Variable Definition and Independent Components

Lorenzo Casini, Alessio Moneta, and Marco Capasso*

In the causal modeling literature, it is well known that ill-defined variables may give rise to ambiguous manipulations. Here, we illustrate how ill-defined variables may also induce mistakes in causal inference when standard causal search methods are applied. To address the problem, we introduce a representation framework, which exploits an independent component representation of the data, and demonstrate its potential for detecting ill-defined variables and avoiding mistaken causal inferences.

1. The Problem of Variable Definition. Some choices of variables may lead to less informative, or even false, causal claims (Spirtes and Scheines 2004; Eberhardt 2016; Woodward 2016). Here is a classic example by Spirtes and Scheines (2004). Consider the following hypothetical data-generating process (fig. 1). Total cholesterol (TC) is a deterministic function (e.g., the sum) of two variables, that is, low-density lipoproteins (LDL) and high-density lipoproteins (HDL), respectively known as "bad" and "good" cholesterol. The two cholesterols, in fact, have different causal roles: LDL causes heart disease (HD), while HDL prevents it. Moreover, assume that HDL and LDL cause, respectively, a disease called "disease 1" (D1) and a disease called "disease 2" (D2). Spirtes and Scheines point out that, if only TC but neither HDL nor LDL is observed, a manipulation of TC with respect to HD is "ambiguous" because it leaves underdetermined the values of TC's underlying determinants, such that the effect on HD is unpredictable.

In applied causal inference, often the variables under study are, like TC, functions of other variables with heterogeneous causal roles. For example,

*To contact the authors, please write to: Lorenzo Casini, Department of Philosophy, University of Turin, Turin, Italy; e-mail: lorenzodotcasini@gmail.com. Alessio Moneta, Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy; e-mail: alessio.moneta@santannapisa.it. Marco Capasso, Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway; e-mail: marco.capasso@nifu.no.

Philosophy of Science, 88 (December 2021) pp. 784–795. 0031-8248/2021/8805-0003$10.00

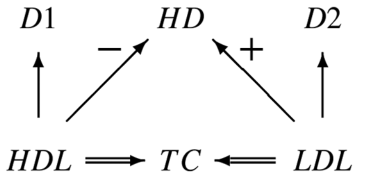

Figure 1. Structure where the manipulation on TC with respect to HD is ambiguous.

macroeconomics deals with aggregate variables such as gross domestic product, foreign sales, total imports, and so on, that are sums or averages of other variables, whose individual causal roles may be multifarious and opaque to the researcher. Often, the researcher cannot observe the underlying microbehaviors simply because statistical agencies provide aggregate data but do not reveal information on the single units. In other cases, collecting microdata may be too complex or costly. Treating aggregate variables as if they had a homogeneous causal role, however, may lead to less informative or false causal claims, as shown by the TC example. We will refer to an aggregate variable incurring such problems as *ill-defined*. Notice, thus, that whether a variable is ill-defined is relative to a variable set. That is, it may be ill-defined in one set but well-defined in another.

The problem of variable definition is often underestimated by the wider public. For instance, insufficient attention has been paid to its consequences for causal inference by constraint-based discovery methods (Spirtes, Glymour, and Scheines 2000; Pearl 2009). We will return to this point in the next section, by showing how the presence of TC in a variable set may lead to wrong causal inferences. To address the problem, we introduce a representation framework—the "independent component" representation—for modeling structures containing two kinds of dependencies, namely, traditional causal dependencies between well-defined variables and dependencies between ill-defined variables and their determinants (see fig. 1). Finally, we demonstrate the potential of this framework for identifying ill-defined variables and reducing the risk of mistaken causal inferences.

2. Causal Search with Ill-Defined Variables. The last decades have witnessed the development and popularization of constraint-based causal discovery methods (Spirtes et al. 2000; Pearl 2009). In this framework, a causal structure is represented as a triple $\langle \mathbf{V}, \mathcal{E}, \mathrm{Pr} \rangle$, where $\langle \mathbf{V}, \mathcal{E} \rangle$ is a directed acyclic graph (DAG) consisting of a set $\mathbf{V}$ of variables and a set $\mathcal{E}$ of edges among them, and Pr is the probability distribution over $\mathbf{V}$ associated to the DAG. The probability distribution Pr is assumed to comply with the Causal Markov Condition (CMC) and, typically, the Causal Faithfulness Condition (CFC). CMC says that

(CMC) For any $V_i \in \mathbf{V} = \{V_1, \dots, V_n\}$, $V_i \perp\!\!\!\perp \mathbf{Non_i} | \mathbf{Par_i}$.

Here $\mathbf{Par_i}$ denotes the set of parents (direct causes) of V_i, and $\mathbf{Non_i}$ denotes the set of nondescendants (noneffects) of V_i. In other words, each variable is probabilistically independent of its noneffects, conditional on its direct causes. CMC presupposes that for every pair of variables in **V**, every common direct cause of the pair is in **V** or has the same value for all units in the population (causal sufficiency). CFC says:

(CFC) $\langle \mathbf{V}, \mathcal{E}, \mathrm{Pr} \rangle$ is such that every conditional independence relation true in Pr is entailed by CMC applied to the true DAG $\langle \mathbf{V}, \mathcal{E} \rangle$.

