Statistica Sinica Preprint No: SS-2019-0207	
Title	Sufficient cause interactions for categorical and ordinal outcomes
Manuscript ID	SS-2019-0207
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202019.0207
Complete List of Authors	Jaffer Zaidi and Tyler VanderWeele
Corresponding Author	Jaffer Zaidi
E-mail	jaffer.zaidi@gmail.com
Notice: Accepted version subject to English editing.	

Sufficient cause interactions for categorical and ordinal outcomes

Jaffer M. Zaidi and Tyler J. VanderWeele

Harvard University

Abstract: The sufficient cause model is extended from binary to categorical and ordinal outcomes to formalize the concept of sufficient cause interaction and synergism in this setting. This extension allows for the derivation of counterfactual and empirical conditions for detecting the presence of sufficient cause interactions for ordinal and categorical outcomes. Some of these conditions are entirely novel in that they cannot be derived from the sufficient cause model for binary outcomes. These empirical conditions enable researchers to discover whether two exposures display synergism for an ordinal or categorical outcome. Likelihood ratio tests that use these derived empirical conditions are developed to infer sufficient cause interaction for ordinal and categorical outcomes. These likelihood ratio tests are used to detect sufficient cause interaction between two major resistance mutations in the development of HIV drug resistance to Etravirine.

Key words and phrases: interaction, sufficient cause, ordinal outcome.

1. Introduction

In this paper, we extend the sufficient cause model defined originally for binary outcomes to categorical and ordinal outcomes, and we derive

the associated empirical and counterfactual conditions associated with sufficient cause interaction. The sufficient cause framework provides a representation of causation in terms of a collection of causal mechanisms, called sufficient causes. A single sufficient cause is constituted of one or more component causes such that when all components of the sufficient cause are present they will together inevitably bring about the outcome. The first crude sufficient cause model appeared in (Cayley, 1853). Rothman popularized the sufficient cause model in epidemiology, and introduced a graphical schematic which is often presented in introductory epidemiology texts (Rothman, 1976). The sufficient cause model has evolved dramatically over the past decade to enable the detection of different forms of interaction (VanderWeele and Robins, 2008; Berzuini and Dawid, 2016; VanderWeele, 2015; VanderWeele and Richardson, 2012; Vanderweele, 2010; Ramsahai, 2013).

Rothman (1976) presented a model for causation as a series of different causal mechanisms each of which that are sufficient to bring about the outcome. In this model, the causal mechanisms are called "sufficient causes," that are defined as the minimal set of actions, events or states of being that jointly initiate a process that will eventually result in the outcome. Many different "sufficient causes" can produce a particular outcome. For instance,

in the course of treatment for HIV-1 viral mutations can arise. Some mutations might on their own make an a particular treatment ineffective, while others might require one or more additional mutations to operate.

Within a deterministic framework, for a binary outcome, suppose we were considering three known potential causes, X_1, X_2, X_3. Suppose hypothetically that it is the case that mutation X_1, and unknown factors A_1 make an individual drug resistant, denoted by binary outcome R. In contrast, mutations X_2 and X_3 will together be sufficient if jointly present with additional unknown factors A_2. A last mechanism might be mutations X_1 and X_2 with additional unknown factors A_3. This provides us with three sufficient causes, denoted A_1X_1, $A_2X_2X_3$, $A_3X_1X_2$, each of which when present will make the individual drug resistant.

In a deterministic sufficient cause model, whenever all of the component causes of a particular sufficient cause are present the outcome will definitely occur, and each component cause is necessary for that particular sufficient cause to bring about the outcome. Sufficient cause $A_2X_2X_3$ has two component causes X_2 and X_3. This particular sufficient cause will not operate if either X_2 or X_3 is not present. This phenomenon whereby two component causes are both needed to cause the outcome to occur is termed synergism. In general it may be logically possible to represent the counterfactual out-

comes across the different causes by different representations of sufficient causes. When it is the case that every such possible sufficient cause representation has a particular conjunction, say X_2X_3, then a "sufficient cause interaction" between X_2 and X_3 is said to be present. In such cases we would then know that synergism must be present between X_2 and X_3. Scientists will want to discover synergism from data, and statisticians have derived empirical conditions to enable the discovery of sufficient cause interactions (VanderWeele and Robins, 2008; Berzuini and Dawid, 2016; VanderWeele, 2015; VanderWeele and Richardson, 2012; Vanderweele, 2010; Ramsahai, 2013).

In this paper, we extend the sufficient cause model to categorical and ordinal outcomes, and we also develop the associated likelihood ratio tests and provide one data application of this theory. This enables researchers to understand which mutations mechanistically interact in the development of HIV-1 drug resistance to Etravirine. For this applied problem the ordinal outcome has three levels: no drug resistance, partial drug resistance, and full drug resistance.

2. Sufficient cause interactions for a specified outcome

Suppose, we have an outcome Y with associated levels $Y \in \{0, 1, 2\}$. We denote binary variables $X_1, \ldots, X_k$ that each take values in $X_i \in \{0, 1\}$. Jointly, $X^k = (X_1, \ldots, X_k)$ take values within $X^k \in \{0, 1\}^k$. The individuals, denoted by symbol ω, compose a population, denoted Ω. We write the potential outcome $Y_{x_1,\ldots,x_k}(\omega)$ of Y for individual ω if for $j = 1, \ldots, k$, each putative cause $X_j \in \{X_1, \ldots, X_k\}$ were set $x_j \in \{0, 1\}$. The data application considers the situation when $k = 2$. For this circumstance, the potential outcome or counterfactual value of an individual ω had X_1 been set to x_1 and X_2 been set to x_2 is denoted $Y_{x_1,x_2}(\omega)$. There are 3^4 potential response types, $\underline{Y}_{x_1x_2}(\omega) = (Y_{11}(\omega), Y_{10}(\omega), Y_{01}(\omega), Y_{00}(\omega))$, that form all the different of types individuals, which we denote $Y_{x_1x_2}(\Omega)$. This table $Y_{x_1x_2}(\Omega)$ is simply all the different permutations of a vector of length four sampling with replacement from the set $\{0, 1, 2\}$.

