal{X}_A and \mathcal{X}_B or (a.ii) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_1} and (b) $\sigma(\mathbb{A} | \mathbb{C}_1) \subset \mathbb{C}_2 \subset (\mathbb{B} \vee \mathbb{C}_1)$. Then (c.i) \mathcal{X}_{C_2} causally isolates \mathcal{X}_A and \mathcal{X}_B or (c.ii) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_2} . ■

Suppose that Y_A , Y_B , Y_{C_1} , and Y_{C_2} are as in Theorem 2.8 and that \mathcal{X}_{C_1} causally isolates \mathcal{X}_A and \mathcal{X}_B or that P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_1} , so that $Y_A \perp Y_B | Y_{C_1}$. Heuristically, Corollary 2.9 states that if (a) knowledge of \mathbb{C}_1 renders \mathbb{A} irrelevant for \mathbb{B} , (b.i) \mathbb{C}_2 contains the information about \mathbb{A} inferred from \mathbb{C}_1 , and (b.ii) \mathbb{C}_2 is contained in \mathbb{B} and \mathbb{C}_1 , then (c) knowledge of \mathbb{C}_2 renders \mathbb{A} irrelevant for \mathbb{B} . In other words, (b.i) and (b.ii) state that \mathbb{C}_2 contains the information in \mathbb{C}_1 relevant for \mathbb{A} but does not contain information not included in \mathbb{B} and \mathbb{C}_1 . Thus, Corollary 2.9 provides a bound on \mathbb{C}_2 sufficient for $Y_A \perp Y_B | Y_{C_2}$ to hold.

If the conditions of Corollary 2.9 hold and if \mathcal{X}_{C_2} does not causally isolate \mathcal{X}_A and \mathcal{X}_B , then it must be that P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_{C_2} . In this case, $Y_A \perp Y_B |$

Y_{C_2} holds, but not because the common causes of \mathcal{X}_A and \mathcal{X}_B (or variables that fully mediate the effects of these common causes on either (or both) \mathcal{X}_A and \mathcal{X}_B) are elements of \mathcal{X}_{C_2} . Instead, conditional independence holds because of a *predictive* relationship that relates Y_A , Y_B , Y_{C_1} , and Y_{C_2} , motivating our terminology designating $Z_{C_2} = Y_{C_2}$ as *predictive proxies*. Parallel to our nomenclature above, we may also call $Z_{C_2} = Y_{C_2}$ *stochastically isolating conditioning instruments*.

Observe that in such situations $Y_A \perp Y_B \mid Y_{C_2}$, but that \mathcal{X}_A and \mathcal{X}_B are not d -separated given \mathcal{X}_{C_2} in the corresponding graph G . Corollary 2.9 plays a particularly helpful role when discussing identification via predictive proxies in Section 3.

3. Identification with Conditioning Instruments

In this work, our focus is on the *total causal effects* of specific settings on a response of interest. We leave to other work the study of *direct* and *indirect* effects, as well as of more refined measures of effects, specifically effects *via* and *exclusive of* subsets of causal variables. We build on results of WC to discuss the structural identification of effects of interest. In particular, we employ the results of Section 2 to provide causal and predictive conditions sufficient to ensure that WC's *conditional exogeneity* assumption (A.2) holds. In turn, these conditions provide significant insight into the generation and construction of covariates.

The next lemma constructs the *total response function* associated with certain settings and a response of interest in canonical systems. As in CW, we write values of settings corresponding to $\Pi_{[a:b]}$ as $z_{[a:b]}$. In particular, when $(i, k) \in \Pi_{b_2}$, we can express response values for $\mathcal{X}_{i,k}$ as $r_{i,k}^\Pi(z_{[0:b_2-1]})$.

Lemma 3.1 Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}, P), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a canonical settable system. For given non-negative integers b_1 and b_2 with $b_1 < b_2$, let $A \subset \Pi_{b_1}$, let $(i, k) \in \Pi_{b_2}$, and let \mathcal{X}_A and $\mathcal{X}_{i,k}$ denote the corresponding settable variables. Let $\Pi_{[0:b_2-1](A:(i,k))} = \Pi_{[0:b_2-1]} \setminus (A \cup \text{ind}(\mathcal{I}_{A:(i,k)}))$, and denote by $\mathcal{C}_{i,k}^{d,A} \equiv \{(g, l) \in \Pi_{[0:b_2-1](A:(i,k))} : \mathcal{X}_{g,l} \xrightarrow{d} \mathcal{X}_{i,k}\}$ the set of indices in $\Pi_{[0:b_2-1](A:(i,k))}$ associated with all direct causes of $\mathcal{X}_{i,k}$. Let $\mathcal{X}_{\mathcal{C}_{i,k}^{d,A}}$ denote the vector of settable variables corresponding to $\mathcal{C}_{i,k}^{d,A}$, and let $z_{\mathcal{C}_{i,k}^{d,A}}$ denote a vector of values for settings of $\mathcal{X}_{\mathcal{C}_{i,k}^{d,A}}$. Then there exists a measurable function $r_{i,k}^A$ that we call the *total response function of $\mathcal{X}_{i,k}$ with respect to \mathcal{X}_A* such that

$$r_{i,k}^A(z_{\mathcal{C}_{i,k}^{d,A}}, z_A) = r_{i,k}^\Pi(z_{[0:b_2-1]}). \quad \blacksquare$$

The total response function of $\mathcal{X}_{i,k}$ with respect to \mathcal{X}_A represents values of a response of interest $Y_{i,k}$ as a function of the setting value z_A and values of settings corresponding to all direct causes of $\mathcal{X}_{i,k}$ that are not $(\mathcal{X}_A, \mathcal{X}_{i,k})$ intercessors. Thus, the difference $r_{i,k}^A(z_{C_{i,k}^{d,A}}, z_A^*) - r_{i,k}^A(z_{C_{i,k}^{d,A}}, z_A)$ serves as a measure of the total effect on $\mathcal{X}_{i,k}$ of an intervention $z_A \rightarrow z_A^*$ to \mathcal{X}_A , setting $Z_{C_{i,k}^{d,A}}$ to $z_{C_{i,k}^{d,A}}$.

In applications, we are often interested in comparable groups of agents indexed by $h = 1, 2, \dots$. In particular, we may be interested in measuring the effect of settings on corresponding responses associated with these agents. Assumption A.1(i) characterizes settable systems of interest to us here.

