Evaluation of “Probabilities of Causation” in Case–Control Studies: Identification and Estimation

Ryusei Shingaki[1] **Haruka Yoshida**[1] **Manabu Kuroki**[1]

[1]Department of Mathematics, Physics, Electrical Engineering and Computer Science, Graduate School of Engineering Science, Yokohama National University, Yokohama, Kanagawa, JAPAN

Abstract

“Probabilities of causation” play a crucial role in practical science. Pearl [2009] defined three types of probabilities of causation in the context of structural causal models: the probability of necessity (PN), the probability of sufficiency (PS), and the probability of necessity and sufficiency (PNS). Furthermore, Tian and Pearl [2000] and Kuroki and Cai [2011] provided the identification conditions for these probabilities under the assumption of monotonicity. However, these identification conditions are described based on “the joint probabilities of observed probabilities” and/or “causal risks”. Therefore, they are not applicable to case–control studies, in which the available statistical data are given in the form of the conditional probabilities of observed variables given an outcome variable. To address this limitation, this paper provides novel identification conditions for the probabilities of causation using (i) two proxy covariates and (ii) a proxy covariate together with an instrumental variable. Remarkably, the use of a proxy covariate and an instrumental variable enables the identification not only of the probabilities of causation but also of the joint distribution of the potential outcome variables. When these probabilities can be evaluated using the proposed identification conditions, new plug-in estimators of these probabilities are presented.

1 INTRODUCTION

A case–control study is a representative type of observational study used to assess whether an exposure variable is associated with an outcome variable. Indeed, a considerable number of textbooks and articles on case–control studies have been published [e.g. Breslow and Day, 1980, Keogh and Cox, 2014, Borgan et al., 2018].

Briefly, case–control studies start with constructing (i) a group of subjects called cases, who have the outcome of interest, and (ii) a group of subjects called controls, who are similar to the cases but do not have the outcome of interest. Researchers then retrospectively determine which subjects in each group were exposed, comparing the frequency of the exposure between the two groups.

A major drawback of case–control studies is the potential for recall bias, whereby subjects with the outcome of interest are more likely to recall and report the exposure status than those without the outcome. Due to their retrospective nature, case–control studies can be used to establish associations between exposure and outcome variables, but not causation. Nevertheless, there are many situations in which we wish to establish causation through case–control studies.

To address this issue, this paper focuses on the identification and estimation of probabilities of causation, which have received increasing attention over the past decades in artificial intelligence, epidemiology, legal reasoning, policy analysis, risk analysis, and statistical science [e.g., Byrne, 2019, Cai and Kuroki, 2005, Cox, 1984, Dawid et al., 2017, Galhotra et al., 2021, Greenland, 1987, Mothilal et al., 2021, Pearl, 2009, Robins and Greenland, 1989, Tian and Pearl, 2000, Watson et al., 2021]. For example, in explainable artificial intelligence (XAI), the notions of “necessity”, “sufficiency”, and “necessity and sufficiency” have been recognized as key components of successful explanations, and the probabilities of causation provide a probabilistic formalization of these concepts [Galhotra et al., 2021, Mothilal et al., 2021, Watson et al., 2021].

Pearl [2009] developed formal semantics for probabilities of causation based on structural causal models and introduced the probability of necessity (PN), the probability of sufficiency (PS), and the probability of necessity and sufficiency (PNS). In addition, Pearl [2009], Tian and Pearl [2000], and Kuroki and Cai [2011] showed how to bound these probabilities using statistical data. Pearl [2009] and Tian and

Proceedings of the 42nd Conference on Uncertainty in Artificial Intelligence (UAI 2026), PMLR 337:6261–6282.

Pearl [2000] also established their identifiability under the monotonicity assumption when the relevant causal risks are identifiable. From the viewpoint of the joint distribution of potential outcome variables, the identification and estimation of the probabilities of causation have also been studied by Kawakami et al. [2023], Shingaki and Kuroki [2021], and Shingaki and Kuroki [2023].

However, these studies are restricted to settings in which the available statistical data are given in the form of the joint distribution of observed variables and/or causal risks. Consequently, the existing identification conditions do not apply to case–control studies where the available data are provided in the form of conditional probabilities of observed variables given an outcome variable [Kuroki et al., 2010]. Although there is substantial demand for drawing causal conclusions from case—control data, the methodological development required to identify probabilities of causation in this setting remains limited.

To overcome this limitation, this paper provides novel identification conditions for probabilities of causation using (i) two proxy covariates and (ii) a proxy covariate together with an instrumental variable (IV). Here, covariates are variables that are not affected by the exposure variable, and proxy covariates are variables used in place of a set of covariates that cannot be measured directly. An instrumental variable is a variable that (i) is not associated with all observed and unobserved covariates, and (ii) is associated with the exposure variable but has no effect on the outcome variable except through its association with the exposure variable. Unlike Tian and Pearl [2000] and Kuroki and Cai [2011], the proposed identification conditions enable us to identify probabilities of causation without the monotonicity assumption. Notably, even when the probability of the outcome variable is unavailable, the use of an instrumental variable and a proxy covariate enables the identification not only of probabilities of causation but also of the joint distribution of the potential outcome variables.

When probabilities of causation are evaluated using the proposed identification conditions, new plug-in estimators of these probabilities are presented. Thus, the results of this paper extend the applicability of statistical causal inference to practical science. The proofs, some numerical experiments, and a case study application are provided in the Supplementary Material.

2 PRELIMINARIES

In this section, we introduce the potential outcome variables used to discuss our problems. We assume that readers are familiar with the basic theory of statistical causal inference [Imbens and Rubin, 2015, Pearl, 2009]. For further details on the basic theory of statistical causal inference and the graph-theoretic terminology used in this paper, we refer readers to Imbens and Rubin [2015] and Pearl [2009].

Let N denote the sample size. We assume that X and Y denote the observed dichotomous exposure and outcome variables, respectively. For $x \in \{x_0, x_1\}$ (where x_1 is the occurrence of an event and x_0 is the non-occurrence of an event) and $y \in \{y_0, y_1\}$ (where y_1 is the occurrence of an outcome and y_0 is the non-occurrence of an outcome), let $p(X = x, Y = y) = p(x, y)$ be the joint probability of $(X, Y) = (x, y)$, $p(X = x) = p(x)$ be the marginal probability of $X = x$, and $p(Y = y \mid X = x) = p(y \mid x) = p(x, y)/p(x)$ be the conditional probability of $Y = y$ given $X = x$. A similar notation is used for other probabilities. Then, in principle, for $x \in \{x_0, x_1\}$, each of the N subjects has a potential outcome $Y_x(i)$ that would have resulted if X had been x for the i-th subject. Here, $Y_x(i) = y$ denotes the counterfactual sentence that "Y would be y had X been x for the i-th subject." The potential outcome variable $Y_x(i)$ is observed only if X is x for the i-th subject, denoted as $X(i) = x$. This property is referred to as consistency [Pearl, 2009] and is formulated as

$$X(i) = x \quad \Longrightarrow \quad Y_x(i) = Y(i) \tag{1}$$

for the i-th subject.

In this paper, we assume the stable unit treatment value assumption [Imbens and Rubin, 2015], which can be summarized as follows: (i) the exposure status of any subject does not affect the outcome status of the other subjects, and (ii) the exposure status of all subjects is comparable. Then, when N subjects in the study are considered as random samples from the population of interest, $Y_x(i)$ can be regarded as a realization of the random variable Y_x. Then, the causal risk of $X = x$ on $Y = y$ is defined as $p(Y_x = y)$.

When an ideal randomized experiment on X is feasible, it is known that X is independent of both Y_{x_0} and Y_{x_1}. Then, the causal risk $p(Y_x = y)$ is identifiable and given by $p(Y_x = y) = p(y \mid x)$. Here, "identifiable" means that the causal quantities, such as $p(Y_x = y)$, can be estimated consistently from the distribution of observed variables. In contrast, when it is difficult to conduct ideal randomized experiments, we can still evaluate the causal risks according to the conditionally ignorable treatment assignment condition [Imbens and Rubin, 2015] or, graphically, the back-door criterion [Pearl, 2009]. In other words, for the exposure variable X, if there exists a set $\boldsymbol{Z}$ of observed variables such that X is conditionally independent of (Y_{x_0}, Y_{x_1}) given $\boldsymbol{Z}$, then we say that treatment assignment is conditionally ignorable given $\boldsymbol{Z}$. In this case, the causal risks $p(Y_x = y)$ are identifiable and given by

$$p(Y_x = y) = \sum_{\boldsymbol{z}} p(y \mid x, \boldsymbol{z})p(\boldsymbol{z}). \tag{2}$$

Although other identification conditions can be used to evaluate causal risks [e.g., Pearl, 2009, Tian and Pearl, 2002], we do not cover them in this paper owing to space constraints.

According to Pearl [2009], we define three probabilities of causation, namely, the probability of necessity (PN), the probability of sufficiency (PS), and the probability of necessity and sufficiency (PNS).

The PN is defined as

$$\mathrm{PN} = p(Y_{x_0} = y_0 \,|\, x_1, y_1), \tag{3}$$

which represents the probability that the outcome of interest would not have occurred $(Y = y_0)$ in the absence of the event of interest ($X = x_0$), given that $X = x_1$ and $Y = y_1$ did occur. The PN has applications in artificial intelligence, epidemiology, and legal reasoning. Epidemiologists have long been concerned with estimating the probability that a certain case of a disease is attributable to a particular exposure, which is usually interpreted counterfactually as "the probability that the disease would not have occurred in the absence of exposure, given that disease and exposure did in fact occur." Such a probability is evaluated by the PN.

The PS is defined as

$$\mathrm{PS} = p(Y_{x_1} = y_1 \,|\, x_0, y_0), \tag{4}$$

which stands for the probability that the outcome of interest would have occurred $(Y = y_1)$ in the presence of the event of interest ($X = x_1$), given that $X = x_0$ and $Y = y_0$ did, in fact, occur. The PS has applications in artificial intelligence, policy analysis, and risk analysis. Policy-makers may well be interested in the dangers that a certain exposure may present to the healthy population [Khoury et al., 1989]. Counterfactually, this notion is interpreted as the "probability that a healthy unexposed subject would have gotten the disease had they been exposed." Such a probability is evaluated by the PS.