An indicator function denoted $I(Y \in S)$ is used to denote a new random variable constructed from Y, which takes value 1 if $Y \in S$ and 0 otherwise. To construct these new binary outcomes, let $A = \{1\}$, $B = \{1, 2\}$, $C = \{2\}$, $D = \{0\}$, $E = \{0, 2\}$, $F = \{0, 1\}$. Specifically, we denote: $Y^L = I(Y \in L)$, where $L \in \{A, B, C, D, E, F\}$. Potential outcome versions of Y^L are defined as $Y^L_{x_1,\ldots,x_k}(\omega) = I(Y_{x_1,\ldots,x_k}(\omega) \in L)$, where $L \in \{A, B, C, D, E, F\}$. The

superscript L in the symbol Y^L does not indicate exponentiation, but rather specifies the condition that is used to construct this new random variable that is constructed from Y. Appendix 1 provides the full list of the different Y^L and $Y^L_{x_1,\ldots,x_k}(\omega)$ without set notation for the reader's convenience. We require the consistency assumption, namely that $Y_{X_1(\omega),\ldots,X_k(\omega)}(\omega) = Y(\omega)$, which states that the value of Y that would have been observed if $X_1, \ldots, X_k$ had been set to what in fact they were is equal to the value of Y that was observed. The consistency assumption for Y^L is implied by the consistency assumption on Y. The disjunctive operator on binary variables $X_1, \ldots, X_k$ is denoted $\vee_{i\in\{1,\ldots k\}} X_i = X_1 \vee \ldots \vee X_k = \max\{X_1, \ldots, X_k\}$. For ease of notation, we shall drop the commas between the intervened variables $\{X_1, \ldots, X_k\}$ in a potential outcome, for example $Y_{x_1,x_2}(\omega) = Y_{x_1x_2}(\omega)$.

The definitions and theorems in this section closely mimic the associated definitions and theorems from VanderWeele and Robins (2008). While this paper is self-contained, a reader that is familiar with VanderWeele and Robins (2008) would recognize that Definitions (2.1)-(2.7) and Theorems (2.1)-(2.5) are the logical extensions to corresponding definitions and theorems presented in VanderWeele and Robins (2008), and we thus keep the exposition very concise. Theorem 2.6 and Corollary 2.1 cannot be derived through the previous framework on sufficient causes (VanderWeele

and Robins, 2008, 2012; Ramsahai, 2013) based upon binary outcomes.

Definition 2.1 (Sufficient cause for a specified outcome)**.** We say that putative binary causes $X_1, \ldots, X_n$ are called sufficient causes for Y^L where $L \in \{A, B, C, D, E, F\}$, if for all values of $x_1, \ldots, x_n \in X^n$ such that $x_1 \times \cdots \times x_n = 1$ we have that $Y^L_{x_1 \cdots x_n}(\omega) = 1$ for all $\omega \in \Omega' \subseteq \Omega$ where $\Omega' \neq \varnothing$.

Definition 2.2 (Minimal sufficient cause for a specified outcome)**.** We say that putative binary causes $X_1, \ldots, X_n$ form a minimal sufficient cause for Y^L where $L \in \{A, B, C, D, E, F\}$, if $X_1, \ldots, X_n$ are sufficient causes for Y^L and no proper subset of $\{X_1, \ldots, X_n\}$ is also a sufficient cause for Y^L.

Definition 2.3. Determinative sufficient causes for a specified outcome A set of sufficient causes $M^L_1, \ldots, M^L_n$ each of which are composed of a product of binary causes for a specified outcome Y^L, where $L \in \{A, B, C, D, E, F\}$, is defined to be determinative for Y^L if for all $\omega \in \Omega$, $Y^L_{x_1 \cdots x_s}(\omega) = 1$ if and only if $M^L_1 \vee M^L_2 \vee \ldots \vee M^L_n = 1$.

Definition 2.4. Non-redundant sufficient causes for a specified outcome A set of determinative sufficient causes $M^L_1, \ldots, M^L_n$ for Y^L, where $L \in \{A, B, C, D, E, F\}$, is called a non-redundant determinative set of minimal sufficient causes if there is no proper subset of $M^L_1, \ldots, M^L_n$ that is also a determinative set of minimal sufficient causes for Y^L.

VanderWeele and Robins (2008) note that minimality and non-redundancy should be distinguished. Minimality concerns components of a given conjunction in that each component is necessary for the conjunction to be sufficient for the outcome to occur. Non-redundancy concerns the disjunction of conjunctions in that each individual conjunction should be present in order for the disjunction to be determinative.

Example 2.1. Suppose an individual is taking treatment for HIV. Viral mutations can occur while the individual takes treatment. Suppose mutation X_1 enables the virus to replicate in particular cells in the human body, and mutation X_2 enables the virus to penetrate these particular cells body, and assume for now that these are the only two mutations that occur. The scientist could ask whether mutations X_1 and X_2 required for this individual for the current treatment to become ineffective in treating HIV, which is known as drug resistance. Alternatively, would mutation X_1 on its own suffice for the individual to develop drug resistance. Scientists also grade drug resistance on an ordinal scale, partial and full. A scientist might believe that mutation X_2 alone sufficient in the development of partial drug resistance for a particular individual, but conjunction X_1 and X_2 are both necessary for full drug resistance for the same drug.

These definitions for sufficient cause for an ordinal or nominal outcome

with three levels generalize the analogous definitions of sufficient cause for a binary outcome. The definitions provided herein are easily adaptable to the case where a researcher is interested in an ordinal outcome Y, where $Y \in \{0, 1, \ldots, j\}$, and such definitions for outcome Y with j levels are presented in the online supplement. A very brief exposition of this generalization is provided in section 4. More general notions of interdependence (Ramsahai, 2013) extended to categorical and ordinal outcomes are also provided in the online supplement. Denote $\bar{X}_i$ as the complement of X_i.