Assumption A.1(i) Let $\mathcal{S} \equiv \{(\Omega, \mathbb{F}, P), (Z, \Pi, r^\Pi, \mathcal{X})\}$ be a canonical settable system. (a) Fix indices k , and $\{j_1, j_2, \dots\}$ such that for all $h = 1, 2, \dots$, and non-negative integers b_1 and b_2 with $b_1 < b_2$, $A_h \equiv \{(h, j_1), (h, j_2), \dots\} \subset \Pi_{b_1}$ and $(h, k) \in \Pi_{b_2}$. Let $\mathcal{D}_h \equiv \mathcal{X}_{A_h}$ and $\mathcal{Y}_h \equiv \mathcal{X}_{h,k}$ denote the corresponding settable variables. (b) Suppose that for all $h = 1, 2, \dots$, $C_{h,k}^{d,A_h} \equiv \{(h, l) \in \Pi_{[0:b_2-1](A_h:(h,k))} : \mathcal{X}_{h,l} \stackrel{d}{\Rightarrow} \mathcal{X}_{h,k}\} = (C_{1h}, B_h) = \{(h, l_1), (h, l_2), \dots, (h, l'_1), (h, l'_2), \dots\}$ and denote by $(\mathcal{Z}_h, \mathcal{U}_h) \equiv (\mathcal{X}_{C_{1h}}, \mathcal{X}_{B_h}) \equiv \mathcal{X}_{C_{h,k}^{d,A_h}}$ the corresponding vector of settable variables. (c) Further, for all $h = 1, 2, \dots$, let $C_{2h} \equiv \{(h, m_1), (h, m_2), \dots\} \subset \Pi_{[0:b_2-1](A_h:(h,k))} \setminus C_{h,k}^{d,A_h}$ and let $\mathcal{W}_h \equiv \mathcal{X}_{C_{2h}}$ denote the corresponding vector of settable variables. Put $C_h \equiv (C_{1h}, C_{2h})$ and $\mathcal{X}_{C_h} \equiv (\mathcal{X}_{C_{1h}}, \mathcal{X}_{C_{2h}}) = (\mathcal{Z}_h, \mathcal{W}_h)$. ■

Thus, j_1, j_2, \dots index causes of interest and k indexes a response of interest associated with a comparable group of agents indexed by $h = 1, 2, \dots$. For given h , we denote by $\mathcal{D}_h \equiv \mathcal{X}_{A_h}$ agent h 's vector of settable variables generating causes of interest and by $\mathcal{Y}_h \equiv \mathcal{X}_{h,k}$ his settable variable generating the response of interest. Further, for all $h = 1, 2, \dots$, we let $\{(h, l_1), (h, l_2), \dots, (h, l'_1), (h, l'_2), \dots\}$ index elements of $\Pi_{[0:b_2-1](A_h:(h,k))}$ corresponding to all direct causes of $\mathcal{X}_{h,k}$ other than those generated by \mathcal{D}_h . For given h , we let $(\mathcal{Z}_h, \mathcal{U}_h) = \mathcal{X}_{C_{h,k}^{d,A_h}}$ denote the corresponding vector of settable variables. We distinguish between \mathcal{Z}_h and \mathcal{U}_h shortly.

From A.1(i), we have that $\mathcal{Z}_h \subset \Pi_{[0:b_2-1](A_h:(h,k))}$ and $\mathcal{Z}_h \stackrel{d}{\Rightarrow} \mathcal{Y}_h$. It follows that for all $h = 1, 2, \dots$, $\mathcal{D}_h \Rightarrow_{\mathcal{S}} \mathcal{Z}_h$. For all $h = 1, 2, \dots$, we treat indexes corresponding to variables that

succeed \mathcal{X}_{A_h} but are not $(\mathcal{X}_{A_h}, \mathcal{X}_{h,k})$ intercessors as elements of $\Pi_{b_2} \cup \Pi_{b_2+1} \cup \dots$ without loss of generality. Thus, we also have that for all $h = 1, 2, \dots$, $\mathcal{D}_h \Rightarrow_S \mathcal{W}_h$.

Let Y_h, D_h, Z_h, U_h , and W_h denote settings (equivalently, responses for canonical systems) of $\mathcal{Y}_h, \mathcal{D}_h, \mathcal{Z}_h, \mathcal{U}_h$, and \mathcal{W}_h respectively. By Lemma 3.1, we can write

$$Y_h \stackrel{c}{=} r_h(D_h, Z_h, U_h) \quad \text{for all } h = 1, 2, \dots,$$

where r_h denotes the unknown total response function of \mathcal{Y}_h with respect to \mathcal{D}_h .

(Following WC, we use the $\stackrel{c}{=}$ symbol instead of the usual equality sign to emphasize that this is a causal relationship.)

Since we consider a comparable group of agents indexed by h , we assume that $r_h = r$ for all $h = 1, 2, \dots$. We continue to suppress explicit mention of attributes from our analysis, keeping in mind that attributes can be incorporated into this framework as discussed in WC (section 2.6) to permit heterogeneity of responses across agents, among other things.

We assume that we only observe a sample from the population, as is often the case in economics and other fields where observational studies are of interest. Following WC, we treat sampling here as generating random positive integers say $\{H_i\}$ governed by the probability measure P . Thus, the sample responses are given by:

$$Y_{H_i} \stackrel{c}{=} r(D_{H_i}, Z_{H_i}, U_{H_i}), \quad i = 1, 2, \dots$$

Engaging in a mild abuse of notation we write $Y_i = Y_{H_i}, D_i = D_{H_i}, Z_i = Z_{H_i}, U_i = U_{H_i}$, and $W_i = W_{H_i}$. In particular, we write:

$$Y_i \stackrel{c}{=} r(D_i, Z_i, U_i), \quad i = 1, 2, \dots$$

Because agents are comparable, we drop explicit reference to the subscript i when referring to the settable variables involved and we write $(\mathcal{Y}, \mathcal{D}, \mathcal{Z}, \mathcal{U}, \mathcal{W})$, keeping in mind that these pertain to a representative agent. Following WC, we say that that \mathcal{S} generates a sample $\{(Y_i, D_i, Z_i, U_i, W_i)\}$ involving settable variables $(\mathcal{Y}, \mathcal{D}, \mathcal{Z}, \mathcal{U}, \mathcal{W})$.