CFC ensures that there is no causal dependence without probabilistic dependence; that is, all probabilistic independencies in the DAG correspond to causal independencies.

Based on these assumptions, constraint-based discovery methods are designed to recover the causal structure from data, by identifying conditional independencies among variables and then causally connecting variables not found to be independent. We now consider examples of simple data-generating processes including one ill-defined variable, TC, and show how using constraint-based methods based on conditional independencies—while ignoring that TC is ill-defined—may lead to mistakes. To anticipate, such mistakes involve apparent violations of CMC or CFC, which the search methods presuppose. Notice, however, that our interest here is not in providing novel counterexamples to CMC and CFC. These violations, in fact, could be avoided by choosing a "more suitable" variable set for causal inference—in this case, one featuring HDL and LDL instead of TC. And indeed, a formulation of CMC requiring that variables be independent of their noneffects conditional on their well-defined direct causes would not incur any violation. Here, however, we do not want to presuppose what counts as an ill-defined variable or a suitable variable set. Our goal is to avoid mistaken causal inferences in virtue of detecting ill-defined variables.

Suppose that, in $\mathbf{V} = \{X, Y, Z\}$, Y is the nondeterministic cause of both X and Z; that is, the true structure is $X \leftarrow Y \rightarrow Z$. If all variables are well-defined, one can infer some properties of the causal structure by testing conditional independencies and applying a constraint-based discovery method. In particular, the independence $X \perp\!\!\!\perp Z | Y$ and CFC allow one to exclude $X \rightarrow Y \leftarrow Z$ from the set of possible structures. Now, let the set of observed variables be $\mathbf{V}' = \{\mathrm{TC}, \mathrm{D1}, \mathrm{D2}\}$. That is, suppose again that one does not observe LDL and HDL but only TC. In this case, too, the true structure is not a collider. Assuming that the dependencies over $\mathbf{V}'$ are causally interpretable, the most plausible structure—the one we wish to rationalize in this article—would be a common cause (i.e., D1 $\leftarrow$ TC $\rightarrow$ D2). However, since

HDL and LDL are independent, LDL $\perp\!\!\!\perp$ HDL, it follows that D1 and D2 are independent, too (i.e., D1 $\perp\!\!\!\perp$ D2). If the true structure is a common cause, this contradicts CFC, which would entail a dependence between the effects of the common cause. Moreover, because D1 and D2 are dependent on (respectively) LDL and HDL, D1 and D2 become dependent upon conditioning on TC (i.e., D1 $\not\!\perp\!\!\!\perp$ D2|TC). For example, suppose one knows that one patient's total cholesterol has increased. Then, knowing that disease 1 is absent gives one relevant information to predict that disease 2 is present. If the true structure is a common cause, this conditional dependence would violate CMC, which would entail the independence of D1 and D2 given their common cause. Based on D1 $\perp\!\!\!\perp$ D2 and D1 $\not\!\perp\!\!\!\perp$ D2|TC, as well as TC $\not\!\perp\!\!\!\perp$ D1 and TC $\not\!\perp\!\!\!\perp$ D2, a constraint-based algorithm (e.g., the algorithms proposed by Spirtes et al. 2000, 84–85, 144–45) will infer an unshielded collider on TC (i.e., D1 → TC ← D2). A researcher applying the algorithm without knowing that TC is the sum of HDL and LDL (which are causes of, respectively, D1 and D2) will thus infer the wrong structure. The reason, ultimately, is that TC is ill-defined in $\mathbf{V}'$.

Similarly, assume that all variables in **V** are well-defined, but now *X* causes *Y*, and *Y* causes *Z*; that is, the true structure is $X \rightarrow Y \rightarrow Z$. Under CMC, it holds that $Z \perp\!\!\!\perp X|Y$, and under CFC, it holds that $X \not\!\perp\!\!\!\perp Z$. Now, consider the set of observed variables $\mathbf{V}'' = \{\text{Da}, \text{TC}, \text{D1}\}$, where Da (not represented in fig. 1), denoting dairies, is a cause of LDL but not of HDL. Again, suppose that one observes TC but neither HDL nor LDL. Here, too, the true structure is not a collider. The most plausible causal interpretation of the dependencies over $\mathbf{V}''$ is a directed path (i.e., Da → TC → D1). However, since Da is a cause of LDL, which is independent of the cause HDL of D1, it holds that Da $\perp\!\!\!\perp$ D1, which violates CFC. Moreover, it holds that Da $\not\!\perp\!\!\!\perp$ D1|TC, which violates CMC. From this, one may again wrongly infer a collider on TC (i.e., Da → TC ← D1). Ultimately, the reason is that TC is ill-defined in $\mathbf{V}''$.

These simple examples show how conditional independencies are sensitive to the presence of ill-defined variables in fork and chain structures, but ill-defined variables are undetectable from conditional independencies only.[1] This may lead to mistaken inferences (i.e., the inference of colliders) if one unreflectively applies constraint-based algorithms.