The PNS is defined as

$$\mathrm{PNS} = p(Y_{x_0} = y_0, Y_{x_1} = y_1), \tag{5}$$

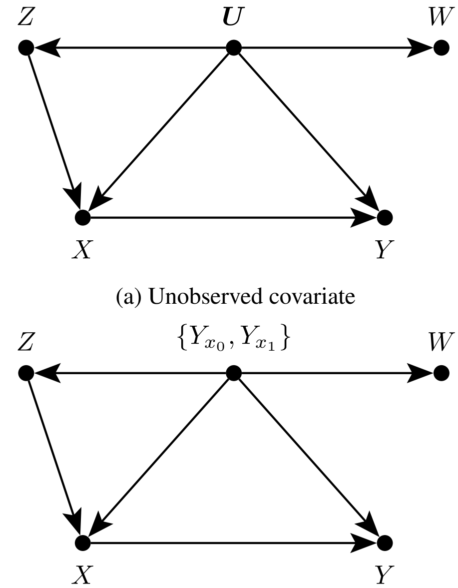
which represents the probability that $X = x_1$ is a necessary and sufficient cause for $Y = y_1$. Therefore, it measures both the sufficiency and the necessity of $X = x_1$ to produce $Y = y_1$.

Since PN, PS, and PNS involve joint probabilities of two potential outcome variables, they are not estimable from statistical data, even under ideal randomized experiments, without any further causal information [Pearl, 2009].

3 IDENTIFICATION

3.1 TWO PROXY COVARIATES

According to Pearl [2009], letting $\boldsymbol{U}$ be the set of all discrete and continuous covariates that could affect X and Y, both observed and unobserved, we consider the identification problem of the probabilities of causation based on the directed graph shown in Figure 1a.



(a) Unobserved covariate

(b) Potential outcomes

Figure 1: Graphical representation for the case of two proxy covariates

Covariates Z and W, in Figure 1a, are measured as proxy covariates of $\boldsymbol{U}$. Here, note that both Z and W can be a set of discrete and/or continuous variables. In Figure 1a, the directed edge from X to Y $(X \to Y)$ indicates that X could affect Y. In addition, the absence of a directed edge from Y to X indicates that Y cannot be a cause of X, and the directed path from $\boldsymbol{U}$ to Y through X ($\boldsymbol{U} \to X \to Y$) indicates that some elements of $\boldsymbol{U}$ could affect Y mediated by X. Figure 1a also graphically represents the data-generating process

$$\begin{aligned} Y = g_y(X, \boldsymbol{U}, \epsilon_y), \quad X = g_x(\boldsymbol{U}, Z, \epsilon_x), \\ Z = g_z(\boldsymbol{U}, \epsilon_z), \quad W = g_w(\boldsymbol{U}, \epsilon_w), \end{aligned} \tag{6}$$

where ϵ_x, ϵ_y, ϵ_z, and ϵ_w are independent random disturbances and they are also independent of $\boldsymbol{U}$. When the structural equation models, such as equation (6), are used to represent the data-generating process, the corresponding directed graph shown in Figure 1a is called a causal diagram.

The situation, such as Figure 1a, is also discussed in Kuroki and Pearl [2014] and Pearl [2010]. However, in this paper, different from Kuroki and Pearl [2014] and Pearl [2010], $\boldsymbol{U}$ can include an uncertain number of all discrete and continuous covariates. Thus, in many situations, it is reasonable to assume the existence of proxy covariates Z and W that W is conditionally independent of $\{X, Y, Z\}$ given $\boldsymbol{U}$, and Z is conditionally independent of Y given $\boldsymbol{U}$ and X. Here, note that such independence assumptions between two observed variables would be affected by partitioning the states of $\boldsymbol{U} \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$.

In the situation shown in Figure 1a, irrespective of the complexity of $\boldsymbol{U} \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$, the impact of $\boldsymbol{U}$ on Y is restricted to the modification of the functional relationships between X and Y. This yields four functions for two dichotomous variables X and Y, and thus the value taken by $\boldsymbol{U} \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$ selects one of these four functions [Pearl, 2009]. Considering these observations, according to Lash et al. [2021], the states of $\boldsymbol{U} \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$ are divided into the following four potential outcome types:

$u_1 = (Y_{x_1} = y_1, Y_{x_0} = y_1)$ represents a "doomed" situation where the exposure status is irrelevant in the sense that the outcome of interest occurs with or without the event of interest.

$u_2 = (Y_{x_1} = y_1, Y_{x_0} = y_0)$ represents a "causative" situation where the outcome of interest occurs if and only if the event of interest occurs.

$u_3 = (Y_{x_1} = y_0, Y_{x_0} = y_1)$ represents a "preventive" situation where the outcome of interest occurs if and only if the event of interest does not occur.

$u_4 = (Y_{x_1} = y_0, Y_{x_0} = y_0)$ represents an "immune" situation where the exposure status is irrelevant in the sense that the outcome of interest does not occur with or without the event of interest.

Here, the assumption of $p(u_3) = 0$ is called monotonicity by Pearl [2009]. However, note that the central aim of this paper is to solve the identification and estimation problems of the probabilities of causation without such an assumption.

According to this partition of the states of $\boldsymbol{U} \cup \{\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_w\}$, it is redefined as a variable U taking a value in $\{u_1, u_2, u_3, u_4\}$. Then, similar to the setting of the noncompliance problem by Pearl [2009, ch.8], throughout the paper, we assume that the recursive factorization of the positive joint probability of X, Y, Z, and W, $p(x, y, z, w)$, is given as

$$\begin{aligned} &p(x, y, z, w) \\ &= \sum_{i=1}^{4} p(y \mid x, u_i) p(x \mid z, u_i) p(z \mid u_i) p(w \mid u_i) p(u_i) \end{aligned} \tag{7}$$

according to Figure 1b. It is also assumed that other joint probabilities of X, Y, Z, and W can be rewritten similarly. In addition, the consistency property implies that

$$\begin{aligned} p(x_1, z, w, u_i \mid y_0) = 0 \quad &\text{for } i = 1, 2, \\ p(x_1, z, w, u_i \mid y_1) = 0 \quad &\text{for } i = 3, 4, \\ p(x_0, z, w, u_i \mid y_0) = 0 \quad &\text{for } i = 1, 3, \\ p(x_0, z, w, u_i \mid y_1) = 0 \quad &\text{for } i = 2, 4. \end{aligned} \tag{8}$$

Thus, for example, the conditional probabilities $p(x_1, z, w \mid y_1)$ and $p(z, w \mid y_1)$ can be rewritten as

$$\begin{aligned} &p(x_1, z, w \mid y_1) \\ &= \sum_{i=1}^{2} p(x_1 \mid y_1, z, u_i) p(z \mid y_1, u_i) p(w \mid u_i) p(u_i \mid y_1), \end{aligned} \tag{9}$$

$$p(z, w \mid y_1) = \sum_{i=1}^{3} p(z \mid y_1, u_i) p(w \mid u_i) p(u_i \mid y_1). \tag{10}$$

In addition, the other joint and marginal probabilities of X, Y, Z, and W can be rewritten similarly.

Under the preparation above, our problem is to evaluate the probabilities of causation using conditional probabilities $p(x, z, w \mid y)$ without relying on the marginal probabilities $p(y)$ or having access to the joint probabilities $p(x, y, z, w)$. To solve the problem, we require the following conditions:

Condition 1 *the positive conditional probabilities $p(x, z, w \mid y)$ of $(X, Z, W) = (x, z, w)$ given $Y = y$ are available for $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$, and $w \in \{w_1, w_2, w_3\}$.*

Condition 2 *for $p(x, y, z, w \mid u) > 0$,*

$$p(x, y, z, w \mid u) = p(x, y, z \mid u) p(w \mid u), \tag{11}$$
$$p(y, z \mid x, u) = p(y \mid x, u) p(z \mid x, u), \tag{12}$$

hold for $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$, $w \in \{w_1, w_2, w_3\}$ and $u \in \{u_1, u_2, u_3, u_4\}$.

Condition 3

$$\begin{pmatrix} 1 & p(w \mid x, y) \\ p(z \mid x, y) & p(z, w \mid x, y) \end{pmatrix} \tag{13}$$

is invertible for $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$, and $w \in \{w_1, w_2, w_3\}$.

Condition 4

$$\begin{pmatrix} 1 & -\dfrac{p(z, w_1 \mid x, y) - p(z \mid x, y) p(w_1 | x, y)}{p(z, w_2 \mid x, y) - p(z \mid x, y) p(w_2 \mid x, y)} \\ 1 & -\dfrac{p(z, w_1 \mid x', y') - p(z \mid x', y') p(w_1 \mid x', y')}{p(z, w_2 \mid x', y') - p(z \mid x', y') p(w_2 \mid x', y')} \end{pmatrix} \tag{14}$$

is invertible for $(x, y, x', y') = (x_1, y_1, x_0, y_1)$, (x_1, y_1, x_0, y_0), (x_1, y_0, x_0, y_1), (x_1, y_0, x_0, y_0) and $z \in \{z_1, z_2\}$.

Conditions 3 and 4 are essentially non-degeneracy requirements on the proxies Z and W, ensuring that they carry sufficiently rich and non-redundant information about the latent structure $\boldsymbol{U}$. More intuitively, Condition 3 requires that different latent strata induce distinct joint distributions of (Z, W), since otherwise they cannot be separated. Condition 4 is a stronger requirement ensuring sufficient joint variation of (Z, W) to distinguish strata within each cell, analogous to rank/completeness conditions in proxy-variable literature [e.g., Kuroki and Pearl, 2014, Miao et al., 2018]. Roughly, the conditions tend to hold when Z and W reflect genuinely different aspects of the underlying factor, and tend to fail when the proxies are weakly informative or nearly redundant.