We now specify how the data are generated. We let \mathbb{N} denote the natural numbers, including zero by convention, \mathbb{N}^+ the positive integers, and $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$.

Assumption A.1(ii) Let a canonical settable system \mathcal{S} generate a sample $\{(Y_i, D_i, Z_i, U_i, W_i)\}$ involving settable variables $(\mathcal{Y}, \mathcal{D}, \mathcal{Z}, \mathcal{U}, \mathcal{W})$. (a) Suppose that the joint distribution

of $(D_i, X_i) \equiv (D_i, Z_i, W_i)$ is H and the conditional distribution of U_i given $(D_i, X_i) = (d, x)$ is $G(\cdot | d, x)$ for all $i = 1, 2, \dots$, where D_i is \mathbb{R}^{k_1} -valued, $k_1 \in \mathbb{N}^+$, Z_i is \mathbb{R}^{k_2} -valued, $k_2 \in \mathbb{N}$, U_i is \mathbb{R}^{k_4} -valued, $k_4 \in \bar{\mathbb{N}}$, and W_i is \mathbb{R}^{k_3} -valued, $k_3 \in \mathbb{N}$. (b) Let the responses $\{Y_i\}$ of \mathcal{Y} be given by :

$$Y_i = \overset{c}{r}(D_i, Z_i, U_i), \quad i = 1, 2, \dots,$$

where r denotes the unknown scalar-valued total response function of \mathcal{Y} with respect to \mathcal{D} . (c) We assume that the realizations of Y_i, D_i, Z_i , and W_i are observed but that those of U_i are not. ■

The identical distribution assumption A.1(a, b) is not necessary to our analysis but significantly simplifies our notation. We employ it in what follows to drop reference to the subscript i from the sample variables. Further, we let $d, x \equiv (z, w)$, and u denote values of settings $\mathcal{D}, \mathcal{X} \equiv (\mathcal{Z}, \mathcal{W})$, and \mathcal{U} respectively.

With r the total response function of \mathcal{Y} with respect to \mathcal{D} , the total effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $Z = z$ and $U = u$, is $r(d^*, z, u) - r(d, z, u)$. Nevertheless, we cannot measure this difference, as the function r is not known. Further, even if r were known, the fact that the realizations of U are not observed prevents us from measuring the total effect on \mathcal{Y} of an intervention $d \rightarrow d^*$ to \mathcal{D} .

One accessible measurement is the conditional average response over the distribution of the unobserved variables given the observed variables. Following WC, we refer to $X = (Z, W)$ as covariates. In fact, proposition 3.1 in WC shows that when $E(Y) < \infty$, the *average response given $(D, X) = (d, x)$* exists, is finite, and for each (d, x) in $\text{supp}(D, X)$ (the support of (D, X)), it is given by

$$\mu(d, x) = \int r(d, z, u) dG(u | d, x).$$

The *average response function* μ is informative about the expected response given realizations (d, x) of (D, X) . However, without further assumptions, it does not permit measurement of the total effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} .

When $E(r(d, Z, U))$ exists and is finite for each d in $\text{supp}(D)$, WC define the *average counterfactual response at d given $X = x$* as

$$\rho(d, x) \equiv E(r(d, Z, U) | X = x) = \int r(d, z, u) dG(u | x),$$

where $G(\cdot | x)$ is the conditional distribution of U given $X = x$. WC show that under suitable assumptions, this conditional expectation has a clear counterfactual

interpretation. In particular, under assumption A.1(i.b, i.c), we have that $\mathcal{D} \Rightarrow_s \mathcal{Z}$ and $\mathcal{D} \Rightarrow_s \mathcal{W}$; thus different values of d do not necessitate different realizations of $X = (W, Z)$. This permits evaluating counterfactual quantities associated with the intervention $d \rightarrow d^*$ to \mathcal{D} . In particular, WC define the *covariate-conditioned average effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $X = x$* denoted by

$$\Delta\rho(d, d^*, x) \equiv \rho(d^*, x) - \rho(d, x)$$

and the *covariate-conditioned average marginal ceteris paribus effect on \mathcal{Y} of \mathcal{D}_j given $(D, X) = (d, x)$* as

$$\xi_j(d, x) \equiv \int D_j r(d, z, u) dG(u | x),$$

where $D_j = (\partial / \partial d_j)$ denotes the partial derivative with respect to d_j , the j th element of d .

When *conditional exogeneity* holds, that is when $D \perp U | X$, we have $dG(u | d, x) = dG(u | x)$ and thus $\Delta\rho(d, d^*, x)$ is *structurally identified* as $\Delta\rho(d, d^*, x) = \Delta\mu(d, d^*, x) \equiv \mu(d^*, x) - \mu(d, x)$. Similarly, when in addition regularity conditions permitting interchange of derivative and integral hold, $\xi_j(d, x)$ is structurally identified as $\xi_j(d, x) = D_j \rho(d, x) = D_j \mu(d, x)$.

We now examine conditions sufficient to ensure that conditional exogeneity holds. In particular, suppose that $(\mathcal{Z}, \mathcal{W})$ causally isolates \mathcal{D} and \mathcal{U} or that P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$. Then, conditional exogeneity holds, and we can thus measure the covariate conditioned average and marginal effects on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $X = x$. Theorem 3.2 builds on a result of WC (theorem 3.3) to formalize this.

Theorem 3.2 Structural Identification with Conditioning Instruments Suppose that assumptions A.1(i(a, b, c), ii(a, b)) hold and that $E(Y) < \infty$. Suppose further that (a) $(\mathcal{Z}, \mathcal{W})$ causally isolates \mathcal{D} and \mathcal{U} or (b) that P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$. Then

(i) For all $(d, x) \in \text{Supp}(D, X)$, $\Delta\rho(d, d^*, x)$ is *structurally identified* as

$$\Delta\rho(d, d^*, x) = \Delta\mu(d, d^*, x).$$

(ii) Suppose further that $D_j r(d, z, u)$ is dominated on a compact neighborhood of d by an integrable function (see assumption A.3 of WC). Then $\xi_j(d, x)$ is *structurally identified* as

$$\xi_j(d, x) = D_j \rho(d, x) = D_j \mu(d, x). \quad \blacksquare$$

When A.1(ii(c)) holds, we say that $\mu(d, x)$ is *stochastically identified*. If further (i) and (ii) in the conclusion of Theorem 3.2 hold, that is, structural identification holds, we say that $\Delta\rho(d, d^*, x)$ and $\xi_j(d, x)$ are *fully identified*.