Theorem 2.1. *For putative binary causes X_1 and X_2 of specified outcome Y^L, where $L \in \{A, B, C, D, E, F\}$, there exist binary variables*

$$A_0^L(\omega), A_1^L(\omega), A_2^L(\omega), A_3^L(\omega), A_4^L(\omega), A_5^L(\omega), A_6^L(\omega), A_7^L(\omega), A_8^L(\omega),$$

which are functions of the counterfactuals $\{Y_{11}^L(\omega), Y_{10}^L(\omega), Y_{01}^L(\omega), Y_{00}^L(\omega)\}$ such that

$$\begin{aligned} Y^L &= A_0^L \vee A_1^L X_1 \vee A_2^L \bar{X}_1 \vee A_3^L X_2 \vee A_4^L \bar{X}_2 \vee A_5^L X_1 X_2 \\ &\quad \vee A_6^L \bar{X}_1 X_2 \vee A_7^L X_1 \bar{X}_2 \vee A_8^L \bar{X}_1 \bar{X}_2, \end{aligned} \tag{2.1}$$

and such that

$$\begin{aligned} Y_{x_1 x_2}^L &= A_0^L \vee A_1^L x_1 \vee A_2^L (1 - x_1) \vee A_3^L x_2 \vee A_4^L (1 - x_2) \vee A_5^L x_1 x_2 \\ &\quad \vee A_6^L (1 - x_1) x_2 \vee A_7^L x_1 (1 - x_2) \vee A_8^L (1 - x_1)(1 - x_2). \end{aligned}$$

The proof of Theorem 2.1 mimics the proof of sufficient cause representation for binary outcomes provided in (VanderWeele and Robins, 2008). For completeness, we provide the proof in the online supplement. We call equation (2.1) a sufficient cause representation of Y^L.

We are able to generalize our definitions provided above and Theorem 2.1 to the situation where the analyst is concerned about defining and analyzing minimum sufficient cause interaction on an ordinal variable with multiple levels, i.e. more than three. This generalization is provided in the supplementary materials. This theorem extends the results provided in VanderWeele and Robins (2008); Theorem 1 also provides a method to construct variables A_i^L as a function of the potential outcomes that together with disjunctions built on the set $\{X_1, X_2, \bar{X}_1, \bar{X}_2\}$ make a determinative set of sufficient causes for Y^L, where $L \in \{A, B, C, D, E\}$. Each of the conjunctions $A_0^L, A_1^L X_1, \ldots, A_8^L \bar{X}_1 \bar{X}_2$ are sufficient to cause Y^L, where $L \in \{A, B, C, D, E, F\}$. The disjunction of all of these conjunctions makes a determinative set of sufficient causes for Y^L, where $L \in \{A, B, C, D, E, F\}$. Similar to the binary outcome context, A_i^L variables could be considered as unknown factors that together with the associated conjunction of $\varnothing, X_1, X_2, \bar{X}_1, \bar{X}_2, X_1\bar{X}_2, \bar{X}_1 X_2, \bar{X}_1\bar{X}_2$ complete the sufficient cause for specified outcome Y^L, where $L \in \{A, B, C, D, E, F\}$.

Now that we have defined sufficient cause for a specified outcome, we define sufficient cause interactions for a specified outcome. Based upon these definitions, counterfactual and empirical conditions are derived to detect the presence of sufficient cause interactions for a specified outcome.

Example 2.2. Consider the drug resistance example presented earlier. Suppose we have two only types of individuals in our population. Individual 1 would develop full drug resistance if she has either of the two mutations, while individual two develops full drug resistance only if she has both mutations. The construction of the variables A_i from Theorem 1 would give us

$$Y^C(\omega) = A_1^C(\omega)X_1(\omega) \vee A_2^C(\omega)X_2(\omega) \vee A_8^C(\omega)X_1(\omega)X_2(\omega),$$

where Y^C denotes full drug resistance. Suppose for these same two individuals, they would develop partial drug resistance if they have either of the two mutations, then an application of Theorem 1 would give us

$$Y^A(\omega) = X_1(\omega) \vee X_2(\omega),$$

where $Y^A(\omega)$ denotes partial drug resistance.

Definition 2.5 (Minimal sufficient cause interaction for a specified outcome)**.** Suppose $F_1 \in \{X_1, \bar{X}_1\}$ and $F_2 \in \{X_2, \bar{X}_2\}$. If in every non-redundant minimal sufficient cause representation for a specified outcome

Y^L,where $L \in \{A, B, C, D, E, F\}$ we are able to find a sufficient cause that contains F_1F_2, then we say that the conjunction F_1F_2 exhibits or displays minimal sufficient cause interaction for outcome Y^L.

Definition 2.6 (Irreducible sufficient cause interactions for a specified outcome)**.** Suppose $F_1 \in \{X_1, \bar{X}_1\}$ and $F_2 \in \{X_2, \bar{X}_2\}$. If in every sufficient cause representation for Y^L, where $L \in \{A, B, C, D, E, F\}$, we are able to find a sufficient cause which contains F_1F_2, then F_1F_2 is said to be irreducible for Y^L.

These two definitions are shown to be equivalent in our case: that is, an irreducible sufficient cause interaction for a specified outcome Y^L is a minimal sufficient cause interactions for a specified outcome Y^L, and vice versa. The theorem and proof demonstrating that the definitions are equivalent replicates the arguments of VanderWeele and Robins (2008), and as such are omitted. Here, we say that the effects of $F_1 \in \{X_1, \bar{X}_1\}$ and $F_2 \in \{X_2, \bar{X}_2\}$ on a specified outcome Y^L, where $L \in \{A, B, C, D, E, F\}$, are synergistic or represent synergism if there is a sufficient cause for Y^L such that F_1F_2 is contained within its conjunction. The rest of the proofs of the theorems and corollaries are collected in the online supplement.

Theorem 2.2. *Suppose $L \in \{A, B, C, D, E, F\}$. There exists an individual $\omega \in \Omega$ for whom $Y^L_{11}(\omega) = 1$ and $Y^L_{10}(\omega) = Y^L_{01}(\omega) = 0$, if and only if*

the conjunction X_1X_2 exhibits sufficient cause interaction for a specified outcome Y^L.

We now consider empirical conditions to detect such sufficient cause interactions. The symbol II is used to denote independence. For example, $Y \amalg X_1$ denotes that Y is marginally independent of X_1, and $Y \amalg X_1 \mid X_2$ denotes that Y is conditionally independent of X_1 given X_2.