The identification result is obtained in two steps. We first identify the conditional probabilities $p(w \mid u)$ and then identify the conditional probabilities $p(u \mid x, y)$.

Lemma 1 *Let Z and W be variables taking the values $z \in \{z_1, z_2\}$ for Z and $w \in \{w_1, w_2, w_3\}$ for W, respectively. Under Conditions 1–4, the conditional probabilities $p(w \mid u)$ are identified by*

$$\begin{aligned}
\begin{pmatrix} p(w_1 \mid u_1) \\ p(w_2 \mid u_1) \end{pmatrix} &= \begin{pmatrix} 1 & -a_{x_1 y_1} \\ 1 & -a_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_1} \\ b_{x_0 y_1} \end{pmatrix}, \\
\begin{pmatrix} p(w_1 \mid u_2) \\ p(w_2 \mid u_2) \end{pmatrix} &= \begin{pmatrix} 1 & -a_{x_1 y_1} \\ 1 & -a_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_1} \\ b_{x_0 y_0} \end{pmatrix}, \\
\begin{pmatrix} p(w_1 \mid u_3) \\ p(w_2 \mid u_3) \end{pmatrix} &= \begin{pmatrix} 1 & -a_{x_1 y_0} \\ 1 & -a_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_0} \\ b_{x_0 y_1} \end{pmatrix}, \\
\begin{pmatrix} p(w_1 \mid u_4) \\ p(w_2 \mid u_4) \end{pmatrix} &= \begin{pmatrix} 1 & -a_{x_1 y_0} \\ 1 & -a_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_0} \\ b_{x_0 y_0} \end{pmatrix},
\end{aligned} \tag{15}$$

where

$$a_{xy} = \frac{p(z, w_1 \mid x, y) - p(z \mid x, y) p(w_1 \mid x, y)}{p(z, w_2 \mid x, y) - p(z \mid x, y) p(w_2 \mid x, y)},$$
$$b_{xy} = \frac{p(z, w_2 \mid x, y) p(w_1 \mid x, y) - p(z, w_1 \mid x, y) p(w_2 \mid x, y)}{p(z, w_2 \mid x, y) - p(z \mid x, y) p(w_2 \mid x, y)}.$$

Using the identification formulas in Lemma 1, we obtain the following identification formulas for $p(u \mid x, y)$.

Lemma 2 *Under Conditions 1–4, the conditional probabilities $p(u \mid x, y)$ are identified by*

$$\begin{aligned}
\begin{pmatrix} p(u_1 \mid x_1, y_1) \\ p(u_2 \mid x_1, y_1) \end{pmatrix} &= \begin{pmatrix} 1 & p(w_1 \mid u_1) \\ 1 & p(w_1 \mid u_2) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ p(w_1 \mid x_1, y_1) \end{pmatrix}, \\
\begin{pmatrix} p(u_1 \mid x_0, y_1) \\ p(u_3 \mid x_0, y_1) \end{pmatrix} &= \begin{pmatrix} 1 & p(w_1 \mid u_1) \\ 1 & p(w_1 \mid u_3) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ p(w_1 \mid x_0, y_1) \end{pmatrix}, \\
\begin{pmatrix} p(u_3 \mid x_1, y_0) \\ p(u_4 \mid x_1, y_0) \end{pmatrix} &= \begin{pmatrix} 1 & p(w_1 \mid u_3) \\ 1 & p(w_1 \mid u_4) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ p(w_1 \mid x_1, y_0) \end{pmatrix}, \\
\begin{pmatrix} p(u_2 \mid x_0, y_0) \\ p(u_4 \mid x_0, y_0) \end{pmatrix} &= \begin{pmatrix} 1 & p(w_1 \mid u_2) \\ 1 & p(w_1 \mid u_4) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ p(w_1 \mid x_0, y_0) \end{pmatrix},
\end{aligned} \tag{16}$$

where the notation $-\top$ stands for a transposed inverse matrix.

From Lemma 2 and the definitions of the potential outcome types, we obtain the following theorem.

Theorem 1 *Under Conditions 1–4, $p(Y_x = y \mid x', y')$ are identifiable for $x, x' \in \{x_0, x_1\}$ and $y, y' \in \{y_0, y_1\}$ as follows:*

$$p(Y_{x_0} = y_0 \mid x_1, y_1) = p(u_2 \mid x_1, y_1), \tag{17}$$
$$p(Y_{x_1} = y_0 \mid x_0, y_1) = p(u_3 \mid x_0, y_1), \tag{18}$$
$$p(Y_{x_0} = y_1 \mid x_1, y_0) = p(u_3 \mid x_1, y_0), \tag{19}$$
$$p(Y_{x_1} = y_1 \mid x_0, y_0) = p(u_2 \mid x_0, y_0). \tag{20}$$

In particular,

$$\mathrm{PN} = p(u_2 \mid x_1, y_1), \quad \mathrm{PS} = p(u_2 \mid x_0, y_0). \tag{21}$$

The proofs of Lemmas 1–2 and Theorem 1 are provided in Supplementary Material A. Note that the PNS is not identifiable by Theorem 1. However, since

$$\mathrm{PNS} = \mathrm{PN} \times p(x_1 \mid y_1) p(y_1) + \mathrm{PS} \times p(x_0 \mid y_0) p(y_0), \tag{22}$$

where $p(y_1) \geq 0$, $p(y_0) \geq 0$, and $p(y_1) + p(y_0) = 1$, and Tian-Pearl bounds [Tian and Pearl, 2000]

$$\begin{aligned} 0 \leq \mathrm{PNS} &\leq p(x_1, y_1) + p(x_0, y_0) \\ &\leq \max \{p(x_1|y_1), p(x_0|y_0)\}, \end{aligned} \tag{23}$$

the bounds are given by

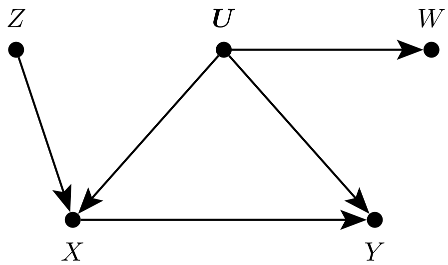
$$\begin{aligned} &\max \{0, \min \{\mathrm{PN} \times p(x_1 \mid y_1), \mathrm{PS} \times p(x_0 \mid y_0)\}\} \\ &\leq \mathrm{PNS} \leq \min \{1, \max \{p(x_1|y_1), p(x_0|y_0)\}, \\ &\qquad \max \{0, \mathrm{PN} \times p(x_1 \mid y_1), \mathrm{PS} \times p(x_0 \mid y_0)\}\}. \end{aligned} \tag{24}$$

Especially, when $p(y_1)$ is additionally available, the joint probabilities of potential outcome variables $p(u_1)$, $p(u_2)$, $p(u_3)$, $p(u_4)$ are identifiable, that is, the PNS is also identifiable.

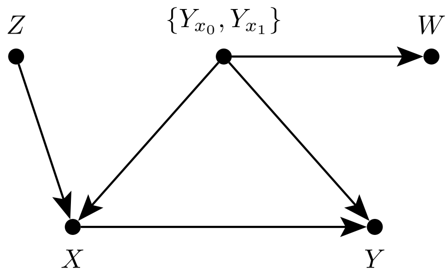
3.2 ONE PROXY COVARIATE TOGETHER WITH ONE INSTRUMENTAL VARIABLE

In this section, consider the situation shown in Figure 2a, in which $\boldsymbol{U}$ is observed through a proxy covariate W, and assume that Figure 2a is re-described as Figure 2b by a similar procedure to the previous section. Then, since Theorem 1 requires us to observe two proxy covariates for $\boldsymbol{U}$ in order to identify the probabilities of causation, some may think that these probabilities are not identifiable because Z is not a proxy covariate for $\boldsymbol{U}$. However, we show that this is not the case.

To see this, in Figure 2a, note that Z satisfies the instrumental variable (IV) assumptions: (i) Z is not independent of X, (ii) Z is independent of $\boldsymbol{U}$, and (iii) Z is conditionally independent of Y given $\boldsymbol{U}$ and X [Greenland, 2000]. Then, the variable Z satisfying the IV assumptions is called an IV. Our idea is to use Z as one of the proxy covariates of Theorem 1 to solve the identification problem and recover the probability of the outcome variable, but not as the statistical tool to derive the IV estimator.



(a) Unobserved Covariate



(b) Potential outcomes

Figure 2: Graphical representation for the case of proxy covariate and one instrumental variable

To do this, when Condition 1 holds, Condition 2 is satisfied for $p(x, y, z, w \mid u) > 0$ ($x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$, $w \in \{w_1, w_2, w_3\}$, and $u \in \{u_1, u_2, u_3, u_4\}$) in Figure 2b (graphically, Z and U are d-connected given X). Thus, noting that equations (11) and (12) hold, if both Conditions 3 and 4 of Theorem 1 hold, $p(Y_x = y \mid x', y')$ is identifiable for $x, x' \in \{x_0, x_1\}$ and $y, y' \in \{y_0, y_1\}$. In addition, although the probability of the outcome variable is not available, the independence between the IV and a proxy covariate enables us to recover the probability of the outcome variable and identify not only the probabilities of causation but also the joint probabilities of potential outcome variables.