3. A Novel Representation Framework. We now introduce a series of definitions, which will allow us to precisely define the notion of an ill-defined variable. First, we introduce a class of data-generating mechanisms inducing the problem of ill-defined variables. We call them "augmented" structural

1. By contrast, no mistake occurs if TC is truly a collider. For instance, the inferred structure over $\mathbf{V}''' = \{\text{Da}, \text{TC}, \text{Ol}\}$, where Ol (olive oil) causes HDL but not LDL, is Da → TC ← Ol, as it should be.

causal models, by which we extend the traditional notion of structural causal models (Pearl 2009; Peters, Janzing, and Schölkopf 2017) to structures including deterministic assignments.

Augmented structural causal model An augmented structural causal model, $\mathfrak{C} \coloneqq (\mathbf{A}_W, \mathbf{A}_I, \Pr)$ consists of a collection $\mathbf{A}_W$ of m assignments, a collection $\mathbf{A}_I$ of k assignments, and a probability distribution Pr such that:

i) the collection of $\mathbf{A}_W$ consists of assignments

$$W_i \coloneqq f_i(\mathbf{Par_i}, S_i), \quad \text{for } i = 1, \dots, m,$$

where $\mathbf{Par_i} \subseteq \mathbf{W} \backslash \{W_i\}$ are called the parents of W_i, and S_i are called *noises*, or *shocks*;

ii) Pr over $\mathbf{S} = \{S_1, \dots, S_m\}$ is such that the shocks are mutually independent; that is, $\Pr(\mathbf{S}) = \Pr(S_1) \cdot \ldots \cdot \Pr(S_m)$. Hence, the S_i are also called *independent components*;

iii) the collection of $\mathbf{A}_I$ consists of assignments

$$I_i \coloneqq f_i(\mathbf{Det_i}), \quad \text{for } i = 1, \dots, k,$$

where $\mathbf{Det_i} \subset \mathbf{V}$ are called *determinants* of I_i.

Model $\mathfrak{C}$ is defined over a set of variables $\mathbf{V} = \mathbf{W} \cup \mathbf{I}$ with cardinality $n = m + k$. We associate to $\mathfrak{C}$ a graph $\mathcal{G}_V$ (see, e.g., fig. 1, where TC is the only variable with a deterministic assignment). The graph $\mathcal{G}_V$ is obtained by creating a node for each element of $\mathbf{V}$ and by drawing a directed edge $\rightarrow$ from each parent in $\mathbf{Par_i}$ (if not empty) to W_i and a modified directed edge $\Rightarrow$ from each determinant in $\mathbf{Det_i}$ to I_i. Henceforth, we restrict our attention to acyclic structures such that $\mathcal{G}_V$ is a modified DAG, to cases where $\mathbf{Det_i}$ has at least two elements, and to assignments $\mathbf{A}_W$ in which the shocks are additive. For simplicity, we also assume that no pair of variables I_i, I_j in $\mathbf{I}$, $I_i \neq I_j$, are linked in $\mathcal{G}_V$ by a bidirected modified "active" (i.e., without colliders) path $I_i \Leftarrow \cdots \Rightarrow I_j$.

By replacing all modified directed edges $\Rightarrow$ with standard edges $\rightarrow$, $\mathcal{G}_V$ becomes a standard DAG, labeled $\tilde{\mathcal{G}}_V$. By removing from $\mathcal{G}_V$ the nodes in $\mathbf{I}$ and the edges connecting $\mathbf{I}$ to $\mathbf{W}$, we obtain a subgraph of $\mathcal{G}_V$, which we denote $\mathcal{G}_W$.

Let us now introduce a particular graph associated with $\mathfrak{C}$, which we call independent component (IC) representation, or $\mathcal{G}^{\mathrm{IC}}$. The graph $\mathcal{G}^{\mathrm{IC}}$ contains edges between shocks and endogenous variables but not among endogenous variables themselves. Despite this apparent limitation, the information in $\mathcal{G}^{\mathrm{IC}}$ will be key to the purpose of our article. Although here we are not concerned with how $\mathcal{G}^{\mathrm{IC}}$ is recovered, we should mention that there exist powerful statistical learning techniques, such as Independent Component Analysis (ICA; Hyvärinen, Karhunen, and Oja 2001), that under certain assumptions (i.e., non-Gaussianity) infer the dependence coefficients, and thus identify the absence of dependencies, between shocks and endogenous variables in $\mathfrak{C}$, and thereby recover the edges in $\mathcal{G}^{\text{IC}}$.

IC representation Consider $\mathfrak{C} \coloneqq (\mathbf{A}_W, \mathbf{A}_I, \Pr)$, with $\mathbf{V} = \mathbf{W} \cup \mathbf{I}$, card($\mathbf{V}$) $= n = m + k$. An IC representation of $\mathfrak{C}$ is a DAG $\mathcal{G}_V^{\text{IC}} = \langle \mathbf{V} \cup \mathbf{S}, \mathcal{E}^{\text{IC}} \rangle$ such that $\mathcal{E}^{\text{IC}}$ consists of the following edges:

i) $S_i \to W_i$, for any $i = 1, \ldots, m$;
ii) $S_i \to W_j$, for any $i \neq j$ such that there is a directed standard path $W_i \to \cdots \to W_j$ in $\mathcal{G}_V$;
iii) $S_i \to I_h$, for any $S_i \in \mathbf{S}$ and any $I_h \in \mathbf{I}$ such that there is a directed modified path $W_i \Rightarrow \cdots \Rightarrow I_h$ in $\mathcal{G}_V$;
iv) $S_i \to I_h$, for any $S_i \in \mathbf{S}$ and any $I_h \in \mathbf{I}$ such that from W_i to I_h in $\mathcal{G}_V$ there is a directed standard path followed by a directed modified path with the same orientation, $W_i \to \cdots \Rightarrow I_h$.