Theorem 2.3. *Suppose V is a set of variables that are sufficient to control for the confounding of the variables of X_1 and X_2 on Y^L, where $L \in \{A, B, C, D, E, F\}$, i.e. $Y^L_{x_1x_2} \amalg \{X_1, X_2\} \mid V$. We can conclude that X_1X_2 exhibit sufficient cause interaction for a specified outcome Y^L if for some value v of V, the following inequality holds:*

$$\begin{aligned} 0 \;<\; & E(Y^L \mid X_1 = 1, X_2 = 1, V = v) \\ & -E(Y^L \mid X_1 = 1, X_2 = 0, V = v) \\ & -E(Y^L \mid X_1 = 0, X_2 = 1, V = v) \end{aligned} \tag{2.2}$$

From here on, we use the shorthand notation $p^L_{x_1x_2}$ to denote $P(Y \in L \mid X_1 = x_1, X_2 = x_2)$ and $p^L_{x_1x_2v}$ to denote $P(Y \in L \mid X_1 = x_1, X_2 = x_2, V = v)$. We could replace X_1 or X_2 by either or both of their complements, and derive similar results for antagonism. The results in Theorem 2.3 generalize the results for identifying synergism for a binary outcome as established

in VanderWeele and Richardson (2012) to categorical or ordinal outcomes under a specified condition. Theorems 2.2 and 2.3 have generalizations for categorical or ordinal outcomes with an arbitrary number of levels that will be presented in Section 5. Our approach allows the researcher to detect sufficient cause interaction between two variables for an ordinal outcome under specified conditions at different levels or an amalgam of different levels of the categorical or ordinal outcome. We provide an example to illustrate Theorem 2.3.

Example 2.3. Consider specified $Y^C = I(Y = 2)$. Now consider the left hand side of the inequality (2.2),

$$\begin{aligned}
& E(Y^C \mid X_1 = 1, X_2 = 1, V = v) - E(Y^C \mid X_1 = 1, X_2 = 0, V = v) \\
& \quad - E(Y^C \mid X_1 = 0, X_2 = 1, V = v) \\
= \; & E(I(Y = 2) \mid X_1 = 1, X_2 = 1, V = v) \\
& \quad - E(I(Y = 2) \mid X_1 = 1, X_2 = 0, V = v) \\
& \quad - E(I(Y = 2) \mid X_1 = 0, X_2 = 1, V = v) \\
= \; & P(Y = 2 \mid X_1 = 1, X_2 = 1, V = v) - P(Y = 2 \mid X_1 = 1, X_2 = 0, V = v) \\
& \quad - P(Y = 2 \mid X_1 = 0, X_2 = 1, V = v) \\
= \; & p^C_{11v} - p^C_{10v} - p^C_{01v}
\end{aligned}$$

Therefore if $p^C_{11v} - p^C_{10v} - p^C_{01v} > 0$, we can say that X_1X_2 exhibit sufficient

cause interaction for the outcome $I(Y = 2)$, or equivalently that X_1X_2 exhibit sufficient cause interaction for the ordinal outcome Y at the level 2.

Following exactly the same steps in Example 1, we can show that if $p^C_{11v} + p^A_{11} - p^C_{10v} - p^A_{10v} - p^C_{01v} - p^A_{01v} > 0$, we can say that X_1X_2 exhibit sufficient cause interaction for the outcome $I(Y \geq 1)$, or equivalently that X_1X_2 exhibit sufficient cause interaction for the ordinal outcome Y at the level 1 or 2. Similarly, if $p^D_{11v} + p^C_{11} - p^D_{10v} - p^C_{10v} - p^D_{01v} - p^C_{01v} > 0$, we can say that X_1X_2 exhibit sufficient cause interaction for the outcome $I(Y \in \{0, 2\})$.

VanderWeele and Robins (2008) demonstrate that if it can be assumed that variables have positive monotonic effects on a binary outcome (i.e. the variables never prevent the outcome), then one can use less stringent tests to detect sufficient cause interaction than if one were unable to make this assumption. We will now examine the analogous results in the case of an ordinal outcome with three levels. Instead, if Y were categorical, exactly the same definitions and results, namely Definitions (2.1)-(2.6) and Theorems (2.1)-(2.3), would hold true. Results that require monotonicity will only work with ordinal outcomes as the next definition requires the outcome to be ordinal. Therefore, Theorems (2.4)-(2.6) and Corollary (2.1) are only valid for ordinal outcomes and cannot be applied to categorical outcomes.

Definition 2.7. Monotonic Effect for a ordinal outcome with three levels

For any two binary variables X_1 and X_2, if for all $\omega \in \Omega$, $Y_{x_1x_2}(\omega)$ is non-decreasing in x_1 for any given $x_2 \in X$, then we say that X_1 has a positive monotonic effect on Y. Similarly, if for all $\omega \in \Omega$, $Y_{x_1x_2}(\omega)$ is non-decreasing in x_2 for any given $x_1 \in X$, then we say that X_2 has a positive monotonic effect on Y.

Theorem 2.4. *Suppose X_1 and X_2 both have positive monotonic effects on ordinal variable Y, and that $Y^B_{x_1x_2} \amalg \{X_1, X_2\} \mid V$. If for some value $v \in V$, we have*

$$p^B_{11v} - p^B_{10v} - p^B_{01v} + p^B_{00v} > 0,$$

then X_1 and X_2 display synergism for outcome $Y^B = I(Y \geq 1)$.

Theorem 2.5. *Suppose X_1 and X_2 both have positive monotonic effects on ordinal variable Y, and that $Y^C_{x_1x_2} \amalg \{X_1, X_2\} \mid V$. If for some value $v \in V$, we have*

$$p^C_{11v} - p^C_{10v} - p^C_{01v} + p^C_{00v} > 0,$$

then X_1 and X_2 display synergism for outcome $Y^C = I(Y = 2)$.

These results are exactly the same as the ones that would have been established had one dichotomized the outcome at the outset and applied the empirical conditions established from VanderWeele and Robins (2008). The next theorem provides a result that one is not able to derive on the bases

of previous literature on sufficient cause interaction or mechanistic interaction (VanderWeele and Robins, 2008; Berzuini and Dawid, 2016; Ramsahai, 2013; VanderWeele and Richardson, 2012). This is a novel result that will enable researchers to discover sufficient cause interaction for specified outcome $Y^A = I(Y = 1)$. The proofs of these results are collected in the online supplement.

Theorem 2.6. *Suppose X_1 and X_2 both have positive monotonic effects on ordinal variable Y, and that $Y^A_{x_1x_2}$ II $\{X_1, X_2\} \mid V$. If for some value $v \in V$, we have at least one of the following inequalities*

$$p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{01v} > 0, \tag{2.3.1}$$

$$p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{10v} > 0, \tag{2.3.2}$$

$$p^A_{11v} - p^A_{10v} - p^A_{01v} > 0, \tag{2.3.3}$$

then X_1 and X_2 display synergism for outcome $Y^A = I(Y = 1)$.