Next, we build on the results of Section 2 to study conditions sufficient for $(\mathcal{Z}, \mathcal{W})$ to causally isolate \mathcal{D} and \mathcal{U} or for P to stochastically isolate \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$. Recall that for all $h = 1, 2, \dots$, we write $\mathcal{D}_h \equiv \mathcal{X}_{A_h}$, $\mathcal{U}_h \equiv \mathcal{X}_{B_h}$, and $(\mathcal{Z}_h, \mathcal{W}_h) = (\mathcal{X}_{C_{1h}}, \mathcal{X}_{C_{2h}}) = \mathcal{X}_{C_h}$. In particular, Corollary 3.3 demonstrates that knowledge of the $\sim C_h$ -causality matrices associated with \mathcal{S} can permit the identification of causal effects of interest.

Corollary 3.3 Structural Identification with Structural Proxies (I) Suppose that assumptions A.1(i(a, b, c), ii(a, b)) holds, that \mathcal{S} is finite, and that $E(Y) < \infty$. For $h = 1, \dots, H$, let $C_{\sim C_h}^{\mathcal{S}}$ denote the $\sim C_h$ -causality matrices associated with \mathcal{S} . Suppose that for all $h = 1, \dots, H$, either $c_{0,(h,j)}^{\mathcal{S}, \sim C_h} = 0$ for all $(h, j) \in A_h$ or $c_{0,(h,l)}^{\mathcal{S}, \sim C_h} = 0$ for all $(h, l) \in B_h$, then (i) and (ii) in the conclusion of Theorem 3.2 hold. \blacksquare

Furthermore, knowledge of the direct causality matrix $C^{d(\mathcal{S})}$ associated with \mathcal{S} alone is sufficient for ensuring the structural identification of certain effects of interest. Corollary 3.4 states this formally.

Corollary 3.4 Structural Identification with Structural Proxies (II) Suppose that assumptions A.1(i(a, b, c), ii(a, b)) holds, that \mathcal{S} is finite, and that $E(Y) < \infty$. For $h = 1, \dots, H$, let P_{C_h} be the path conditional matrix given \mathcal{X}_{C_h} associated with \mathcal{S} . Suppose that for all $h = 1, \dots, H$, either $p_{0,(h,j)}^{C_h} = 0$ for all $(h, j) \in A_h$ or $p_{0,(h,l)}^{C_h} = 0$ for all $(h, l) \in B_h$, then (i) and (ii) in the conclusion of Theorem 3.2 hold. \blacksquare

Corollary 3.4 has a convenient graphical representation. In particular, the covariate-conditioned average effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $X = x$ is structurally identified if for all $h = 1, \dots, H$, (a) there does not exist a $(\mathcal{X}_0, \mathcal{X}_{hj})$ -path that does not contain elements of C_h for all $(h, j) \in A_h$ or (b) there does not exist a $(\mathcal{X}_0, \mathcal{X}_{hl})$ -path that does not contain elements of C_h for all $(h, l) \in B_h$. Under the regularity conditions in Theorem 3.2 (ii), conditions (a) and (b) also are sufficient for the covariate-conditioned average marginal ceteris paribus effect on \mathcal{Y} of \mathcal{D}_j given $(D, X) = (d, x)$ to be

structurally identified. When further assumptions are imposed on the probability measure P (see Pearl 2000, p. 16-17, 68-69; and CW section 5), Corollary 3.4 contains the “back door criterion” (Pearl, 1995, 2000) as a special case. In addition, an application of Corollary 3.4 delivers the “front door criterion” (Pearl, 1995, 2000), where the total effect of interest is decomposed into multiple effects each identified via Corollary 3.4. Chalak and White (2007a, sections 4.1.2 and 4.2.3) discuss such identification schemes for the special case of linear structural equations systems.

We now turn our attention to the structural identification of causal effects via predictive proxies. First, we accommodate in our structure settable variables \mathcal{V}_h for $h = 1, 2, \dots$.

Assumption A.1(i) (d) For $h = 1, 2, \dots$, let $\tilde{C}_{2h} \equiv \{(h, \tilde{m}_1), (h, \tilde{m}_2), \dots\} \subset \Pi_{[0:b_2-1](A_h:(h,k))} \setminus \mathcal{C}_{h,k}^{d,A_h}$ and let $\mathcal{V}_h \equiv \mathcal{X}_{\tilde{C}_{2h}}$ denote the corresponding vector of settable variables. Put $\tilde{C}_h = (C_{1h}, \tilde{C}_{2h})$ and $\mathcal{X}_{\tilde{C}_h} = (\mathcal{X}_{C_{1h}}, \mathcal{X}_{\tilde{C}_{2h}}) = (\mathcal{Z}_h, \mathcal{V}_h)$. ■

Next, Assumption A.1(ii(d)) accommodates random variables V_i associated with \mathcal{V} in our sample.

Assumption A.1(ii) Let a canonical settable system \mathcal{S} generate a sample $\{(Y_i, D_i, Z_i, U_i, W_i, V_i)\}$ involving settable variables $(\mathcal{Y}, \mathcal{D}, \mathcal{Z}, \mathcal{U}, \mathcal{W}, \mathcal{V})$. (d) Suppose that the joint distribution of $(D_i, \tilde{X}_i) \equiv (D_i, Z_i, V_i)$ is \tilde{H} and the conditional distribution of U_i given $(D_i, \tilde{X}_i) = (d, \tilde{x})$ is $\tilde{G}(\cdot | d, \tilde{x})$ for all $i = 1, 2, \dots$, where D_i, Z_i, U_i are as in A.1(ii.a) and V_i is \mathbb{R}^{k_4} -valued, $k_4 \in \mathbb{N}$. (e) The realizations of V_i are observed, and those of Y_i, D_i, Z_i , and W_i are as in A.1(ii(c)). ■