Let us illustrate this definition relative to figure 2, where $\mathbf{W} = \{\text{HDL, LDL, D1, D2, HD}\}$ and $\mathbf{I} = \{\text{TC}\}$. (i) There is a shock for each variable in $\mathbf{W}$. Some shocks (e.g., S_{D1}) only hit one variable (D1). Others are common to multiple variables. (ii) For any variable (e.g., HDL), its shock (S_{HDL}) also hits all of its descendants, if any (D1, HD). (iii) Any shock to a determinant of a variable I_i in $\mathbf{I}$ (e.g., S_{HDL}) also hits I_i (TC). (iv) If $\mathbf{V}$ contained a cause of a determinant of I_i (e.g., dairies, Da, which causes LDL), its shock (S_{Da}) would also hit I_i (TC).

One may also define $\mathcal{G}^{\text{IC}}$ relative to any subset $\mathbf{O}$ of variables in $\mathbf{V}$, namely, $\mathcal{G}_O^{\text{IC}} = \langle \mathbf{O} \cup \mathbf{S}_O, \mathcal{E}_O^{\text{IC}} \rangle$. The set $\mathbf{S}_O$ is obtained by removing from $\mathbf{S}$ those shocks, which $\mathfrak{C}$ assigns to variables in $\mathbf{W}$ that are not in $\mathbf{O}$, and by adding those shocks, which $\mathfrak{C}$ assigns to variables in $\mathbf{W}$ that are determinants of variables in $\mathbf{I} \cap \mathbf{O}$. The set $\mathcal{E}_O^{\text{IC}}$ is obtained by removing from $\mathcal{E}^{\text{IC}}$ all of those edges whose tails are not in $\mathbf{S}_O$. For any variable set $\mathbf{O}$, we call "idiosyncratic" a shock to a variable X in $\mathcal{G}_{\mathbf{O}}^{\text{IC}}$ that is a parent of X and of no other variable. We may now define ill- and well-defined variables:

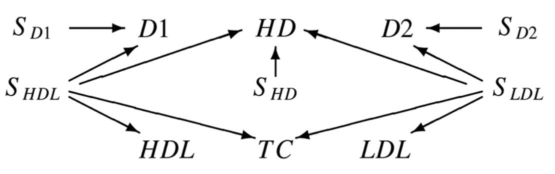

Figure 2. $\mathcal{G}_V^{\text{IC}}$ corresponding to the DAG $\mathcal{G}_V$ in figure 1.

Ill- and well-defined variables Let $\mathfrak{C}$ over $\mathbf{V} = \mathbf{W} \cup \mathbf{I}$ contain the assignment $I \coloneqq f(\mathbf{Det}_I)$, card($\mathbf{Det}_I$) ≥ 2. Let $\mathbf{Des}_I$ denote the set of all descendants of determinants of I in $\mathcal{G}_V$.[2] Assume $I \in \mathbf{O} \subseteq \mathbf{V}$. Then, I is *ill-defined* in $\mathbf{O}$ if and only if, for some $\text{Des}_j \in \mathbf{Des}_I$, there exists a variable Y such that (i) $Y \in \mathbf{O}$, (ii) $Y \neq I$, (iii) Y belongs to a (possibly empty) active path from Det_i to Des_j in $\mathcal{G}_V$ (i.e., $\text{Det}_i \rightarrow \cdots \rightarrow \text{Des}_j$ or $\text{Det}_i \rightarrow \cdots \Rightarrow \text{Des}_j$), and (iv) $\mathcal{G}^{\text{IC}}_{\{I,Y\}}$ contains no shock S_Y common to I, Y, for which $S_Y \perp\!\!\!\perp Y|I$ in $\mathfrak{C}$. Any variable in $\mathbf{O}$ that is not ill-defined in $\mathbf{O}$ is *well-defined* in $\mathbf{O}$.

For instance, TC is well-defined in {Da, TC} because Da is neither a determinant of TC nor a descendant of a determinant of TC and vice versa. By contrast, TC is ill-defined in {HDL, TC} because HDL is a determinant of TC, and $S_{\text{HDL}} \not\perp\!\!\!\perp \text{HDL}|\text{TC}$. Also, TC is ill-defined in {TC, D1} and {TC, HD} because D1 and HD are effects of determinants of TC, and (respectively) $S_{D1} \not\perp\!\!\!\perp \text{D1}|\text{TC}$ and $S_{\text{HD}} \not\perp\!\!\!\perp \text{HD}|\text{TC}$. More generally, a variable I is ill-defined in $\mathbf{O}$ if and only if $\mathbf{O}$ also contains a variable Y among I's determinants or their descendants, and $\mathcal{G}^{\text{IC}}_{\{I,Y\}}$ contains no shock S_Y on I, Y, such that I screens off S_Y from Y in $\mathfrak{C}$. This lack of screening off intuitively captures the idea that a manipulation of I with respect to Y is ambiguous. In turn, to explain the lack of screening off, we need the following proposition (proof in appendix):