Theorem 1 identifies the conditional probabilities of the potential outcomes. To identify the joint probabilities of the potential outcome types, it remains to recover the marginal probability of the outcome variable. This can be achieved under the following additional condition:

Condition 5

$$\frac{R_{w_i} R_{z_j} R_{w_k z_l \mid y_0} - R_{w_k} R_{z_l} R_{w_i z_j \mid y_0}}{R_{w_k} R_{z_l} \Delta_{w_i z_j} - R_{w_i} R_{z_j} \Delta_{w_k z_l}} \tag{25}$$

is a constant value within the range $(0, 1)$ *for any combination of* (w_i, z_j, w_k, z_l) *satisfying* $(w_i, z_j) \neq (w_k, z_l)$ $(i, k = 1, 2, 3; j, l = 1, 2)$*, where*

$$R_w = p(w \mid y_1) - p(w \mid y_0), \tag{26}$$

$$R_z = p(z \mid y_1) - p(z \mid y_0), \tag{27}$$

$$R_{wz|y} = p(w, z \mid y) - p(w \mid y) p(z \mid y), \tag{28}$$

$$\Delta_{wz} = R_{wz|y_1} - R_{wz|y_0} \tag{29}$$

for $y \in \{y_0, y_1\}$*,* $z \in \{z_1, z_2\}$ *and* $w \in \{w_1, w_2, w_3\}$*.*

Under Conditions 1–5, the joint probabilities of the potential outcome types are identified:

Theorem 2 *Under Conditions 1–5, the joint probabilities of potential outcome variables* $p(u_1)$*,* $p(u_2)$*,* $p(u_3)$*, and* $p(u_4)$ *are identifiable.*

The proof is given in the Supplementary Material A.

Condition 5 is not an additional restriction on proxies but an overidentifying restriction arising from the proxy covariate structrure together with an instrumental variable. The constancy corresponds to agreement across multiple moment equations, and the (0,1) constraint ensures a valid probability. We note that Condition 5 is sensitive to misspecification of the structural assumptions.

Unlike Theorem 1, Theorem 2 also provides situations where not only PN, PS, and PNS, but also the joint probabilities of potential outcome variables $p(u_1)$, $p(u_2)$, $p(u_3)$, and $p(u_4)$ are identifiable since $p(y_1)$ can be recovered under additional testable conditions.

4 ESTIMATION

When $p(Y_x = y \mid x', y')$ is identifiable under the proposed conditions (for $x, x' \in \{x_0, x_1\}$ and $y \in \{y_0, y_1\}$), the resulting estimation problem reduces to that of a singular model. Consequently, these probabilities cannot be evaluated by standard statistical estimation methods such as maximum likelihood estimation. To address this issue, we propose new plug-in estimators of $p(Y_x = y \mid x', y')$ (for $x, x' \in \{x_0, x_1\}$ and $y \in \{y_0, y_1\}$). The proposed plug-in estimators are derived in two steps: (i) estimating $p(w \mid u)$, and (ii) estimating $p(u \mid x, y)$ using the results from the first step.

4.1 ESTIMATING $p(w \mid u)$

Let $\widehat{a}_{xy}$ and $\widehat{b}_{xy}$ denote consistent estimators of a_{xy} and b_{xy}, respectively, obtained by replacing $p(z \mid x, y)$, $p(w \mid x, y)$, and $p(z, w \mid x, y)$ of a_{xy} and b_{xy} with their sample counterparts $\widehat{p}(z \mid x, y)$, $\widehat{p}(w \mid x, y)$, and $\widehat{p}(z, w \mid x, y)$, respectively (for $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$, $w \in \{w_1, w_2, w_3\}$). Then, the consistent estimator of

$p(w \mid u)$ are given by

$$\begin{pmatrix} \widehat{p}(w_1 \mid u_1) \\ \widehat{p}(w_2 \mid u_1) \end{pmatrix} = \begin{pmatrix} 1 & -\widehat{a}_{x_1 y_1} \\ 1 & -\widehat{a}_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{b}_{x_1 y_1} \\ \widehat{b}_{x_0 y_1} \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(w_1 \mid u_2) \\ \widehat{p}(w_2 \mid u_2) \end{pmatrix} = \begin{pmatrix} 1 & -\widehat{a}_{x_1 y_1} \\ 1 & -\widehat{a}_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{b}_{x_1 y_1} \\ \widehat{b}_{x_0 y_0} \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(w_1 \mid u_3) \\ \widehat{p}(w_2 \mid u_3) \end{pmatrix} = \begin{pmatrix} 1 & -\widehat{a}_{x_1 y_0} \\ 1 & -\widehat{a}_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{b}_{x_1 y_0} \\ \widehat{b}_{x_0 y_1} \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(w_1 \mid u_4) \\ \widehat{p}(w_2 \mid u_4) \end{pmatrix} = \begin{pmatrix} 1 & -\widehat{a}_{x_1 y_0} \\ 1 & -\widehat{a}_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{b}_{x_1 y_0} \\ \widehat{b}_{x_0 y_0} \end{pmatrix}. \quad (30)$$

4.2 ESTIMATING $p(u \mid x, y)$

By replacing $p(w_1 \mid u)$ and $p(w_1 \mid x, y)$ in equation (16) with $\widehat{p}(w_1 \mid u)$ and $\widehat{p}(w_1 \mid x, y)$, respectively ($x \in \{x_0, x_1\}, y \in \{y_0, y_1\}, u \in \{u_1, u_2, u_3, u_4\}$), the consistent estimators $\widehat{p}(u \mid x, y)$ of $p(u \mid x, y)$ are given by

$$\begin{pmatrix} \widehat{p}(u_1 \mid x_1, y_1) \\ \widehat{p}(u_2 \mid x_1, y_1) \end{pmatrix} = \begin{pmatrix} 1 & \widehat{p}(w_1 \mid u_1) \\ 1 & \widehat{p}(w_1 \mid u_2) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ \widehat{p}(w_1 \mid x_1, y_1) \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(u_1 \mid x_0, y_1) \\ \widehat{p}(u_3 \mid x_0, y_1) \end{pmatrix} = \begin{pmatrix} 1 & \widehat{p}(w_1 \mid u_1) \\ 1 & \widehat{p}(w_1 \mid u_3) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ \widehat{p}(w_1 \mid x_0, y_1) \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(u_3 \mid x_1, y_0) \\ \widehat{p}(u_4 \mid x_1, y_0) \end{pmatrix} = \begin{pmatrix} 1 & \widehat{p}(w_1 \mid u_3) \\ 1 & \widehat{p}(w_1 \mid u_4) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ \widehat{p}(w_1 \mid x_1, y_0) \end{pmatrix},$$
$$\begin{pmatrix} \widehat{p}(u_2 \mid x_0, y_0) \\ \widehat{p}(u_4 \mid x_0, y_0) \end{pmatrix} = \begin{pmatrix} 1 & \widehat{p}(w_1 \mid u_2) \\ 1 & \widehat{p}(w_1 \mid u_4) \end{pmatrix}^{-\top} \begin{pmatrix} 1 \\ \widehat{p}(w_1 \mid x_0, y_0) \end{pmatrix}. \quad (31)$$

Therefore, from equations (17)-(20), we can estimate $p(Y_x = y \mid x', y')$ as follows:

$$\widehat{p}(Y_{x_0} = y_0 \mid x_1, y_1) = \widehat{p}(u_2 \mid x_1, y_1),$$
$$\widehat{p}(Y_{x_1} = y_0 \mid x_0, y_1) = \widehat{p}(u_3 \mid x_0, y_1),$$
$$\widehat{p}(Y_{x_0} = y_1 \mid x_1, y_0) = \widehat{p}(u_3 \mid x_1, y_0),$$
$$\widehat{p}(Y_{x_1} = y_1 \mid x_0, y_0) = \widehat{p}(u_2 \mid x_0, y_0).$$

4.3 ESTIMATING $p(u)$

Here, in the context of Theorem 2, by replacing $p(w \mid y)$, $p(z \mid y)$ and $p(w, z \mid y)$ in equation (25) of Theorem 2 with the sample probabilities $\widehat{p}(w \mid y)$, $\widehat{p}(z \mid y)$ and $\widehat{p}(w, z \mid y)$, respectively ($y \in \{y_0, y_1\}, z \in \{z_1, z_2\}, w \in \{w_1, w_2, w_3\}$), the consistent estimator $\widehat{p}(y_1)$ of $p(y_1)$ is given by

$$\frac{\widehat{R}_{w_i} \widehat{R}_{z_j} \widehat{R}_{w_k z_l \mid y_0} - \widehat{R}_{w_k} \widehat{R}_{z_l} \widehat{R}_{w_i z_j \mid y_0}}{\widehat{R}_{w_k} \widehat{R}_{z_l} \widehat{\Delta}_{w_i z_j} - \widehat{R}_{w_i} \widehat{R}_{z_j} \widehat{\Delta}_{w_k z_l}} \quad (32)$$

for $(w_k, z_l) \neq (w_i, z_j)$, where

$$\widehat{R}_w = \widehat{p}(w \mid y_1) - \widehat{p}(w \mid y_0), \quad (33)$$
$$\widehat{R}_z = \widehat{p}(z \mid y_1) - \widehat{p}(z \mid y_0), \quad (34)$$
$$\widehat{R}_{wz \mid y} = \widehat{p}(w, z \mid y) - \widehat{p}(w \mid y)\widehat{p}(z \mid y), \quad (35)$$
$$\widehat{\Delta}_{wz} = \widehat{R}_{wz \mid y_1} - \widehat{R}_{wz \mid y_0} \quad (36)$$

for $y \in \{y_0, y_1\}$, $z \in \{z_1, z_2\}$ and $w \in \{w_1, w_2, w_3\}$. Thus, noting that $p(x \mid y)$ is estimable by the sample probabilities $\widehat{p}(x \mid y)$ for $x \in \{x_0, x_1\}, y \in \{y_0, y_1\}$, the consistent estimator $\widehat{p}(u)$ of the joint probabilities of potential outcome variables $p(u)$ is given by

$$\widehat{p}(u) = \widehat{p}(Y_x = y \mid x', y')\widehat{p}(x' \mid y')\widehat{p}(y') + \widehat{p}(Y_{x'} = y' \mid x, y)\widehat{p}(x \mid y)\widehat{p}(y) \quad (37)$$

for $u \in \{u_1, u_2, u_3, u_4\}$ and the corresponding pair (x, y) ($x \in \{x_0, x_1\}, y \in \{y_0, y_1\}$).