We now state a structural identification result via predictive proxies. Observe that now we have two vectors of potential covariates: $X \equiv (Z, W)$ and $\tilde{X} \equiv (Z, V)$. Consequently, we let $\tilde{\mu}(d, \tilde{x}) \equiv \int r(d, z, u) d\tilde{G}(u | d, \tilde{x})$ denote the average response given $(D, \tilde{X}) = (d, \tilde{x})$ and we put $\Delta\tilde{\mu}(d, d^*, \tilde{x}) \equiv \tilde{\mu}(d^*, \tilde{x}) - \tilde{\mu}(d, \tilde{x})$. Similarly, we let $\tilde{\rho}(d, \tilde{x}) \equiv \int r(d, z, u) d\tilde{G}(u | \tilde{x})$ denote the average counterfactual response at d given $\tilde{X} = \tilde{x}$. Also, we let $\Delta\tilde{\rho}(d, d^*, \tilde{x}) \equiv \tilde{\rho}(d^*, \tilde{x}) - \tilde{\rho}(d, \tilde{x})$ denote the covariate-conditioned average effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $\tilde{X} = \tilde{x}$ and $\tilde{\xi}_j(d, \tilde{x}) \equiv \int D_j r(d, z, u) d\tilde{G}(u | \tilde{x})$ denote the covariate-conditioned average marginal ceteris paribus effect on \mathcal{Y} of \mathcal{D}_j given $(D, \tilde{X}) = (d, \tilde{x})$.

Corollary 3.5 Structural Identification with Predictive Proxies Suppose that assumptions A.1(i(a, b, c, d), ii(a, b, d)) hold and that $E(Y) < \infty$. Let $\mathcal{D}, \mathcal{U}, (\mathcal{Z}, \mathcal{W}), (\mathcal{Z}, \mathcal{V})$ generate σ -algebras $\mathbb{D}, \mathbb{U}, \mathbb{W}$, and $\mathbb{V} \in \mathbf{F}$. Suppose that (a.i) $(\mathcal{Z}, \mathcal{V})$ causally isolates \mathcal{D}

and \mathcal{U} or (a.ii) that P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{V})$; and (b.i) $\sigma(\mathbb{D} \mid \mathbb{V}) \subset \mathbb{W} \subset (\mathbb{U} \vee \mathbb{V})$ or (b.ii) $\sigma(\mathbb{U} \mid \mathbb{V}) \subset \mathbb{W} \subset (\mathbb{D} \vee \mathbb{V})$. Then

(i) For all $(d, x) \in \text{Supp}(D, X)$, $\Delta\tilde{\rho}(d, d^*, \tilde{x})$ is structurally identified as

$$\Delta\tilde{\rho}(d, d^*, \tilde{x}) = \Delta\tilde{\mu}(d, d^*, \tilde{x})$$

(ii) Suppose further that $D_j r(d, z, u)$ is dominated on a compact neighborhood of d by an integrable function, $\tilde{\xi}_j(d, \tilde{x})$ is *structurally identified* as

$$\tilde{\xi}_j(d, \tilde{x}) = D_j \tilde{\rho}(d, \tilde{x}) = D_j \tilde{\mu}(d, \tilde{x})$$

(iii) Conclusions (i) and (ii) in Theorem 3.2 hold. ■

Consider the case where A.1(ii(e)) fails because the realizations of V_i are not observed but where A.1(i.c) holds so that the realizations of W_i are observed. In particular, suppose that $(\mathcal{Z}, \mathcal{V})$ causally isolates \mathcal{D} and \mathcal{U} but that $(\mathcal{Z}, \mathcal{W})$ does not causally isolate \mathcal{D} and \mathcal{U} . Then it suffices that $\sigma(\mathbb{D} \mid \mathbb{V}) \subset \mathbb{W} \subset (\mathbb{U} \vee \mathbb{V})$ or $\sigma(\mathbb{U} \mid \mathbb{V}) \subset \mathbb{W} \subset (\mathbb{D} \vee \mathbb{V})$ to ensure that P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$ and thus that $\Delta\rho(d, d^*, x)$ and $\xi_j(d, x)$ are structurally identified. Heuristically, it suffices that the information in \mathbb{W} relates to the information in \mathbb{D} , \mathbb{U} , and \mathbb{V} in the sense discussed in Section 2.2. Observe that it is a predictive relationship between Y_i , D_i , Z_i , and W_i that ensures conditional exogeneity here. In particular, since $(\mathcal{Z}, \mathcal{W})$ does not causally isolate \mathcal{D} and \mathcal{U} , we have that the conditions for Pearl’s back door criterion do not hold here. One thus cannot conclude that structural identification holds on this basis. Nevertheless, Corollary 3.5 ensures that the covariate-conditioned average effect and marginal average effect on \mathcal{Y} of the intervention $d \rightarrow d^*$ to \mathcal{D} given $X = x$ are indeed structurally identified.

Alternatively, suppose that A.1(ii(e)) also holds so that the realizations of both W_i and V_i are observed. Corollary 3.5 gives that $\Delta\tilde{\rho}(d, d^*, \tilde{x})$ and $\tilde{\xi}_j(d, \tilde{x})$ are structurally identified and so are $\Delta\rho(d, d^*, x)$ and $\xi_j(d, x)$. This “over-identification” provides a basis for constructing tests for conditional exogeneity. It is also of direct interest to construct the optimal set of covariates that ensures conditional exogeneity and delivers an efficient estimator of the conditional average and marginal effects. We leave these questions to other work. Finally, we note that although Theorem 3.2 and Corollary 3.5 are concerned with average effects, our results generalize to accommodate effects on other features of the distribution of the response, such as its variance, quantiles, or its probability distribution (see WC, section 4).

4. Conclusion

We build on results of WC and CW to study the specification of *proxies* that can act as *conditioning instruments* to support the *structural identification* of causal effects of interest. In particular, we provide causal and predictive conditions sufficient for *conditional exogeneity* to hold, thus permitting structural identification of effects of interest.

We build on results in CW to provide two procedures for inferring conditional causal isolation among vectors of settable variables in canonical settable systems. Our first procedure employs the $\sim A$ -causality matrix associated with \mathcal{S} to state necessary and sufficient conditions for conditional causal isolation. Our second procedure employs the direct causality matrix $C^{d(\mathcal{S})}$ associated with \mathcal{S} to infer conditional independence relationships. The second procedure contains as a special case the notion of d -separation introduced in the artificial intelligence literature (see Pearl, 1995, 2000). It follows from CW Theorem 4.6 that these two procedures permit verifying conditional independence relationships among certain vectors of random variables in canonical settable systems.