Corollary 2.1. *Suppose X_1 and X_2 both have positive monotonic effects on ordinal variable Y, and that $Y^A_{x_1x_2}$ II $\{X_1, X_2\} \mid V$. If for some value $v \in V$,*

we have at least one of the following inequalities

$$2 \cdot p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} > 1, \tag{2.4.1}$$

$$p^A_{11v} - p^D_{11v} - p^C_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} > 0 \tag{2.4.2}$$

$$2p^A_{11v} - p^A_{10v} - p^A_{01v} - p^D_{00v} - p^C_{00v} > 0 \tag{2.4.3}$$

$$p^A_{11v} - p^A_{10v} - p^A_{01v} > 0, \tag{2.4.4}$$

then X_1 *and* X_2 *display synergism for outcome* $Y^A = I(Y = 1)$.

We demonstrate in the online supplement that if $2 \cdot p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} > 1$, then $p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{01v} > 0$ and $p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{01v} > 0$. The converse is not true. This implies that the empirical conditions (2.3.1) and (2.3.2) are weaker than the empirical condition (2.4.1). Conditions (2.4.1)-(2.4.3) are shown to be equivalent to one another in the online supplement. The only circumstance in which we would use condition (2.4.1) instead of conditions (2.3.1) and (2.3.2) is when we do not have data on the outcome $I(Y = 2)$. We would like to note here that Theorems (2.4)-(2.6) are derived in the online supplement through arguments made from the sufficient cause framework and monotonicity. We also derived the same inequality constraints using a different approach that modifies the theory provided in Ramsahai (2013) using convex polytopes from binary outcomes to categorical and ordinal outcomes, and we find that

the empirical conditions presented in this paper are the only inequalities that were observed. A more detailed explanation of how the Ramsahai approach is adapted to the ordinal outcome setting is provided in the online supplement.

2.1 Inference for Sufficient Cause Interaction for Ordinal Outcomes

Previous authors have used likelihood ratio tests to conduct hypothesis tests on moment conditions that stem from problems in causal inference (Ramsahai, 2013; Ramsahai et al., 2011), including sufficient cause interaction for binary outcomes (Ramsahai, 2013). The saturated Bernoulli model is also proposed to detect sufficient cause interactions in the setting of binary outcomes in the presence of covariates (VanderWeele and Richardson, 2012; VanderWeele and Robins, 2008; Vansteelandt et al., 2012). Researchers have also used Bonferonni corrections for testing multiple moment conditions in causal inference literature (Wang et al., 2017). The approach taken here follows likelihood ratio tests (Ramsahai et al., 2011; Ramsahai, 2013). In the setting of a composite null, the likelihood ratio test statistics' asymptotic distribution is obtained assuming the true parameter is on the boundary of the null hypothesis (Van der Vaart, 2000; Ramsahai, 2013; Ramsahai

et al., 2011; Drton, 2009). For the tests considered in the data analysis, the likelihood ratio test follows a weighted mixture of χ^2-distributions (Ramsahai, 2013). A description of the asymptotics of likelihood ratio tests under multiple inequality constraints is available in Silvapulle and Sen (2011). Likelihood ratio tests that use Theorems (2.3)-(2.5) are closely related to previously proposed likelihood ratio tests under inequality constraints for sufficient cause interaction (Ramsahai, 2013). Theorems (2.3)-(2.6) provide the alternative space to each of the specified forms of sufficient cause interaction. As usual, the complement of the alternative space is the null space.

Theorems (2.3)-(2.5) only involve a statistical test with a single inequality constraint. For example, the hypothesis test $H_0 : p^A_{11} - p^A_{10} - p^A_{01} \leq 0$ versus $H_1 : p^A_{11} - p^A_{10} - p^A_{01} > 0$ is a test to establish whether X_1 and X_2 display synergism for outcome $Y^A = I(Y = 1)$. For such hypothesis tests that involve a single inequality constraint, the null space is a half-space (Ramsahai et al., 2011; Ramsahai, 2013; Self and Liang, 1987). Throughout, we let t denote the observed value of the likelihood ratio statistic. For these tests with a single inequality constraint, following Self and Liang, the p-value of the likelihood ratio test is $P(\chi^2_1 > t)/2$ for positive t and 1 otherwise (Self and Liang, 1987; Ramsahai, 2013). On the other hand, for Theorem 2.6,

the null space is defined by the intersection of three half-spaces, each of which is defined through an inequality constraint. To use Theorem 2.6, the associated null space is the intersection of the following three inequalities:

$$\begin{aligned} p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{01v} &\leq 0, \\ p^A_{11v} - p^A_{10v} - p^A_{01v} + p^A_{00v} + p^C_{00v} - p^C_{10v} &\leq 0, \\ p^A_{11v} - p^A_{10v} - p^A_{01v} &\leq 0. \end{aligned}$$

A similar situation arises in falsification of the binary instrumental variable model (Ramsahai et al., 2011). The correct p-value depends upon where the true parameter lies on the boundary of the null space (Ramsahai et al., 2011). If the true parameter lies only on the boundary of one of the inequality constraints, then the correct p-value is $P(\chi^2_1 > t)/2$ (Ramsahai et al., 2011). While it is clear what is the correct asymptotic sampling distribution if the true parameter lies on the boundary of multiple inequality constraints in the context of the falsification of binary instrumental variable model (Ramsahai et al., 2011), it is not always the case that the statistician will be able to easily derive the associated asymptotic sampling distribution when there is more than one inequality constraint that defines the null space. This is particularly true in the setting where there are a large number of inequality constraints.