5 NUMERICAL EXPERIMENT

In this section, we present a numerical experiment to examine the properties of the proposed estimation method for PN, PS, and PNS. We consider the causal diagram shown in Figure 2a, where the joint distribution of (X, Y, Z, W, U) is specified in Table 1. To simulate a case–control study based on Figure 2b and Table 1, under the scenario in which (X, Y, Z, W) are observed but U is unobserved, we proceed as follows. In Step 1, we generate a sufficiently large dataset of (X, Y, Z, W) according to Table 1. From this dataset, we extract two samples of size N: one consisting of individuals with $Y = y_1$ (the case group) and the other consisting of individuals with $Y = y_0$ (the control group). Note that under Table 1, $p(y_0) = 0.6875$ and $p(y_1) = 0.3125$. In Step 2, based on the case and control groups obtained in Step 1, we compute the proposed estimators for $N = 1000$, 5000, 10000, and 50000. Since the true values of PN and PS under Table 1 are 0.875 and 0.500, respectively, the sample means of the corresponding estimators are expected to be close to these values. Likewise, because the true value of PNS is 0.4375, its estimator is expected to concentrate around 0.4375.

Table 2 and Figure 3 show the basic statistics and boxplots of the estimated PN, PS, PNS based on $5,000$ replications with the given sample size N. Here, the probabilities of causation are estimated using bootstrap replication so that the estimates remain within the range $[0, 1]$, because if the determinant of equation (13) or (14) is close to zero, Condition 2 or 3 may be violated.

From Table 2 and Figure 3, the sample means of the estimated PN, PS, and PNS are close to their true values, and the sample standard errors decrease as the sample size increases. In addition, the interquartile ranges of the estimated PN, PS, and PNS become narrower as the sample size increases, while still covering the true values. Furthermore, the estimated PN, PS, and PNS appear to be approximately symmetrically distributed. Indeed, their asymptotic normality

Table 1: Parameter setting in the numerical experiment.

	(a) $p(W\|U)$			(b) $p(Y=y_1\|X,U)$		(c) $p(X=x_1\|Z,U)$		(d) $p(U)$
	w_1	w_2	w_3	x_0	x_1	z_1	z_2	
u_1	7/10	1/5	1/10	1	1	7/10	3/10	1/16
u_2	3/20	1/20	4/5	0	1	3/20	17/20	7/16
u_3	5/10	2/5	1/10	1	0	5/10	5/10	1/16
u_4	1/10	1/5	7/10	0	0	1/10	9/10	7/16

	(e) $p(Z)$
z_1	1/2
z_2	1/2

Table 2: Basic statistics in the numerical experiments.

	(a) estimated PN				(b) estimated PS			
	$N=1,000$	$N=5,000$	$N=10,000$	$N=50,000$	$N=1,000$	$N=5,000$	$N=10,000$	$N=50,000$
Minimum	0.052	0.286	0.576	0.767	0.000	0.017	0.099	0.355
1st Quantile	0.831	0.844	0.847	0.852	0.327	0.421	0.454	0.482
Median	0.882	0.886	0.884	0.877	0.464	0.485	0.496	0.500
Mean	0.877	0.885	0.886	0.879	0.454	0.473	0.489	0.499
3rd Quantile	0.933	0.930	0.924	0.902	0.588	0.542	0.533	0.517
Maximum	1.000	1.000	1.000	0.998	0.993	0.763	0.731	0.614
s.e.	0.074	0.060	0.053	0.038	0.190	0.105	0.067	0.027

	(c) estimated PNS			
	$N=1,000$	$N=5,000$	$N=10,000$	$N=50,000$
Minimum	0.006	0.132	0.157	0.375
1st Quantile	0.326	0.389	0.409	0.428
Median	0.422	0.430	0.435	0.438
Mean	0.407	0.423	0.433	0.438
3rd Quantile	0.497	0.465	0.461	0.448
Maximum	0.745	0.601	0.562	0.507
s.e.	0.122	0.065	0.042	0.015

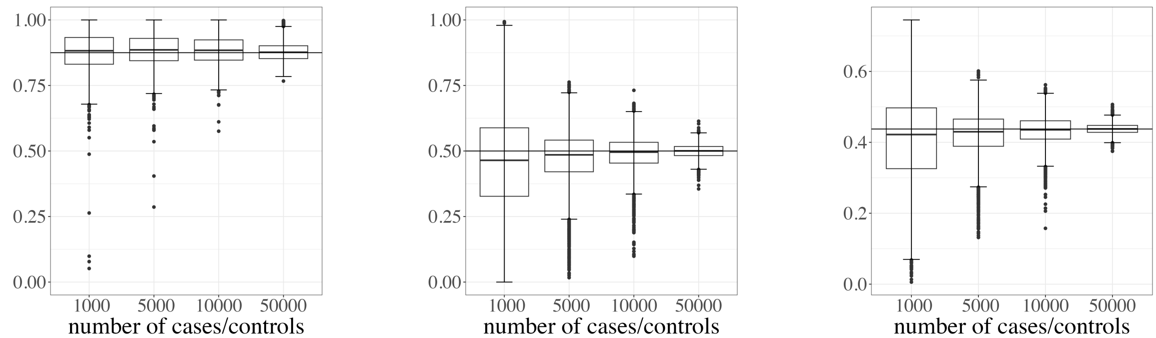
can be established by arguments similar to those in Shingaki and Kuroki [2023].

6 DISCUSSION

Probabilities of causation arise in various areas of practical science, such as the identification of prevented and preventable proportions [Yamada and Kuroki, 2017] and proportions of benefit and harm [Mueller et al., 2022], the explainable AI (XAI) problems [Galhotra et al., 2021, Mothilal et al., 2021, Watson et al., 2021], the evaluation problems of traffic conflicts [Yamada and Kuroki, 2019], personalized decision-making problems [Mueller and Pearl, 2023], and unit selection problems [Li and Pearl, 2019]. These applications indicate that the identification and estimation of such probabilities consititute an important topic in statistical causal inference. To address these problems in the context of case–control studies, this paper provides novel identification conditions for the probabilities of causation using (i) two proxy covariates, and (ii) a proxy covariate together with an instrumental variable (IV). In particular, even without the knowledge of the probability of the outcome variable, the independence between the instrumental variable and a proxy covariate enables the identification of not only the probabilities of causation, but also the joint distribution of the potential outcome variables.

To the best of our knowledge, although there is a substantial need to address these problems without relying on the monotonicity assumption in case–control studies, there has been little work on the identification of probabilities of causation in this setting. The bias-breaker approach of Geneletti et al. [2009] and the related developments by Stanghellini et al. [2025] aim to render selection ignorable in case–control studies in order to recover causal risks. In contrast, our proxy-based approach exploits invertibility conditions on outcome-conditional distributions to identify probabilities of causation and, in the instrumental-variable setting, the joint probabilities of potential outcomes. The two approaches are therefore complementary, addressing different causal quantities within case–control studies.

In addition, estimation problems involving singular models often arise in statistical causal inference, yet they remain largely unresolved in the existing literature. To address this issue, we proposed new plug-in estimators for the probabili-



(a) PN (b) PS (c) PNS

Figure 3: Boxplots of estimated "probabilities of causation".

ties of causation based on the proposed identification conditions. Because the proposed estimators involve ratios and matrix inversions, finite-sample variability may be amplified when denominators are close to zero, leading to infeasible estimates. In practice, bootstrap resampling can be used to obtain feasible estimates. Alternatively, regularized or constrained estimators [e.g., Shingaki and Kuroki, 2024] may provide a useful remedy, although at the cost of introducing some bias.

Acknowledgements

This research was partially supported by a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science, Grant Numbers 24K14851. We thank the anonymous reviewers for their insightful comments.

References

Ørnulf Borgan, Norman E. Breslow, Nilanjan Chatterjee, Mitchell H. Gail, Alastair Scott, and Christopher J Wild. *Handbook of statistical methods for case-control studies*. CRC Press, 2018.

Norman E. Breslow and Nicholas E. Day. *Statistical methods in cancer research. Volume I — The analysis of case-control studies*. IARC Scientific Publication, 1980.

Ruth M. J. Byrne. Counterfactuals in explainable artificial intelligence (xai): Evidence from human reasoning. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 6276–6282. International Joint Conferences on Artificial Intelligence Organization, 2019.

Zhihong Cai and Manabu Kuroki. Variance estimators for three "probabilities of causation". *Risk Analysis*, 25(6): 1611–1620, 2005.

David Card. Using geographic variations in college proximity to estimate the return to schooling. In L.N. Christofides, L.N. Christofides, J. Vanderkamp, E.K. Grant, and R. Swidinsky, editors, *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*. University of Toronto Press, 1995.

Louis Anthony Cox. Probability of causation and the attributable proportion risk. *Risk Analysis*, 4(3):221–230, 1984.

A. Philip Dawid, Monica Musio, and Rossella Murtas. The probability of causation. *Law, Probability and Risk*, 16 (4):163–179, 2017.

Sainyam Galhotra, Romila Pradhan, and Babak Salimi. Explaining black-box algorithms using probabilistic contrastive counterfactuals. In *Proceedings of the 2021 International Conference on Management of Data*, SIGMOD '21, page 577–590, New York, NY, USA, 2021. Association for Computing Machinery.

Sara Geneletti, Sylvia Richardson, and Nicky Best. Adjusting for selection bias in retrospective, case–control studies. *Biostatistics*, 10(1):17–31, 01 2009. ISSN 1465-4644.

Sander Greenland. Variance estimators for attributable fraction estimates consistent in both large strata and sparse data. *Statist. Med.*, 6(6):701–708, 1987.

Sander Greenland. An introduction to instrumental variables for epidemiologists. *International Journal of Epidemiology*, 29(4):722–729, 08 2000.

Guido W. Imbens and Donald B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

Yuta Kawakami, Ryusei Shingaki, and Manabu Kuroki. Identification and estimation of the probabilities of potential outcome types using covariate information in studies with non-compliance. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):12234–12242, Jun. 2023.

Ruth H. Keogh and David R. Cox. *Case-Control Studies*. Cambridge University Press, 2014.