We also build on results of van Putten and Van Schuppen (1985) to provide necessary and sufficient conditions that preserve conditional independence relationships when enlarging or reducing the σ -algebras generated by the conditioning vector of variables. A corollary of this result delivers a predictive relationship among variables of interest sufficient to ensure that P is conditionally stochastically isolating.

We build on these results to distinguish between two categories of conditioning instruments. The first renders the causes of interest and the unobserved direct causes of the response of interest conditionally causally isolated. Hence we refer to these as *causally isolating conditioning instruments* or *structural proxies*. We provide two procedures that permit verifying structural identification from the $\sim A$ -causality matrix associated with the structural proxies, as well as from the direct causality matrix $C^{d(\mathcal{S})}$. The $C^{d(\mathcal{S})}$ -based procedure contains as a special case the “back door criterion” introduced in the artificial intelligence literature (see Pearl 1995, 2000). The second category of conditioning instruments ensures that P stochastically isolates the causes of interest and the unobserved direct causes, given the vector of conditioning instruments. We refer to this second category of conditioning instruments as *stochastically isolating conditioning instruments* or *predictive proxies*. We are unaware of any treatment of this category in the literature other than in WC and Chalak and White (2007a).

Here we focus on measuring total causal effects in canonical systems. We leave to other work analysis of the measurement of *direct*, *indirect*, and the more refined notions of causality *via* and *exclusive of* sets of variables. We provide formal causal and predictive conditions sufficient to ensure that conditional exogeneity holds. We leave to other work the study of testing for conditional exogeneity, a key condition for structural identification. We also leave aside discussion of how to construct an optimal set of

conditioning instruments that support efficient estimation of causal effects of interest. The procedures of Section 2 should prove helpful in suggesting and testing for causal models. We also leave a formal treatment of this topic to other work.

5. Mathematical Appendix

Proof of Proposition 2.2 Let n be a positive integer such that $n \leq \sum_{h=1}^H J_h$ and let $\{(h_1, j_1), \dots, (h_n, j_n)\}$ be a set of n distinct elements of $\{0, (1,1), \dots, (H, J_H)\}$ such that $c_{(h_1, j_1), (h_2, j_2)}^{d(\mathcal{S})} \times c_{(h_2, j_2), (h_3, j_3)}^{d(\mathcal{S})} \dots \times c_{(h_n, j_n), (h_1, j_1)}^{d(\mathcal{S})} = 1$. From Definition 2.1, it follows that there exist a set of n distinct settable variables $\{\mathcal{X}_{h_1, j_1}, \dots, \mathcal{X}_{h_n, j_n}\}$ such that $\mathcal{X}_{h_1, j_1} \xrightarrow{d} \mathcal{X}_{h_2, j_2} \xrightarrow{d} \dots \xrightarrow{d} \mathcal{X}_{h_n, j_n} \xrightarrow{d} \mathcal{X}_{h_1, j_1}$, that is, $\{\mathcal{X}_{h_1, j_1}, \dots, \mathcal{X}_{h_n, j_n}\}$ belongs to a $d(\mathcal{S})$ -cycle. However, this yields a contradiction with the definition of recursiveness (CW, definition 2.5), thus completing the proof. ■

Proof of Proposition 2.3 (a) The proof follows from CW (definition 4.3). First, we let $\underline{a} = \min\{b: \text{there exists } (h, j) \in \Pi_b \cap A\}$ and $\underline{b} = \min\{b: \text{there exists } (i, k) \in \Pi_b \cap B\}$. (i) Suppose that $0 \in A$, it follows that $\underline{a} = 0$, and $\underline{b} \neq 0$. Suppose further that $c_{0, (i, k)}^{S, \sim C} = 0$ for all $(i, k) \in B$. Let $C_2 \equiv C \cap \text{ind}(\mathcal{I}_{\{0\}; B})$, then Definition 2.2 gives that $\mathcal{X}_0 \xrightarrow{\sim C_2} \mathcal{X}_B$ and thus \mathcal{X}_A and \mathcal{X}_B are conditionally causally isolated given \mathcal{X}_C . (ii) Suppose that $0 \in B$; it follows that $\underline{a} \neq 0$, $\underline{b} = 0$. Suppose further that $c_{0, (h, j)}^{S, \sim C} = 0$ for all $(h, j) \in A$. Let $C_1 \equiv C \cap \text{ind}(\mathcal{I}_{\{0\}; A})$; then Definition 2.2 gives that $\mathcal{X}_0 \xrightarrow{\sim C_1} \mathcal{X}_A$. Thus \mathcal{X}_A and \mathcal{X}_B are conditionally causally isolated given \mathcal{X}_C . (iii) Suppose that $0 \notin A \cup B$; it follows that $\underline{a} \neq 0$, $\underline{b} \neq 0$. Suppose further that $c_{0, (h, j)}^{S, \sim C} = 0$ for all $(h, j) \in A$ or $c_{0, (i, k)}^{S, \sim C} = 0$ for all $(i, k) \in B$; then Definition 2.2 gives that $\mathcal{X}_0 \xrightarrow{\sim C_1} \mathcal{X}_A$ or $\mathcal{X}_0 \xrightarrow{\sim C_2} \mathcal{X}_B$. Thus \mathcal{X}_A and \mathcal{X}_B are conditionally causally isolated given \mathcal{X}_C . (b) Let P be any probability measure on (Ω, \mathbb{F}) and suppose that either (a.i), (a.ii), or (a.iii) hold. It then follows from CW (theorem 4.6) that $Y_A \perp Y_B \mid Y_C$. ■

Proof of Lemma 2.5 For given $(h, j) \in A$ and $(i, k) \in B$, let $C_{(h, j); (i, k)} = C_{A; B} \cap \text{ind}(\mathcal{I}_{(h, j); (i, k)})$, and let $\mathcal{X}_{C_{(h, j); (i, k)}}$ be the corresponding vector of settable variables. Let $\mathcal{I}_{(h, j); (i, k)}^{C_{(h, j); (i, k)}}$ denote the $(\mathcal{X}_{h, j}, \mathcal{X}_{i, k})$ intercessors for paths through $\mathcal{X}_{C_{(h, j); (i, k)}}$, and let $\mathcal{X}_{\underline{C}_{(h, j); (i, k)}}$ denote the $(\mathcal{X}_{h, j}, \mathcal{X}_{i, k})$ intercessors not belonging to paths through $\mathcal{X}_{C_{(h, j); (i, k)}}$. Suppose that $p_{(h, j), (i, k)}^C = 0$. Since $C \subset \prod_{1 \leq i \leq \max(\bar{a}, \bar{b}) - 1} \setminus (A \cup B)$, we have $(h, j), (i, k) \notin C$. From Definition 2.2, it follows that $c_{(h, j), (i, k)} = 0$ and there does not exist a subset $\{(h_1, j_1), \dots,$

$(h_n, j_n)\}$ of $\Pi \setminus C$ with $1 \leq n \leq \sum_{h=1}^H J_h$ such that $c_{(h,j),(h_1,j_1)} \times c_{(h_1,j_1),(h_2,j_2)} \times \dots \times c_{(h_n,j_n),(i,k)} = 1$.