For our data analysis, the only instance where we examine a null space that is defined through multiple inequality constraints stems from Theorem 2.6. For these types hypothesis test, the asymptotic sampling distribution changes upon where one assumes the true parameter lies on the null space. To construct a p-value based upon Theorem 2.6, we use Theorem 3 of Self and Liang (1987) to find that the asymptotic sampling distribution is $w_{0,3}\chi^2_0 + w_{1,3}\chi^2_1 + w_{2,3}\chi^2_2 + w_{3,3}\chi^2_3$ if the true parameter lies on the boundary of all three half-spaces. The weights can be calculated using equations (4.8) and (4.9) in Shapiro (1985). If the true parameter lies only the boundary of only two of the three half spaces, then the asymptotic sampling distribution is given by $w_{0,2}\chi^2_0 + w_{1,2}\chi^2_1 + w_{2,2}\chi^2_2$. Finally, if the true parameter lies only the boundary of only one of the three half spaces, then we can use the the earlier p-value of $P(\chi^2_1 > t)/2$. To get one p-value, one can use the least favorable configuration p-value $= \sup_{\mathbf{p}\in\mathbf{p}_0} P(T > t)$, where T is the likelihood-ratio test statistic as defined in the online supplement, t is the observed test statistic, $\mathbf{p}$ is the parameter space defined in the online supplement, and $\mathbf{p}_0$ is the null parameter space defined by the relevant inequality constraints. The statistics and econometrics literature for calculating the asymptotic sampling distribution of test statistics in the analysis of moment structures under inequality constraints is vast (Drton, 2009; Geyer et al.,

1994; Shapiro, 1985; Wolak, 1991; Dardanoni and Forcina, 1998; Silvapulle and Sen, 2011). In the situation where analytic formulas to calculate the weights of the χ^2 distributions are not available, Monte-Carlo methods can used to determine the weights to a pre-specified degree of precision (Dardanoni and Forcina, 1998; Silvapulle and Sen, 2011). The likelihood and likelihood ratio test statistic are provided in the online supplement.

3. Application to HIV Drug Resistance

HIV drug resistance arises when viral mutations make particular drugs less effective in controlling HIV (Tang and Shafer, 2012). For our data analysis, consider the data from the Stanford HIV drug resistance database on 484 viral isolates in Table 1, which presents information on viral resistance to the NRTI Etravirine (Tang and Shafer, 2012). The two viral mutations under consideration are X_1, which denotes presence of mutation 103 R and X_2 which denotes presence of mutation 179 D. The scientific question under consideration is whether mutations 103 R and 179 D interact synergystically to confer drug resistance to Etravirine. Our outcome Y is an ordinal outcome with three levels: no drug resistance, partial drug resistance, and full drug resistance. Here, no drug resistance is labeled 0, partial drug resistance is labeled 1, and full drug resistance is labeled 2. We

Table 1: Drug Resistance to Etravirine by mutation category

$X_1=0,X_2=0$			$X_1=0,X_2=1$			$X_1=1,X_2=0$			$X_1=1,X_2=1$		
$n_{0,0}$	P	F	$n_{0,1}$	P	F	$n_{1,0}$	P	F	$n_{1,1}$	P	F
445	74	57	17	2	1	10	2	1	12	7	4

Here, $n_{x_1,x_2} = \sum_{i=1}^{n} I(X_{1i} = x_1, X_{2i} = x_2)$. P denotes individuals with partial drug resistance, and F denotes full drug resistance.

Table 2: Likelihood ratio test of drug resistance

Outcome	Null Hypothesis	LRT	p-value
$I(Y=1)$	$p_{11}^A - p_{10}^A - p_{01}^A \leq 0$	1.477	0.112
$I(Y\geq 1)$	$p_{11}^B - p_{10}^B - p_{01}^B \leq 0$	4.218	0.020
$I(Y=2)$	$p_{11}^C - p_{10}^C - p_{01}^C \leq 0$	0.925	0.168

assume that there is no confounding between Y and X_1, X_2. This means that $Y_{x_1x_2}$ II (X_1, X_2). A laboratory experiment provides a contingency table, summarized in Table 1, on HIV drug resistance by mutation category (Tang and Shafer, 2012).

Table 2 provides the likelihood ratio test statistics and associated p-values for assessing sufficient cause interaction for each of the specified outcome levels $I(Y = 1)$, $I(Y \geq 1)$, and $I(Y = 2)$. Table 3 provides the same information, but assumes that the effects of X_1 and X_2 are positive

Table 3: Likelihood ratio test of drug resistance under monotonicity

Outcome	Null Hypothesis	LRT	p-value
	$p^A_{11} - p^A_{10} - p^A_{01} \leq 0$		
$I(Y=1)$	$p^A_{11} - p^A_{10} - p^A_{01} + p^A_{00} + p^C_{00} - p^C_{01} \leq 0$	4.704	$0.057^\star$
	$p^A_{11} - p^A_{10} - p^A_{01} + p^A_{00} + p^C_{00} - p^C_{10} \leq 0$		
$I(Y \geq 1)$	$p^B_{11} - p^B_{10} - p^B_{01} + p^B_{00} \leq 0$	10.624	< 0.005
$I(Y=2)$	$p^C_{11} - p^C_{10} - p^C_{01} + p^C_{00} \leq 0$	2.585	0.054

The p-value annotated with a $\star$ is obtained under the least favorable configuration.

monotonic for Y. If one is unwilling to make any monotonicity assumptions, there is no statistical evidence that X_1 and X_2 have a synergistic effect on $Y^C = I(Y = 2)$ or $Y^A = I(Y = 1)$, but there is some evidence for as synergistic effect on $Y^B = I(Y \geq 1)$. If one were willing to assume that X_1 and X_2 have positive monotonic effects on the outcome, the evidence for a synergistic effect of X_1 and X_2 on specified outcome $I(Y = 1)$ and $I(Y = 2)$ is stronger, although the p-values are slightly above the nominal 0.05 rejection threshold. In this situation, since we have evidence that X_1 and X_2 have a synergistic effect on $I(Y \geq 1)$, and thus also that X_1 and X_2 have a synergistic effect on either or both of the outcomes $I(Y = 1)$ and $I(Y = 2)$, since for $p^B_{11} - p^B_{10} - p^B_{01} > 0$, either or both of these two

inequalities $p_{11}^C - p_{10}^C - p_{01}^C > 0$ or $p_{11}^A - p_{10}^A - p_{01}^A > 0$ needs to hold. Yet, given the current sample size, we are unable to detect whether that synergistic effect occurs either for $I(Y = 1)$ or $I(Y = 2)$.

Discussions on whether the proposed exposures have monotonic effects on the outcome should occur with the scientific investigators. Such assumptions could have scientific justification, and would enable researchers to use less stringent conditions to draw the same inferences. Primary mutations, such as the two investigated in this paper, "directly decrease the susceptibility of the virus to an antiretroviral treatment"(Tang and Shafer, 2012). To the best of our knowledge, it is not known if mutation 103 R and 179 D are never preventative for partial or full drug resistance for every individual taking Etravirine as part of their treatment for HIV. Monotonicity assumptions can be falsified from the data, but they are never completely verifiable.