Muin J Khoury, W Dana Flanders, Sander Greenland, and Myron J Adams. On the measurement of susceptibility in epidemiologic studies. *American Journal of Epidemiology*, 129(1):183–190, 1989.

Manabu Kuroki and Zhihong Cai. Statistical analysis of 'probabilities of causation' using co-variate information. *Scand. J. Statist.*, 38(3):564–577, 2011.

Manabu Kuroki and Judea Pearl. Measurement bias and effect restoration in causal inference. *Biometrika*, 101(2): 423–437, 2014.

Manabu Kuroki, Zhihong Cai, and Zhi Geng. Sharp bounds on causal effects in case-control and cohort studies. *Biometrika*, 97(1):123–132, 2010.

Timothy L. Lash, Tyler J. VanderWeele, Sebastien Haneuse, and Kenneth J. Rothman. *Modern Epidemiology*. Lippincott Williams & Wilkins, 4th edition, 2021.

Ang Li and Judea Pearl. Unit selection based on counterfactual logic. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 1793–1799. International Joint Conferences on Artificial Intelligence Organization, 2019.

Wang Miao, Zhi Geng, and Eric J Tchetgen Tchetgen. Identifying causal effects with proxy variables of an unmeasured confounder. *Biometrika*, 105(4):987–993, 12 2018. ISSN 0006-3444.

Ramaravind Kommiya Mothilal, Divyat Mahajan, Chenhao Tan, and Amit Sharma. Towards unifying feature attribution and counterfactual explanations: Different means to the same end. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, pages 652–663, New York, NY, USA, 2021. Association for Computing Machinery.

Scott Mueller and Judea Pearl. Personalized decision making – a conceptual introduction. *Journal of Causal Inference*, 11(1):20220050, 2023.

Scott Mueller, Ang Li, and Judea Pearl. Causes of effects: Learning individual responses from population data. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 2712–2718. International Joint Conferences on Artificial Intelligence Organization, 2022.

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

Judea Pearl. On measurement bias in causal inference. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pages 425–432, Arlington, Virginia, USA, 2010. AUAI Press.

James Robins and Sander Greenland. The probability of causation under a stochastic model for individual risk. *Biometrics*, 45(4):1125–1138, 1989.

Ryusei Shingaki and Manabu Kuroki. Identification and estimation of joint probabilities of potential outcomes in observational studies with covariate information. In *Advances in Neural Information Processing Systems*, volume 34, pages 26475–26486. Curran Associates, Inc., 2021.

Ryusei Shingaki and Manabu Kuroki. Probabilities of potential outcome types in experimental studies: Identification and estimation based on proxy covariate information. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10):12287–12294, 2023.

Ryusei Shingaki and Manabu Kuroki. Identification and estimation of "causes of effects" using covariate-mediator information. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pages 3574–3582. PMLR, 2024.

Elena Stanghellini, Marco Doretti, and Taiki Tezuka. A note on "simple graphical rules to assess selection bias in general-population and selected-sample treatment effects" by m. b. mathur and i. shpitser. *American Journal of Epidemiology*, 194(3):562–564, 2025. ISSN 0002-9262.

Jin Tian and Judea Pearl. Probabilities of causation: Bounds and identification. *Annals of Mathematics and Artificial Intelligence*, 28(1):287–313, 2000.

Jin Tian and Judea Pearl. A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, pages 567–573, USA, 2002. American Association for Artificial Intelligence.

David S. Watson, Limor Gultchin, Ankur Taly, and Luciano Floridi. Local explanations via necessity and sufficiency: unifying theory and practice. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 1382–1392. PMLR, 2021.

Kentaro Yamada and Manabu Kuroki. Counterfactual-based prevented and preventable proportions. *Journal of Causal Inference*, 5(2):20160020, 2017.

Kentaro Yamada and Manabu Kuroki. New traffic conflict measure based on a potential outcome model. *Journal of Causal Inference*, 7(1):20180001, 2019.

Evaluation of "Probabilities of Causation" in Case–Control Studies: Identification and Estimation (Supplementary Material)

Ryusei Shingaki[1] **Haruka Yoshida**[1] **Manabu Kuroki**[1]

[1]Department of Mathematics, Physics, Electrical Engineering and Computer Science, Graduate School of Engineering Science, Yokohama National University, Yokohama, Kanagawa, JAPAN

A PROOFS

From Conditions 1 and 1, and by the consistency property, we have

$$p(z, w \mid x_1, y_1) = p(z \mid u_1, x_1)p(w \mid u_1)p(u_1 \mid x_1, y_1) + p(z \mid u_2, x_1)p(w \mid u_2)p(u_2 \mid x_1, y_1), \tag{38}$$
$$p(z, w \mid x_0, y_1) = p(z \mid u_1, x_0)p(w \mid u_1)p(u_1 \mid x_0, y_1) + p(z \mid u_3, x_0)p(w \mid u_3)p(u_3 \mid x_0, y_1), \tag{39}$$
$$p(z, w \mid x_1, y_0) = p(z \mid u_3, x_1)p(w \mid u_3)p(u_3 \mid x_1, y_0) + p(z \mid u_4, x_1)p(w \mid u_4)p(u_4 \mid x_1, y_0), \tag{40}$$
$$p(z, w \mid x_0, y_0) = p(z \mid u_2, x_0)p(w \mid u_2)p(u_2 \mid x_0, y_0) + p(z \mid u_4, x_0)p(w \mid u_4)p(u_4 \mid x_0, y_0) \tag{41}$$

for $z \in \{z_1, z_2\}$ and $w \in \{w_1, w_2, w_3\}$, since

$$p(x_1, z, u_i \mid y_0) = 0 \quad \text{for } i = 1, 2, \tag{42}$$
$$p(x_1, z, u_i \mid y_1) = 0 \quad \text{for } i = 3, 4, \tag{43}$$
$$p(x_0, z, u_i \mid y_0) = 0 \quad \text{for } i = 1, 3, \tag{44}$$
$$p(x_0, z, u_i \mid y_1) = 0 \quad \text{for } i = 2, 4 \tag{45}$$

hold for $z \in \{z_1, z_2\}$, $w \in \{w_1, w_2, w_3\}$, and $u \in \{u_1, u_2, u_3, u_4\}$.

A.1 PROOF OF LEMMA 1: IDENTIFICATION OF $p(w \mid u)$

From Condition 3, note that

$$\begin{pmatrix} 1 & p(w \mid x_1, y_1) \\ p(z \mid x_1, y_1) & p(z, w \mid x_1, y_1) \end{pmatrix}$$

is invertible for $z \in \{z_1, z_2\}$ and $w \in \{w_1, w_2, w_3\}$. Then, from equation (38), regarding $p(z \mid x_1, y_1)$, $p(w \mid x_1, y_1)$, and $p(z, w \mid x_1, y_1)$, we have

$$\begin{aligned} &\begin{pmatrix} 1 & p(w \mid x_1, y_1) \\ p(z \mid x_1, y_1) & p(z, w \mid x_1, y_1) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ p(z \mid u_1, x_1) & p(z \mid u_2, x_1) \end{pmatrix} \begin{pmatrix} p(u_1 \mid x_1, y_1) & 0 \\ 0 & p(u_2 \mid x_1, y_1) \end{pmatrix} \begin{pmatrix} 1 & p(w \mid u_1) \\ 1 & p(w \mid u_2) \end{pmatrix} \end{aligned} \tag{46}$$

for $w \in \{w_1, w_2\}$. Then, we derive

$$\begin{pmatrix} 1 & p(w_2 \mid u_1) \\ 1 & p(w_2 \mid u_2) \end{pmatrix} \begin{pmatrix} 1 & p(w_2 \mid x_1, y_1) \\ p(z \mid x_1, y_1) & p(z, w_2 \mid x_1, y_1) \end{pmatrix}^{-1} \begin{pmatrix} 1 & p(w_1 \mid x_1, y_1) \\ p(z \mid x_1, y_1) & p(z, w_1 \mid x_1, y_1) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & p(w_2 \mid u_1) \\ 1 & p(w_2 \mid u_2) \end{pmatrix} \begin{pmatrix} 1 & \dfrac{p(z, w_2 \mid x_1, y_1)p(w_1 \mid x_1, y_1) - p(z, w_1 \mid x_1, y_1)p(w_2 \mid x_1, y_1)}{p(z, w_2 \mid x_1, y_1) - p(z \mid x_1, y_1)p(w_2 \mid x_1, y_1)} \\ 0 & \dfrac{p(z, w_1 \mid x_1, y_1) - p(z \mid x_1, y_1)p(w_1 \mid x_1, y_1)}{p(z, w_2 \mid x_1, y_1) - p(z \mid x_1, y_1)p(w_2 \mid x_1, y_1)} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & p(w_2 \mid u_1) \\ 1 & p(w_2 \mid u_2) \end{pmatrix} \begin{pmatrix} 1 & b_{x_1 y_1} \\ 0 & a_{x_1 y_1} \end{pmatrix} = \begin{pmatrix} 1 & b_{x_1 y_1} + p(w_2 \mid u_1) a_{x_1 y_1} \\ 1 & b_{x_1 y_1} + p(w_2 \mid u_2) a_{x_1 y_1} \end{pmatrix} = \begin{pmatrix} 1 & p(w_1 \mid u_1) \\ 1 & p(w_1 \mid u_2) \end{pmatrix},$$

i.e.,

$$p(w_1 \mid u_1) = b_{x_1 y_1} + p(w_2 \mid u_1) a_{x_1 y_1}, \quad p(w_1 \mid u_2) = b_{x_1 y_1} + p(w_2 \mid u_2) a_{x_1 y_1}, \tag{47}$$

where

$$a_{xy} = \frac{p(z, w_1 \mid x, y) - p(z \mid x, y)p(w_1 \mid x, y)}{p(z, w_2 \mid x, y) - p(z \mid x, y)p(w_2 \mid x, y)}, \quad b_{xy} = \frac{p(z, w_2 \mid x, y)p(w_1 \mid x, y) - p(z, w_1 \mid x, y)p(w_2 \mid x, y)}{p(z, w_2 \mid x, y) - p(z \mid x, y)p(w_2 \mid x, y)}$$