We thus have $\mathcal{I}_{(h,j):(i,k)}^{C_{(h,j):(i,k)}} = \mathcal{I}_{(h,j):(i,k)}$ and $\mathcal{X}_{\underline{C}_{(h,j):(i,k)}} = \emptyset$. Further, using CW's notation, there exist measurable functions $\tilde{r}_{C_{(h,j):(i,k):(i,k)}}^{C_{(h,j):(i,k)},(h,j)}$ and $\tilde{r}_{i,k}^{C_{(h,j):(i,k)},(h,j)}$ such that:

$$\begin{aligned} & r_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, z_{h,j}, \mathcal{Y}_{(h,j):C_{(h,j):(i,k)}}, \mathcal{Y}_{\underline{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \mathcal{Y}_{\bar{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k):(i,k)}}) \\ &= \tilde{r}_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \tilde{r}_{C_{(h,j):(i,k):(i,k)}}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}})), \end{aligned}$$

otherwise $\mathcal{X}_{h,j} \stackrel{d}{\Rightarrow} \mathcal{X}_{i,k}$ or there exists $\{\mathcal{X}_{h_1,j_1}, \dots, \mathcal{X}_{h_n,j_n}\} \subset \mathcal{I}_{(h,j):(i,k)} \setminus C$ with $1 \leq n \leq \sum_{h=1}^H J_h$ such that $\mathcal{X}_{h,j} \stackrel{d}{\Rightarrow} \mathcal{X}_{h_1,j_1} \stackrel{d}{\Rightarrow} \dots \stackrel{d}{\Rightarrow} \mathcal{X}_{h_n,j_n} \stackrel{d}{\Rightarrow} \mathcal{X}_{i,k}$.

Thus, for all (a) $z_{[0:b_1](h,j)}$; and (b) $z_{h,j}$ and $z_{h,j}^*$ with $z_{h,j} \neq z_{h,j}^*$, we have:

$$\begin{aligned} & r_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, z_{h,j}, \mathcal{Y}_{(h,j):C_{(h,j):(i,k)}}, \mathcal{Y}_{\underline{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \mathcal{Y}_{\bar{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k):(i,k)}}) \\ & \quad - r_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, z_{h,j}, \mathcal{Y}_{(h,j):C_{(h,j):(i,k)}}, \mathcal{Y}_{\underline{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \mathcal{Y}_{\bar{C}_{(h,j):(i,k)}}), \\ & \quad r_{C_{(h,j):(i,k):(i,k)}}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, z_{h,j}, \mathcal{Y}_{(h,j):C_{(h,j):(i,k)}}, \mathcal{Y}_{\underline{C}_{(h,j):(i,k)}}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \mathcal{Y}_{\bar{C}_{(h,j):(i,k)}})) \\ &= \tilde{r}_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \tilde{r}_{C_{(h,j):(i,k):(i,k)}}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}})) \\ & \quad - \tilde{r}_{i,k}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}}, \tilde{r}_{C_{(h,j):(i,k):(i,k)}}^{C_{(h,j):(i,k)},(h,j)} (z_{[0:b_1](h,j)}, \mathcal{Y}_{C_{(h,j):(i,k)}})) \\ &= 0. \end{aligned}$$

It follows from definition 2.8(II) in CW that $\mathcal{X}_{h,j} \stackrel{\sim C_{(h,j):(i,k)}}{\Rightarrow}_S \mathcal{X}_{i,k}$. Since $\in A$ and $(i, k) \in B$ are arbitrary, we have $\mathcal{X}_{h,j} \stackrel{C_{(h,j):(i,k)}}{\Rightarrow}_S \mathcal{X}_{i,k}$ for all $(h, j) \in A$ and $(i, k) \in B$, i.e. $\mathcal{X}_A \stackrel{\sim C_{AB}}{\Rightarrow}_S \mathcal{X}_B$. ■

Proof of Corollary 2.6 Suppose that either (i), (ii), or (iii) hold. From Lemma 2.5 and Definition 2.2, we then have that conditions (i), (ii) or (iii) of Proposition 2.3 hold. The proof then follows from Proposition 2.3(b). ■

Proof of Theorem 2.8 Theorem 4.6 in CW gives that (a) \mathcal{X}_C causally isolates \mathcal{X}_A and \mathcal{X}_B or (b) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_C if and only if $Y_A \perp Y_B \mid Y_C$, that is, \mathbb{A} is conditionally independent of \mathbb{B} given \mathbb{C} . The result is then immediate from theorem 3.3 in Van Putten and Van Schuppen (1985). ■

Proof of Corollary 2.9 Theorem 4.6 in CW gives that (a) \mathcal{X}_C causally isolates \mathcal{X}_A and \mathcal{X}_B or (b) P stochastically isolates \mathcal{X}_A and \mathcal{X}_B given \mathcal{X}_C if and only if $Y_A \perp Y_B \mid Y_C$, that is, \mathbb{A} is conditionally independent of \mathbb{B} given \mathbb{C} . The result is then immediate from proposition 3.4(e) in Van Putten and Van Schuppen (1985). ■