4. Generalizations and Extensions

In this section, we allow our ordinal outcome Y to take values $Y \in \{0, 1, \ldots, j\}$. If we want to investigate whether putative binary causes X_1 and X_2 have synergistic effects on outcome $Y^y = I(Y \geq y)$, where $0 < y \leq j$, then assuming no confounding between putative causes X_1 and X_2 on Y^y, we

need to check that $P(Y \geq y \mid X_1 = 1, X_2 = 1) - P(Y \geq y \mid X_1 = 1, X_2 = 0) - P(Y \geq y \mid X_1 = 0, X_2 = 1) > 0$. The proof of this result is similar to the proof of Theorem 2.3 and is provided in the online supplement.

Theorem 4.1. *If we can assume that X_1 and X_2 have positive monotonic effects on Y, then $P(Y \geq y \mid X_1 = 1, X_2 = 1) - P(Y \geq y \mid X_1 = 1, X_2 = 0) - P(Y \geq y \mid X_1 = 0, X_2 = 1) + P(Y \geq y \mid X_1 = 0, X_2 = 0) > 0$ implies X_1 and X_2 display synergism for $I(Y \geq 1)$.*

Let $S \subseteq \{1, 2, \ldots, n-1\}$, $Y^S = I(Y \in S)$, $S^+ = \max(S)$ and $Y^{S^+} = I(Y > S^+)$. If we wish to check X_1 and X_2 have synergistic effects on $Y^S = I(Y \in S)$, where S is an arithmetic progression with common difference one, and both X_1 and X_2 have positive monotonic effects on Y, then we need to check whether if at least one of the following three inequalities hold:

$$\begin{aligned} & P(Y \in S \mid X_1 = 1, X_2 = 1) \\ & -P(Y \in S \mid X_1 = 1, X_2 = 0) \\ & -P(Y \in S \mid X_1 = 0, X_2 = 1) > 0, \end{aligned}$$

$$\begin{aligned} P(Y \in S \mid X_1 = 1, X_2 = 1) \quad & - \quad P(Y \in S \mid X_1 = 1, X_2 = 0) \\ -P(Y \in S \mid X_1 = 0, X_2 = 1) \quad & + \quad P(Y \in S \mid X_1 = 0, X_2 = 0) \\ +P(Y > S^+ \mid X_1 = 0, X_2 = 0) \quad & - \quad P(Y > S^+ \mid X_1 = 1, X_2 = 0) > 0, \end{aligned}$$

$$\begin{aligned} P(Y \in S \mid X_1 = 1, X_2 = 1) \quad &- \quad P(Y \in S \mid X_1 = 1, X_2 = 0) \\ -P(Y \in S \mid X_1 = 0, X_2 = 1) \quad &+ \quad P(Y \in S \mid X_1 = 0, X_2 = 0) \\ +P(Y > S^+ \mid X_1 = 0, X_2 = 0) \quad &- \quad P(Y > S^+ \mid X_1 = 0, X_2 = 1) > 0, \end{aligned}$$

Here, the second and third inequalities are implied by $P(Y \in S \mid X_1 = 1, X_2 = 1) - P(Y \notin S \mid X_1 = 1, X_2 = 1) - P(Y \in S \mid X_1 = 1, X_2 = 0) - P(Y \in S \mid X_1 = 0, X_2 = 1) + P(Y \in S \mid X_1 = 0, X_2 = 0) > 0$. This last inequality can be shown to be equivalent to $2 \cdot P(Y \in S \mid X_1 = 1, X_2 = 1) - P(Y \in S \mid X_1 = 1, X_2 = 0) - P(Y \in S \mid X_1 = 0, X_2 = 1) + P(Y \in S \mid X_1 = 0, X_2 = 0) > 1$. If X_1 or X_2 do not have positive monotonic effects on Y, or S is not an arithmetic progression with common difference one, then only the first inequality out of the three listed above remains valid for determining if X_1 and X_2 have synergistic effects on $Y^S = I(Y \in S)$. The full set of definitions, theorems, and proofs associated with this generalization are provided in the online supplement.

For this section, we allow $(Y_{11}, Y_{10}, Y_{01}, Y_{00})$ to have a distribution function $P(Y_{11} \in y_{11}, Y_{10} \in y_{10}, Y_{01} \in y_{01}, Y_{00} \in y_{00})$, where $y_{11}, y_{10}, y_{01}, y_{00}$ are all subsets of $\mathbf{R}$. Proofs of next two results are provided in the online supplement.

Definition 4.1 (Generalized Positive Monotonicity)**.** We say that X_1 has a positive monotonic effect on $Y \in y_c$ for any fixed $y_c \subset \mathbf{R}$ if there is no individual $\omega \in \Omega$ such that $Y_{x_1x_2}(\omega) \notin y_c$ and $Y_{x_3x_2}(\omega) \in y_c$ for all $x_1 > x_3$ for any fixed x_2. Similarly, we say that X_2 has a positive monotonic effect on $Y \in y_c$ for some $y_c \subset \mathbf{R}$ if there is no individual $\omega \in \Omega$ such that $Y_{x_1x_2}(\omega) \notin y_c$ and $Y_{x_1x_3}(\omega) \in y_c$ for all $x_2 > x_3$ for any fixed x_1. If X_1 and X_2 each individually have a positive monotonic effect on $Y \in y_c$ for any fixed $y_c \subset \mathbf{R}$ then we say that X_1 and X_2 have positive monotonic effects on $Y \in y_c$.

Theorem 4.2. *Suppose* $Y_{x_1x_2} \amalg X_1X_2$. *Here,* y_a *is any subset of* $\mathbf{R}$. *The contrast*

$$P\left(Y \in y_a \mid X_1 = 1, X_2 = 1\right) - P\left(Y \in y_a \mid X_1 = 1, X_2 = 0\right) - P\left(Y \in y_a \mid X_1 = 0, X_2 = 1\right)$$

is equal to

$$P\left(Y_{11} \in y_a, Y_{10} \notin y_a, Y_{01} \notin y_a\right) - P\left(Y_{11} \notin y_a, Y_{10} \in y_a, Y_{01} \notin y_a\right) - P\left(Y_{11} \in y_a, Y_{10} \in y_a, Y_{01} \in y_a\right) - P\left(Y_{11} \notin y_a, Y_{10} \notin y_a, Y_{01} \in y_a\right).$$

Theorem 4.3. *Suppose* $Y_{x_1x_2} \amalg X_1X_2$ *and suppose* X_1 *and* X_2 *have positive*

monotonic effects on $Y \in y_c$. *For any* y_c *that is a subset of* $\mathbf{R}$, *the contrast*

$$P\left(Y \in y_c \mid X_1 = 1, X_2 = 1\right) - P\left(Y \in y_c \mid X_1 = 1, X_2 = 0\right)$$
$$- P\left(Y \in y_c \mid X_1 = 0, X_2 = 1\right) + P\left(Y \in y_c \mid X_1 = 0, X_2 = 0\right)$$

is equal to $P\left(Y_{11} \in y_c, Y_{10} \notin y_c, Y_{01} \notin y_c, Y_{00} \notin y_c\right)$.