Proposition 1 Let $\mathfrak{C}$ over $\mathbf{V} = \mathbf{W} \cup \mathbf{I}$ contain the assignment $I \coloneqq f(\mathbf{Det}_I)$, card($\mathbf{Det}_{\text{I}}$) ≥ 2. Assume CMC and CFC in $\mathcal{G}_W$. Then, for any Det_i, Des_i, Anc_i, where Des_i is a descendant of Det_i, and Anc_i is an ancestor of Det_i, it holds that $\text{Anc}_i \not\perp\!\!\!\perp \text{Des}_i|I$, except for a parameter set Θ (characterizing the assignments in $\mathfrak{C}$) that violates CFC in $\tilde{\mathcal{G}}_V$.[3]

We can also define a graph $\mathcal{G}_O = \langle \mathbf{O}, \mathcal{E}_O \rangle$ representing the structure over $\mathbf{O}$, where $\mathcal{E}_O$ consists of the following edges. First, $\mathcal{G}_O$ has a modified edge $X \Rightarrow Y$ if and only if there is a directed path $X \Rightarrow \cdots \Rightarrow Y$ in $\mathcal{G}_V$, and no variable between X and Y is in $\mathbf{O}$. Next, let the tail $\diamond$ of the arrow $X \diamond\!\!\rightarrow Y$ indicate that X is ill-defined in $\{X, Y\}$. Then, $\mathcal{G}_O$ has an edge $X \diamond\!\!\rightarrow$ Y for any $\langle X, Y, Z \rangle$ for which $X, Y \in \mathbf{O}$, $Z \in \mathbf{V}$, $Z \notin \mathbf{O}$, and $\mathcal{G}_V$ features a path $X \Leftarrow Z \rightarrow Y$, unless $\mathcal{G}^{\text{IC}}_O$ has a shock S common to X, Y for which $S \perp\!\!\!\perp Y|X$ in $\mathfrak{C}$, in which case $X \rightarrow Y$ is in $\mathcal{G}_O$. Furthermore, $\mathcal{G}_O$ has a standard edge $X \rightarrow Y$ if $\mathcal{G}_V$ has a directed path from X to Y featuring standard edges $\rightarrow$ or modified edges $\Rightarrow$, and no variable between X and Y is in $\mathbf{O}$. Finally, $\mathcal{G}_O$ has a bidirected edge $X \longleftrightarrow Y$ if and only if $\mathcal{G}_V$ has an active path $X \leftarrow \cdots \leftarrow Z \rightarrow \cdots \rightarrow Y$ featuring standard

2. Notice that $\mathbf{Det}_I \subseteq \mathbf{Des}_I$ by definition of "descendant."

3. We do not assume CFC in $\tilde{\mathcal{G}}_V$. For such a Θ, I counts as well-defined in our framework, as the manipulation of I with respect to Des_i is not ambiguous.

or modified edges, and only X, Y on that path are in **O**. No further edges are in $\mathcal{G}_O$.

Illustrated in relation to figure 1, $\mathcal{G}_{\{\text{HDL,TC,LDL}\}}$ is HDL $\Rightarrow$ TC $\Leftarrow$ LDL, and $\mathcal{G}_{\{\text{HDL,LDL,HD}\}}$ is HDL $\rightarrow$ HD $\leftarrow$ LDL. The two problematic structures with ill-defined variables from section 2, namely, $\mathcal{G}_{\{\text{TC},D1,D2\}}$ and $\mathcal{G}_{\{\text{Da,TC},D1\}}$, are represented as, respectively, D1 $\leftarrow\!\diamond$ TC $\diamond\!\rightarrow$ D2 and Da $\rightarrow$ TC $\diamond\!\rightarrow$ D1. Finally, let us define the notions of ill- and well-defined causes:

> **Ill- and well-defined causes** For any **X**, $Y \in$ **O**, X is an *ill-defined cause* of Y in **O** if and only if $\mathcal{G}_{\{X,Y\}}$ contains the edge $X \diamond\!\rightarrow Y$. For any X, $Y \in$ **O**, X is a *well-defined cause* of Y in **O** if and only if Y is well-defined in $\{X, Y\}$, and $\mathcal{G}_{\{X,Y\}}$ contains the edge $X \rightarrow Y$.

For instance, HDL is a well-defined cause of HD in {HDL, HD}.[4] By contrast, TC is an ill-defined cause of HD in {TC, HD}.

4. Identification. We now illustrate the applicability of our framework to detecting ill-defined variables and improving causal inference. We begin with a condition under which one may unambiguously identify ill-defined variables.