Proof of Lemma 3.1 Let $B = \text{ind}(\mathcal{I}_{A:\{(i,k)\}})$, we permute the arguments of $r_{i,k}^\Pi$ to write $r_{i,k}^{B,A}(z_{[0:b_2-1](A:(i,k))}, z_A, z_B) \equiv r_{i,k}^\Pi(z_{[0:b_2-1]})$. Substituting for $z_B = y_B = r_B^{B,A}(z_{[0:b_2-1](A:(i,k))}, z_A)$, we have $r_{i,k}^{B,A}(z_{[0:b_2-1](A:(i,k))}, z_A, z_B) = r_{i,k}^{B,A}(z_{[0:b_2-1](A:(i,k))}, z_A, r_B^{B,A}(z_{[0:b_2-1](A:(i,k))}, z_A)) = \tilde{r}_{i,k}^A(z_{[0:b_2-1](A:(i,k))}, z_A)$. Fix $(g, l) \in \Pi_{[0:b_2-1](A:(i,k))}$ such that $\mathcal{X}_{g,l} \xrightarrow[d]{S} \mathcal{X}_{i,k}$. Let $z_{[0:b_2-1](A:(i,k))(g,l)}$ denote values for settings corresponding to all elements of $\Pi_{[0:b_2-1](A:(i,k))(g,l)} = \Pi_{[0:b_2-1](A:(i,k))} \setminus \{(g, l)\}$. Then the function $\tilde{r}_{i,k}^A(z_{[0:b_2-1](A:(i,k))}, z_A)$ is constant in $z_{g,l}$ for every $(z_{[0:b_2-1](A:(i,k))(g,l)}, z_A)$. Thus there exists a measurable function $\tilde{r}_{i,k}^{A,(g,l)}$ such that $\tilde{r}_{i,k}^{A,(g,l)}(z_{[0:b_2-1](A:(i,k))(g,l)}, z_A) = \tilde{r}_{i,k}^A(z_{[0:b_2-1](A:(i,k))}, z_A)$ for all $z_{[0:b_2-1](A:(i,k))(g,l)}$ and z_A . Since $(g, l) \in \Pi_{[0:b_2-1](A:(i,k))}$ is arbitrary, it follows that there exists a measurable function $r_{i,k}^A$ such that $r_{i,k}^A(z_{C_{i,k}^{d,A}}, z_A) = \tilde{r}_{i,k}^A(z_{[0:b_2-1](A:(i,k))}, z_A)$ for all $z_{[0:b_2-1](A:(i,k))(C_{i,k}^{d,A})}$, $z_{C_{i,k}^{d,A}}$, and z_A using the obvious notation for $z_{[0:b_2-1](A:(i,k))(C_{i,k}^{d,A})}$. ■

Proof of Theorem 3.2 The proof follows from theorem 3.3 in WC. (i) Suppose that A.1(i(a, b, c), ii(a, b)) hold; then assumptions A.1(a, b, c.i) in WC hold. In particular, under assumptions A.1(i(b, c)), we have that $\mathcal{D} \xrightarrow[S]{} \mathcal{Z}$ and $\mathcal{D} \xrightarrow[S]{} \mathcal{W}$. Also, from Theorem 4.6 in CW we have that (a) $(\mathcal{Z}, \mathcal{W})$ causally isolates \mathcal{D} and \mathcal{U} or (b) P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$, if and only if $U \perp D \mid X$. Hence assumption A.2 in WC holds. Since $E(Y) < \infty$, the result follows from WC (theorem 3.3 (i)). (ii) Suppose further that $D_j r(d, z, u)$ is dominated on a compact neighborhood of d by an integrable function, then assumption A.3 in WC holds and the result follows from WC (theorem 3.3 (ii)). ■

Proof of Corollary 3.3 Since $0 \notin A_h \cup B_h$ for all $h = 1, \dots, H$, Proposition 2.3 gives that \mathcal{D}_h and \mathcal{U}_h are conditionally causally isolated given $(\mathcal{Z}_h, \mathcal{W}_h)$ and thus $D \perp U \mid X$. The result then follows from Theorem 3.2. ■

Proof of Corollary 3.4 Since $0 \notin A_h \cup B_h$, for all $h = 1, \dots, H$, Corollary 2.6 gives that \mathcal{D}_h and \mathcal{U}_h are conditionally causally isolated given $(\mathcal{Z}_h, \mathcal{W}_h)$ and thus $D \perp U \mid X$. The result then follows from Theorem 3.2. ■

Proof of Corollary 3.5 Suppose that (a.i) or (a.ii); and (b.i) or (b.ii) hold. Then, from Corollary 2.9, we have that $(\mathcal{Z}, \mathcal{W})$ causally isolates \mathcal{D} and \mathcal{U} or that P stochastically isolates \mathcal{D} and \mathcal{U} given $(\mathcal{Z}, \mathcal{W})$. The result then follows from Theorem 3.2. ■

References

Bang-Jensen, J. and G. Gutin (2001). *Digraphs: Theory, Algorithms and Applications*. London: Springer-Verlag.

Blundell, R. and J. Powell (2003), “Endogeneity in Nonparametric and Semiparametric Regression Models,” in M. Dewatripoint, L. Hansen, and S. Turnovsky (eds.), *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol II. New York: Cambridge University Press, 312-357.

Chalakov, K., and H. White (2007a), “An Extended Class of Instrumental Variables for the Estimation of Causal Effects,” UCSD Department of Economics Discussion Paper.

Chalakov, K. and H. White (2007b), “Independence and Conditional Independence in Causal Systems,” UCSD Department of Economics Discussion Paper.

Heckman, J. and R. Robb (1985), “Alternative Methods for Evaluating the Impact of Interventions,” in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*. Cambridge: Cambridge University Press, 146-245.

Heckman, J. and B. Honore (1990), “The Empirical Content of the Roy Model,” *Econometrica*, 58, 1121-1149.

Heckman, J. and E. Vytlacil (2005), “Structural Equations, Treatment Effects, and Econometric Policy Evaluation,” *Econometrica*, 73, 669-738.

Matzkin, R. (2003), “Nonparametric Estimation of Nonadditive Random Functions,” *Econometrica*, 71, 1339-1375.

Pearl, J. (1995), “Causal Diagrams for Empirical Research,” *Biometrika*, 82, 669-710 (with Discussion).

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press.

Rosenbaum, P. R. (2002). *Observational Studies*. 2nd ed., Berlin: Springer-Verlag.

Rosenbaum, P. R. and D. Rubin (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70, 41-55.

Rubin, D. (1974), "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies," *Journal of Educational Psychology*, 66, 688-701.

Van Putten, C. and J.H. Van Schuppen, (1985), "Invariance Properties of the Conditional Independence Relation," *The Annals of Probability*, 13, 934-945.

White, H. and K. Chalak (2006), "A Unified Framework for Defining and Identifying Causal Effects," UCSD Department of Economics Discussion Paper.