5. Discussion

Our extension of the sufficient cause model to ordinal and categorical outcomes enables researchers to investigate more complex scientific questions. In addition, we derive novel empirical conditions that in some situations would be more powerful in testing sufficient cause interaction for ordinal outcomes than applying the previously formulated empirical conditions for a binary outcome to a dichotomized ordinal outcome. The interpretations of sufficient cause interaction are far stronger than the corresponding interpretations of tests for statistical interaction. We applied these novel tests for sufficient cause interaction to detect whether viral mutation 103R and 179D interacted synergystically to confer partial, full or any drug resistance to Etravirine.

6. Appendix

6.1 Appendix 1: Notation

We denote $Y^A = I(Y \in \{1\})$, $Y^B = I(Y \in \{1,2\})$, $Y^C = I(Y \in \{2\})$, $Y^D = I(Y \in \{0\})$, $Y^E = I(Y \in \{0,2\})$, $Y^F = I(Y \in \{0,2\})$. Potential outcome versions of Y^A, Y^B, Y^C, Y^D, Y^E, and Y^F are defined as $Y^A_{x_1,\ldots,x_k}(\omega) = I(Y_{x_1,\ldots,x_k}(\omega) \in \{1\})$, $Y^B_{x_1,\ldots,x_k} = I(Y_{x_1,\ldots,x_k}(\omega) \in \{1,2\})$, $Y^C_{x_1,\ldots,x_k} = I(Y_{x_1,\ldots,x_s}(\omega) \in \{2\})$, $Y^D_{x_1,\ldots,x_s} = I(Y_{x_1,\ldots,x_k}(\omega) \in \{0\})$, $Y^E_{x_1,\ldots,x_s} = I(Y_{x_1,\ldots,x_k}(\omega) \in \{0,2\})$, and $Y^F_{x_1,\ldots,x_s} = I(Y_{x_1,\ldots,x_k}(\omega) \in \{0,1\})$.

Supplementary Materials

Definitions and theorems of the more general theorems are provided in the online supplement. Proofs of all Theorems and Corollaries are collected in the online supplement.

Acknowledgements

The authors acknowledge useful discussions and correspondence with Denis Agniel, Ron Bosch, Mathias Drton, James M. Robins, Robert W. Shafer, Eric J. Tchetgen Tchetgen, and Linbo Wang. This research was supported by the National Institutes of Health, U.S.A.

References

Berzuini, C. and A. P. Dawid (2016). Stochastic mechanistic interaction. *Biometrika 103*(1), 89–102.

Cayley, A. (1853). Xxxvii. note on a question in the theory of probabilities. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 6*(39), 259–259.

Dardanoni, V. and A. Forcina (1998). A unified approach to likelihood inference on stochastic orderings in a nonparametric context. *Journal of the American Statistical Association 93*(443), 1112–1123.

Drton, M. (2009). Likelihood ratio tests and singularities. *The Annals of Statistics 37*(2), 979–1012.

Geyer, C. J. et al. (1994). On the asymptotics of constrained m-estimation. *The Annals of Statistics 22*(4), 1993–2010.

Ramsahai, R., S. Lauritzen, et al. (2011). Likelihood analysis of the binary instrumental variable model. *Biometrika 98*(4), 987.

Ramsahai, R. R. (2013). Probabilistic causality and detecting collections of interdependence patterns. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 75*(4), 705–723.

Rothman, K. J. (1976). Causes. *American Journal of Epidemiology 104*(6), 587–592.

Self, S. G. and K.-Y. Liang (1987). Asymptotic properties of maximum likelihood estimators and

likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association 82*(398), 605–610.

Shapiro, A. (1985). Asymptotic distribution of test statistics in the analysis of moment structures under inequality constraints. *Biometrika 72*(1), 133–144.

Silvapulle, M. J. and P. K. Sen (2011). *Constrained statistical inference: Order, inequality, and shape constraints*, Volume 912. John Wiley & Sons.

Tang, M. W. and R. W. Shafer (2012). HIV-1 antiretroviral resistance. *Drugs 72*(9), e1–e25.

Van der Vaart, A. W. (2000). *Asymptotic statistics*, Volume 3. Cambridge university press.

VanderWeele, T. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.

Vanderweele, T. J. (2010). Sufficient cause interactions for categorical and ordinal exposures with three levels. *Biometrika 97*(3), 647–659.

VanderWeele, T. J. and T. S. Richardson (2012). General theory for interactions in sufficient cause models with dichotomous exposures. *Annals of statistics 40*(4), 2128.

VanderWeele, T. J. and J. M. Robins (2008). Empirical and counterfactual conditions for sufficient cause interactions. *Biometrika 95*(1), 49–61.

VanderWeele, T. J. and J. M. Robins (2012). Stochastic counterfactuals and stochastic sufficient causes. *Statistica Sinica 22*(1), 379.

Vansteelandt, S., T. J. VanderWeele, and J. M. Robins (2012). Semiparametric tests for suf-

ficient cause interaction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 74*(2), 223–244.

Wang, L., J. M. Robins, and T. S. Richardson (2017). On falsification of the binary instrumental variable model. *Biometrika 104*(1), 229–236.

Wolak, F. A. (1991). The local nature of hypothesis tests involving inequality constraints in nonlinear models. *Econometrica: Journal of the Econometric Society*, 981–995.

Harvard University, Department of Biostatistics, 677 Huntington Avenue, Kresge Building, Boston, Massachusetts, 02115

E-mail: (jaffer.zaidi@gmail.com)

Harvard University, Department of Epidemiology, Department of Biostatistics, 677 Huntington Avenue, Kresge Building, Boston, Massachusetts, 02115

E-mail: (tvanderw@hsph.harvard.edu)