> **Proposition 2: Sufficient condition for ill-definedness** Consider $\mathfrak{C}$ over **V**, and $\mathbf{O} = \{X, Y, Z\} \subseteq \mathbf{V}$. Assume CMC and CFC in $\mathcal{G}_W$. Also assume (i) $X \perp\!\!\!\perp Z$, (ii) $X \not\perp\!\!\!\perp Y$, $Y \not\perp\!\!\!\perp Z$, $X \not\perp\!\!\!\perp Z | Y$, and (iii) $\mathcal{G}_O^{\text{IC}}$ has no idiosyncratic shock on Y. Then, Y is ill-defined in **O** with two determinants in **V**, and $\mathcal{G}_O$ is $X \leftarrow\!\diamond Y \diamond\!\rightarrow Z$.

For instance, applied to $\mathbf{V}' = \{\text{TC, D1, D2}\}$, this condition establishes that TC is an ill-defined common cause of D1 and D2 (i.e., D1 $\leftarrow\!\diamond$TC$\diamond\!\rightarrow$ D2), since D1 $\perp\!\!\!\perp$ D2, D1 $\not\perp\!\!\!\perp$ TC, TC $\not\perp\!\!\!\perp$ D2, D1 $\not\perp\!\!\!\perp$ D2|TC, and $\mathcal{G}_{V'}{}^{\text{IC}}$ has no idiosyncratic shock to TC. Proposition 2 is easily generalizable to cases with more than two determinants.

If one observes no effects of independent determinants of the ill-defined variable; for instance, in $\mathbf{V}'' = \{\text{Da, TC, D1}\}$, the above condition is not applicable. Nonetheless, one may still reduce the ambiguity concerning ill-defined variables and partially recover the causal structure. To this end, let us assume that determinism induces dependencies (DD):

> **(DD)** For any I and any $\text{Det}_i \in \mathbf{Det}_I$ in $\mathfrak{C}$, it holds that $I \not\perp\!\!\!\perp \text{Det}_i$.

4. At the same time, HDL is not a (well-defined) cause of TC in {HDL, TC}, because TC is not well-defined in that set.

In other words, there are probabilistic dependencies between variables with deterministic assignments and their determinants. This assumption is only violated by canceling paths from determinants to determined variables. Its satisfaction requires (similarly to CFC) the absence of special parameterizations. For simplicity, we also assume that **O** contains no determinants of variables in **O**, such that $\mathcal{E}_O$ contains no modified edges $\Rightarrow$.[5] Then, one may identify well-defined variables:

> **Proposition 3: Sufficient condition for well-definedness** Consider $\mathfrak{C}$ over **V**, and $\mathbf{O} \subseteq \mathbf{V}$. Assume DD. Assume CMC and CFC in $\mathcal{G}_W$. Assume that no determinant of ill-defined variables in **O** is in **O**. Then, a variable X is well-defined in **O** if for any Y in **O**, $X \neq Y$, and one of i–iv holds: (i) $X \perp\!\!\!\perp Y$; (ii) in $\mathcal{G}^{\text{IC}}_{\{X,Y\}}$, X is not a child of an idiosyncratic shock, and X, Y are children of a common shock S, such that $S \perp\!\!\!\perp Y|X$; (iii) in $\mathcal{G}^{\text{IC}}_{\{X,Y\}}$, X is the only child of an idiosyncratic shock; (iv) in $\mathcal{G}^{\text{IC}}_{\{X,Y\}}$, X, Y are children of idiosyncratic shocks, and there is $\mathbf{Z} \subset \mathbf{O}$ such that $X \perp\!\!\!\perp Y|\mathbf{Z}$ and no $Z_i \in \mathbf{Z}$ is the child of an idiosyncratic shock in $\mathcal{G}^{\text{IC}}_{\{X,Z_i\}}$.

For instance, Da (which, to recall, causes LDL but not HDL) is well-defined in $\mathbf{V}''$, since (i) Da $\perp\!\!\!\perp$ D1, and (ii) $\mathcal{G}^{\text{IC}}_{\{\text{Da},\text{TC}\}}$ contains a shock S common to Da, TC, such that $S \perp\!\!\!\perp$ TC|Da, and no idiosyncratic shock to Da, from which one may infer Da $\rightarrow$ TC. Next, one can identify putative ill-defined variables:

> **Proposition 4: Necessary condition for ill-definedness** Consider $\mathfrak{C}$ over **V** and its associated graph $\mathcal{G}_V$. Assume DD. Assume CMC and CFC in $\mathcal{G}_W$. Let X be ill-defined in $\mathbf{O} = \{X, Y\}$ with $\mathbf{Det}_X \cap \mathbf{O} = \emptyset$, $\mathbf{O} \subseteq \mathbf{V}$. Then (i) $X \not\perp\!\!\!\perp Y$; (ii) in $\mathcal{G}^{\text{IC}}_O$, X, Y are children of a common shock; (iii.a) in $\mathcal{G}^{\text{IC}}_O$, X is child of an idiosyncratic shock, or (iii.b) in $\mathcal{G}^{\text{IC}}_O$, X is not a child of an idiosyncratic shock and there is a set of shocks **S** on X such that $X \perp\!\!\!\perp Y|\mathbf{S}$.