for $x \in \{x_0, x_1\}$ and $y \in \{y_0, y_1\}$. Similarly, since

$$\begin{pmatrix} 1 & p(w \mid x_0, y_1) \\ p(z|x_0, y_1) & p(z, w \mid x_0, y_1) \end{pmatrix}, \quad \begin{pmatrix} 1 & p(w \mid x_1, y_0) \\ p(z|x_1, y_0) & p(z, w \mid x_1, y_0) \end{pmatrix}, \quad \begin{pmatrix} 1 & p(w \mid x_0, y_0) \\ p(z|x_0, y_0) & p(z, w \mid x_0, y_0) \end{pmatrix}$$

are invertible for $z \in \{z_1, z_2\}$ and $w \in \{w_1, w_2, w_3\}$, from equations (39)-(41), we have

$$p(w_1 \mid u_1) = b_{x_0 y_1} + p(w_2 \mid u_1) a_{x_0 y_1}, \quad p(w_1 \mid u_3) = b_{x_0 y_1} + p(w_2 \mid u_3) a_{x_0 y_1}, \tag{48}$$

$$p(w_1 \mid u_3) = b_{x_1 y_0} + p(w_2 \mid u_3) a_{x_1 y_0}, \quad p(w_1 \mid u_4) = b_{x_1 y_0} + p(w_2 \mid u_4) a_{x_1 y_0}, \tag{49}$$

$$p(w_1 \mid u_2) = b_{x_0 y_0} + p(w_2 \mid u_2) a_{x_0 y_0}, \quad p(w_1 \mid u_4) = b_{x_0 y_0} + p(w_2 \mid u_4) a_{x_0 y_0}. \tag{50}$$

Thus, from Condition 4, $p(w_1 \mid u_1)$ and $p(w_2 \mid u_1)$ are identifiable from equations (47) and (48) as

$$\begin{pmatrix} p(w_1 \mid u_1) \\ p(w_2 \mid u_1) \end{pmatrix} = \begin{pmatrix} 1 & -a_{x_1 y_1} \\ 1 & -a_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_1} \\ b_{x_0 y_1} \end{pmatrix}. \tag{51}$$

Similarly, $p(w_1 \mid u_2)$, $p(w_2 \mid u_2)$, $p(w_1 \mid u_3)$, $p(w_2 \mid u_3)$, $p(w_1 \mid u_4)$, and $p(w_2 \mid u_4)$ are identifiable from equations (47)-(50) as

$$\begin{pmatrix} p(w_1 \mid u_2) \\ p(w_2 \mid u_2) \end{pmatrix} = \begin{pmatrix} 1 & -a_{x_1 y_1} \\ 1 & -a_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_1} \\ b_{x_0 y_0} \end{pmatrix}, \tag{52}$$

$$\begin{pmatrix} p(w_1 \mid u_3) \\ p(w_2 \mid u_3) \end{pmatrix} = \begin{pmatrix} 1 & -a_{x_1 y_0} \\ 1 & -a_{x_0 y_1} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_0} \\ b_{x_0 y_1} \end{pmatrix}, \tag{53}$$

$$\begin{pmatrix} p(w_1 \mid u_4) \\ p(w_2 \mid u_4) \end{pmatrix} = \begin{pmatrix} 1 & -a_{x_1 y_0} \\ 1 & -a_{x_0 y_0} \end{pmatrix}^{-1} \begin{pmatrix} b_{x_1 y_0} \\ b_{x_0 y_0} \end{pmatrix}, \tag{54}$$

respectively.

A.2 PROOF OF LEMMA 2: IDENTIFICATION OF $p(u \mid x, y)$

First, equation (46) provides

$$\begin{pmatrix} 1 & p(w \mid x_1, y_1) \\ p(z \mid x_1, y_1) & p(z, w \mid x_1, y_1) \end{pmatrix} \begin{pmatrix} 1 & p(w \mid u_1) \\ 1 & p(w \mid u_2) \end{pmatrix}^{-1} = \begin{pmatrix} p(u_1 \mid x_1, y_1) & p(u_2 \mid x_1, y_1) \\ p(z \mid u_1, x_1)p(u_1 \mid x_1, y_1) & p(z \mid u_2, x_1)p(u_2 \mid x_1, y_1) \end{pmatrix}. \tag{55}$$

Comparing the first columns of both sides, we see that $p(u_1 \mid x_1, y_1)$ and $p(u_2 \mid x_1, y_1)$ are identifiable. Similarly, from

$$\begin{pmatrix} 1 & p(w \mid x_0, y_1) \\ p(z \mid x_0, y_1) & p(z, w \mid x_0, y_1) \end{pmatrix} \begin{pmatrix} 1 & p(w \mid u_1) \\ 1 & p(w \mid u_3) \end{pmatrix}^{-1} = \begin{pmatrix} p(u_1 \mid x_0, y_1) & p(u_3 \mid x_0, y_1) \\ p(z \mid u_1, x_0)p(u_1 \mid x_0, y_1) & p(z \mid u_3, x_0)p(u_3 \mid x_0, y_1) \end{pmatrix}, \tag{56}$$

$$\begin{pmatrix} 1 & p(w \mid x_1, y_0) \\ p(z \mid x_1, y_0) & p(z, w \mid x_1, y_0) \end{pmatrix} \begin{pmatrix} 1 & p(w \mid u_3) \\ 1 & p(w \mid u_4) \end{pmatrix}^{-1} = \begin{pmatrix} p(u_3 \mid x_1, y_0) & p(u_4 \mid x_1, y_0) \\ p(z \mid u_3, x_1)p(u_3 \mid x_1, y_0) & p(z \mid u_3, x_1)p(u_3 \mid x_1, y_0) \end{pmatrix}, \tag{57}$$

$$\begin{pmatrix} 1 & p(w \mid x_0, y_0) \\ p(z \mid x_0, y_0) & p(z, w \mid x_0, y_0) \end{pmatrix} \begin{pmatrix} 1 & p(w \mid u_2) \\ 1 & p(w \mid u_4) \end{pmatrix}^{-1} = \begin{pmatrix} p(u_2 \mid x_0, y_0) & p(u_4 \mid x_0, y_0) \\ p(z \mid u_2, x_0)p(u_4 \mid x_0, y_0) & p(z \mid u_4, x_0)p(u_4 \mid x_0, y_0) \end{pmatrix}, \tag{58}$$

we see that $p(u_1 \mid x_0, y_1)$, $p(u_3 \mid x_0, y_1)$, $p(u_3 \mid x_1, y_0)$, $p(u_4 \mid x_1, y_0)$, $p(u_2 \mid x_0, y_0)$, and $p(u_4 \mid x_0, y_0)$ are identifiable. Note that $p(u_3 \mid x_1, y_1)$, $p(u_4 \mid x_1, y_1)$, $p(u_2 \mid x_0, y_1)$, $p(u_4 \mid x_0, y_1)$, $p(u_1 \mid x_1, y_0)$, $p(u_2 \mid x_1, y_0)$, $p(u_1 \mid x_0, y_0)$, and $p(u_3 \mid x_0, y_0)$ are evaluated by zeros from equations (42)-(45). From the consideration, we know that $p(u \mid x, y)$ are identifiable for $x \in \{x_0, x_1\}$, $y \in \{y_0, y_1\}$, and $u \in \{u_1, u_2, u_3, u_4\}$.

A.3 PROOF OF THEOREM 1

By Lemma 2, $p(Y_x = y \mid x', y')$ is identifiable for $(x, y, x', y') = (x_1, y_1, x_0, y_1)$, (x_1, y_1, x_0, y_0), (x_1, y_0, x_0, y_1), (x_1, y_0, x_0, y_0) as

$$\begin{aligned} p(Y_{x_0} = y_0 \mid x_1, y_1) &= p(u_2 \mid x_1, y_1) + p(u_4 \mid x_1, y_1) = p(u_2 \mid x_1, y_1), && (59) \\ p(Y_{x_1} = y_0 \mid x_0, y_1) &= p(u_3 \mid x_0, y_1) + p(u_4 \mid x_0, y_1) = p(u_3 \mid x_0, y_1), && (60) \\ p(Y_{x_0} = y_1 \mid x_1, y_0) &= p(u_1 \mid x_1, y_0) + p(u_3 \mid x_1, y_0) = p(u_3 \mid x_1, y_0), && (61) \\ p(Y_{x_1} = y_1 \mid x_0, y_0) &= p(u_1 \mid x_0, y_0) + p(u_2 \mid x_0, y_0) = p(u_2 \mid x_0, y_0). && (62) \end{aligned}$$

A.4 PROOF OF THEOREM 2

Under Conditions 1–4, note that if $p(y_1)$ is given in Theorem 1, then the joint probabilities of potential outcome variables $p(u_1)$, $p(u_2)$, $p(u_3)$, and $p(u_4)$ are identifiable. Taking this into account, in this section, we demonstrate that the probability $p(y)$ of the outcome variable can be recovered through the use of one proxy covariate and one instrumental variable, shown in Figure 2 of the main part.