For instance, TC and D1 are such that (i) TC $\not\perp\!\!\!\perp$ D1. Moreover, in $\mathcal{G}^{\text{IC}}_{\{\text{TC},D1\}}$ they are (ii) children of a common shock and (iii.a) children of idiosyncratic shocks. Therefore, TC and D1 qualify as putatively ill-defined. Assuming the absence of bidirected modified paths, $\mathcal{G}_{\{\text{TC},D1\}}$ cannot be TC $\Leftarrow \cdots \Rightarrow$ D1. Therefore, only three structures are possible, namely, TC $\diamond\!\!\rightarrow$ D1, TC $\leftarrow\!\!\diamond$ D1, and TC $\longleftrightarrow$ D1. The ambiguity may be resolved by enlarging $\mathbf{V}''$ until a sufficient set **Z** of common causes of TC, D1 is found that screens them off, or (given **Z**) the dependence between TC and D1 is oriented such that one is a

5. Of course, there is no a priori guarantee that **O** contains no determinants. Although one could easily relax this assumption, and thereby obtain a more general result, this would require a lengthier proof. For reasons of space, here we prioritize simplicity over generality.

well-defined cause of the other (i.e., TC $\rightarrow$ D1 or TC $\leftarrow$ D1), or enough effects of determinants of TC or D1 are observed as to remove the idiosyncratic shock on TC or D1, such that either TC $\diamond\!\!\rightarrow$ D1 or TC $\leftarrow\!\!\diamond$ D1 holds.

5. Conclusion. The problem of variable definition is known to be responsible for ambiguous manipulations. Furthermore, we showed that it can lead to mistaken causal inferences by standard constraint-based causal discovery methods. To address the problem, we introduced a novel representation framework suitable for structures including ill-defined variables, that is, the IC representation. We argued that recovering the IC representation can unambiguously identify ill-defined variables, under certain assumptions, or at least exclude that certain variables are ill-defined and consequently reduce the risk of mistakes. Given recent advances in statistical techniques (e.g., ICA) by which one may recover the IC representation, our proposal holds great promise. Therefore, we strongly invite further research on the subject.

Appendix

Proof of Proposition 1. Assume per absurdum that there exist Anc_i, Des_i of Det_i, such that $\mathrm{Anc}_i \perp\!\!\!\perp \mathrm{Des}_i | I$ for any set of parameters Θ in $\mathfrak{C}$. This is possible only if one of A–C holds: (A) Det_i suffices to determine I, such that I renders Det_i irrelevant to Anc_i, Des_i. This requires $\mathrm{card}(\mathbf{Det}_I) = 1$, contradicting $\mathrm{card}(\mathbf{Det}_I) \geq 2$. (B) $\mathrm{card}(\mathbf{Det}_I) \geq 2$ and for some $\mathrm{Det}_j \in \mathbf{Det}_I$, there is no directed path $\mathrm{Det}_j \rightarrow \cdots \rightarrow \mathrm{Des}_i$. Then, Det_j would act as an exogenous noise on I, such that the edge $\mathrm{Det}_i \Rightarrow I$ would be observationally indistinguishable from a standard edge $\mathrm{Det}_i \rightarrow I$. Holding CFC in $\mathbf{W}$, and since I behaves like a child of Det_i, we would have $\mathrm{Anc}_i \not\perp\!\!\!\perp \mathrm{Des}_i | I$, contradicting our starting hypothesis. (C) $\mathrm{card}(\mathbf{Det}_I) \geq 2$ and for any $\mathrm{Det}_j \in \mathbf{Det}_I$, there is a directed path $\mathrm{Det}_j \rightarrow \cdots \rightarrow \mathrm{Des}_i$. Then, there exists a parameter set Θ such that $\mathrm{Anc}_i \perp\!\!\!\perp \mathrm{Des}_i | I$. For instance, assume $\mathrm{card}(\mathbf{Det}_I) = 2$ and a generalized additive model where $I = f(\mathrm{Det}_1) + g(\mathrm{Det}_2)$, $\mathrm{Det}_1 = f'(\mathrm{Anc}) + S_{\mathrm{Det1}}$, $\mathrm{Det}_2 = g'(\mathrm{Anc}) + S_{\mathrm{Det2}}$, $\mathrm{Des} = f''(\mathrm{Det}_1) + g''(\mathrm{Det}_2) + S_{\mathrm{Des}}$. Then, $\mathrm{Anc} \perp\!\!\!\perp \mathrm{Des} | I$ holds if and only if (i) $f(\mathrm{Det}_1) = g(\mathrm{Det}_2)$ and $f'(\mathrm{Anc}) = g'(\mathrm{Anc})$ or (ii) $f(\mathrm{Det}_1) = g(\mathrm{Det}_2)$ and $f''(\mathrm{Det}_1) = g''(\mathrm{Det}_2)$. This point generalizes to larger cardinalities. Finally, since I is a parent of neither Anc_i nor Des_i in $\tilde{\mathcal{G}}_V$, any such parameter set Θ realizing $\mathrm{Anc}_i \perp\!\!\!\perp \mathrm{Des}_i | I$ necessarily violates CFC in $\tilde{\mathcal{G}}_V$. QED

Proof of Proposition 2. Let $^*\!\!\rightarrow$ denote one among $\rightarrow$, $\leftrightarrow$, and $\diamond\!\!\rightarrow$. Assume per absurdum that i–iii are true, but Y is well-defined. CMC and ii entail that $\mathcal{G}_V$ contains paths linking X, Y and Y, Z. CFC and i entail that $\mathcal{G}_V$ contains no path linking X, Z. Then, $\mathcal{G}_O$ contains only two edges, one connecting X, Y, and one connecting Y, Z. Among the possible structures in

$\mathcal{G}_O$, $X *\!\!\rightarrow Y \rightarrow Z$, $X \leftarrow Y \leftarrow\!\!* Z$, $X \leftarrow\!\!* Y \rightarrow Z$, and $X \leftarrow Y *\!\!\rightarrow Z$ contradict i, and $X *\!\!\rightarrow Y \leftarrow\!\!* Z$ contradicts iii. In all other structures, that is, $X \leftarrow\!\!\diamond Y \diamond\!\!\rightarrow Z$, $X *\!\!\rightarrow Y \diamond\!\!\rightarrow Z$, and $X \leftarrow\!\!\diamond Y \leftarrow\!\!* Z$, Y is ill-defined. The latter two contradict iii. Thus, $\mathcal{G}_O$ is $X \leftarrow\!\!\diamond Y \diamond\!\!\rightarrow Z$, and $\mathbf{Det}_Y$ has precisely two elements in $\mathbf{V}$ (one causing X and one causing Z); otherwise, $\mathcal{G}_O^{\mathrm{IC}}$ would contain an idiosyncratic shock on Y associated to its extra determinant(s), violating iii. As a corollary, $\mathcal{G}_O^{\mathrm{IC}}$ contains idiosyncratic shocks on X and Z. QED

Proof of Proposition 3. (i) From the definition of ill-defined variable, for any $I \in \mathbf{V}$, $\mathcal{G}_V$ contains a directed path from some $\mathrm{Det}_i \in \mathbf{Det}_I$ to some descendant Des_j of Det_i. Under CFC and DD, I is ill-defined only if $\mathbf{O}$ contains some Y on that path, such that $I \not\perp\!\!\!\perp Y$. Hence, if $\mathbf{O} = \{X, Y\}$ and $X \perp\!\!\!\perp Y$, then X is well-defined. (ii) In $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$, X is ill-defined and not a child of an idiosyncratic shock only if $\mathcal{G}_V$ contains directed paths from each $\mathrm{Det}_i \in \mathbf{Det}_X$ to Y. Then, $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$ contains no shock S common to X, Y, such that $S \perp\!\!\!\perp Y|X$. Since this contradicts ii, X cannot be ill-defined. (iii) If X is the only child of an idiosyncratic shock in $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$, then $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$ contains a shock common to X, Y. Then, X is ill-defined in $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$ only if $\mathbf{O}$ contains a node $\mathrm{Det}_i \in \mathbf{Det}_X$, which is not a child of an idiosyncratic shock, contradicting the assumption that $\mathbf{Det}_X \cap \mathbf{O} = \varnothing$. Hence, X is well-defined. (iv) Suppose per absurdum that X is ill-defined, entailing a directed path $\mathrm{Det}_i \rightarrow \cdots \rightarrow Y$ in $\mathcal{G}_V$. Since $X \perp\!\!\!\perp Y|\mathbf{Z}$, some $Z_i \in \mathbf{Z} \subset \mathbf{O}$ is on that path. Then, Z_i is a child of an idiosyncratic shock in $\mathcal{G}^{\mathrm{IC}}_{\{X,Z_i\}}$, contradicting iv. Hence, X is well-defined. QED

Proof of Proposition 4. Preamble: From the definition of ill-defined variable, and from $\mathbf{Det}_X \cap \mathbf{O} = \varnothing$, it follows that $\mathcal{G}_{\{X,\, Y\}}$ is $X \diamond\!\!\rightarrow Y$. (i) Under CFC and DD, the preamble implies $X \not\perp\!\!\!\perp Y$. (ii) By definition of IC representation, $\mathcal{G}^{\mathrm{IC}}_{\{X,Y\}}$ contains at least one common shock to X, Y due to a latent determinant of X. (iii) If $\mathcal{G}_V$ contains a determinant of X not linked to Y by a directed path, then X is a child of an idiosyncratic shock (iii.a). If, on the contrary, all determinants of X are linked to Y by directed paths in $\mathcal{G}_V$, then X is not a child of an idiosyncratic shock. Additionally, given $X \perp\!\!\!\perp Y|\mathbf{Det}_X$, it follows that there is a set $\mathbf{S}$ of shocks on X's determinants, such that $X \perp\!\!\!\perp Y|\mathbf{S}$ (iii.b). QED

REFERENCES

Eberhardt, F. 2016. "Green and Grue Causal Variables." *Synthese* 193:1029–46.

Hyvärinen, A., J. Karhunen, and E. Oja. 2001. *Independent Component Analysis*. New York: Wiley.

Pearl, J. 2009. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press.

Peters, J., D. Janzing, and B. Schölkopf. 2017. *Elements of Causal Inference*. Cambridge, MA: MIT Press.

Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge MA: MIT Press.

Spirtes, P., and R. Scheines. 2004. "Causal Inference of Ambiguous Manipulations." *Philosophy of Science* 71:833–45.

Woodward, J. 2016. "The Problem of Variable Choice." *Synthese* 193:1047–72.