To see this, first, note that Z is independent of W. Then, we have

$$\begin{aligned} & p(w, z) - p(w)p(z) \\ &= p(w, z \mid y_1)p(y_1) + p(w, z \mid y_0)(1 - p(y_1)) \\ &\quad - \left(p(w \mid y_1)p(y_1) + p(w \mid y_0)(1 - p(y_1))\right)\left(p(z \mid y_1)p(y_1) + p(z \mid y_0)(1 - p(y_1))\right) \\ &= p(w, z \mid y_0) - p(w \mid y_0)p(z \mid y_0) \\ &\quad + \left(p(w, z \mid y_1) - p(w, z \mid y_0) - p(w \mid y_0)p(z \mid y_1) - p(w \mid y_1)p(z \mid y_0) + 2p(w \mid y_0)p(z \mid y_0)\right)p(y_1) \\ &\quad - (p(w \mid y_1) - p(w \mid y_0))(p(z \mid y_1) - p(z \mid y_0))p(y_1)^2 = 0 \end{aligned} \tag{63}$$

for $z \in \{z_1, z_2\}$ and $w \in \{w_1, w_2, w_3\}$. Here, letting

$$R_w = p(w \mid y_1) - p(w \mid y_0), \quad R_z = p(z \mid y_1) - p(z \mid y_0), \quad R_{wz|y} = p(w, z \mid y) - p(w \mid y)p(z \mid y) \tag{64}$$

for $z \in \{z_1, z_2\}$, $w \in \{w_1, w_2, w_3\}$, and $y \in \{y_0, y_1\}$, equation (63) can be rewritten as

$$R_{wz|y_0} + \left(R_{wz|y_1} - R_{wz|y_0} + R_w R_z\right)p(y_1) - R_w R_z p(y_1)^2 = 0. \tag{65}$$

Then, for $(w_i, z_j) \neq (w_k, z_l)$, we have

$$\begin{pmatrix} R_{w_i} R_{z_j} & R_{w_i z_j|y_0} - R_{w_i z_j|y_1} - R_{w_i} R_{z_j} \\ R_{w_k} R_{z_l} & R_{w_k z_l|y_0} - R_{w_k z_l|y_1} - R_{w_k} R_{z_l} \end{pmatrix} \begin{pmatrix} p(y_1)^2 \\ p(y_1) \end{pmatrix} = \begin{pmatrix} R_{w_i z_j|y_0} \\ R_{w_k z_l|y_0} \end{pmatrix}. \tag{66}$$

If

$$\begin{pmatrix} R_{w_i} R_{z_j} & R_{w_i z_j|y_0} - R_{w_i z_j|y_1} - R_{w_i} R_{z_j} \\ R_{w_k} R_{z_l} & R_{w_k z_l|y_0} - R_{w_k z_l|y_1} - R_{w_k} R_{z_l} \end{pmatrix} \tag{67}$$

Table 3: Parameter setting in the additional numerical experiment.

	(a) $p(W\|U)$			(b) $p(Y=y_1\|X,U)$		(c) $p(X=x_1\|Z,U)$	
	w_1	w_2	w_3	x_0	x_1	z_1	z_2
u_1	7/10	1/5	1/10	1	1	7/10	3/10
u_2	3/20	1/20	4/5	0	1	3/20	17/20
u_3	5/10	2/5	1/10	1	0	5/10	5/10
u_4	1/10	1/5	7/10	0	0	1/10	9/10

	(e) $p(Z)$
z_1	1/2
z_2	1/2

is invertible, we have

$$\begin{pmatrix} p(y_1)^2 \\ p(y_1) \end{pmatrix} = \begin{pmatrix} R_{w_i}R_{z_j} & R_{w_iz_j|y_0} - R_{w_iz_j|y_1} - R_{w_i}R_{z_j} \\ R_{w_k}R_{z_l} & R_{w_kz_l|y_0} - R_{w_kz_l|y_1} - R_{w_k}R_{z_l} \end{pmatrix}^{-1} \begin{pmatrix} R_{w_iz_j|y_0} \\ R_{w_kz_l|y_0} \end{pmatrix}. \tag{68}$$

If the solution regarding $(p(y_1), p(y_1)^2)$ of equation (68) lies within the range $(0,1)$ for any combination of $(w_i, z_j) \neq (w_k, z_l)$, and the square of the second row in the left hand side of equation (68) is equal to the first row in the left hand side of equation (68), we know that $p(y_1)$ is identifiable, and is given by

$$p(y_1) = \frac{R_{w_i}R_{z_j}R_{w_kz_l|y_0} - R_{w_k}R_{z_l}R_{w_iz_j|y_0}}{R_{w_k}R_{z_l}\left(R_{w_iz_j|y_1} - R_{w_iz_j|y_0}\right) - R_{w_i}R_{z_j}\left(R_{w_kz_l|y_1} - R_{w_kz_l|y_0}\right)} \tag{69}$$

from equation (68). Thus, the joint probabilities of potential outcome variables $p(u)$ is given by

$$p(u_1) = p(Y_{x_1} = y_1 \mid x_0, y_1)p(x_0 \mid y_1)p(y_1) + p(Y_{x_0} = y_1 \mid x_1, y_1)p(x_1 \mid y_1)p(y_1), \tag{70}$$
$$p(u_2) = p(Y_{x_1} = y_1 \mid x_0, y_0)p(x_0 \mid y_0)p(y_0) + p(Y_{x_0} = y_0 \mid x_1, y_1)p(x_1 \mid y_1)p(y_1), \tag{71}$$
$$p(u_3) = p(Y_{x_1} = y_0 \mid x_0, y_1)p(x_0 \mid y_1)p(y_1) + p(Y_{x_0} = y_1 \mid x_1, y_0)p(x_1 \mid y_0)p(y_0), \tag{72}$$
$$p(u_4) = p(Y_{x_1} = y_0 \mid x_0, y_0)p(x_0 \mid y_0)p(y_0) + p(Y_{x_0} = y_0 \mid x_1, y_0)p(x_1 \mid y_0)p(y_0). \tag{73}$$

B ADDITIONAL NUMERICAL EXPERIMENTS

In this section, we present additional numerical experiments to further examine the properties of our proposed estimation method for the probabilities of the potential outcome types, i.e., $p(u_1)$, $p(u_2)$, $p(u_3)$, and $p(u_4)$.

For simplicity, we assume that X, Y, Z, W, and U are discrete variables and consider the directed acyclic graph shown in Figure 2b, where the conditional probabilities regarding (X, Y, Z, W, U) are specified in Table 3. To simulate a case–control study based on Figure 2b and Table 3, under the scenario in which (X, Y, Z, W) are observed but U is unobserved, we proceed as follows:

Step 1 For sufficiently large N, generate a dataset $\{(X_i, Y_i, Z_i, W_i)\}_{i=1}^{N}$.

Step 2 Based on this dataset, extract two samples of size N: one consisting of individuals with $Y = y_1$ (the case group) and the other consisting of individuals with $Y = y_0$ (the control group).

We examine the properties of the proposed estimators for the following cases:

(a) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 7/16, 1/16, 7/16)$,

(b) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 1/16, 1/16, 13/16)$.

(c) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (5/16, 5/16, 5/16, 1/16)$,

(d) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/4, 1/4, 1/4, 1/4)$,

Tables 5–8 and Figure 4 present the basic statistics and the boxplots of the estimated $p(u_1)$, $p(u_2)$, $p(u_3)$, and $p(u_4)$) for $5,000$ replications with the given sample size N. Here, the probabilities of potential outcome types are estimated using

Table 4: Parameter setting of $p(Z, W|U)$ in the sensitivity analysis.

(a) $\rho = 0$

	u_1			u_2			u_3			u_4		
	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3
z_1	0.350	0.100	0.050	0.075	0.025	0.400	0.250	0.200	0.050	0.050	0.100	0.350
z_2	0.350	0.100	0.050	0.075	0.025	0.400	0.250	0.200	0.050	0.050	0.100	0.350

(b) $\rho = 0.1$

	u_1			u_2			u_3			u_4		
	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3
z_1	0.371	0.091	0.038	0.092	0.027	0.381	0.273	0.188	0.040	0.062	0.109	0.329
z_2	0.329	0.109	0.062	0.058	0.023	0.419	0.227	0.212	0.060	0.038	0.091	0.371

(c) $\rho = 0.2$

	u_1			u_2			u_3			u_4		
	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3
z_1	0.393	0.081	0.026	0.109	0.03	0.361	0.296	0.175	0.029	0.074	0.119	0.307
z_2	0.307	0.119	0.074	0.041	0.02	0.439	0.204	0.225	0.071	0.026	0.081	0.393

(d) $\rho = 0.3$

	u_1			u_2			u_3			u_4		
	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3	w_1	w_2	w_3
z_1	0.415	0.069	0.016	0.125	0.033	0.342	0.320	0.160	0.020	0.084	0.131	0.285
z_2	0.285	0.131	0.084	0.025	0.017	0.458	0.180	0.240	0.080	0.016	0.069	0.415

bootstrap replications to ensure that the estimates remain within the range $[0, 1]$, because if equation (13) or (14) is not invertible, Conditions 3 or 4 may be violated.

From Tables 5–8 and Figure 4, the sample means of the estimated probabilities of potential outcome types are close to the true values, and the corresponding standard errors decrease as the sample size increases. In addition, the interquartile ranges of the estimated probabilities of potential outcome types are narrower while still containing the true values as the sample size increases. In contrast, the estimated probabilities of potential outcome types appear to be asymmetrically distributed when the true values are far from 0.5. In such situations, the asymptotic normality may not hold for finite sample sizes. These results suggest that the proposed estimation method provides consistent estimators of the probabilities of potential outcome types, although larger sample sizes may be required for the asymptotic normality to become apparent in some cases.

C SENSITIVITY ANALYSIS WITH RESPECT TO VIOLATIONS OF THE INDEPENDENCE ASSUMPTION

The identification result for the instrumental-variable setting presented in Section 3.2 is derived under the assumption that the instrumental variable Z and the proxy variable W are independent. The numerical experiments in Section 5 were conducted under this assumption. Since this assumption may be violated in practice due to the presence of unmeasured common causes between Z and W, it is of interest to investigate the sensitivity of the proposed estimators to such violations.

To this end, we conducted an additional simulation study by introducing positive marginal correlations between Z and W. The data-generating process was based on Figure 2. The conditional distributions $p(Y \mid X, U)$ and $p(X \mid Z, U)$ were fixed according to Table 1, and the distribution of the latent variable was fixed as $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/4, 1/4, 1/4, 1/4)$.

For each target value $\rho = \mathrm{Corr}(Z, W) \in \{0, 0.1, 0.2, 0.3\}$, the conditional distributions $p(Z, W \mid U = u)$ were constructed using iterative proportional fitting (IPF). Specifically, IPF was employed to preserve the specified marginal distributions of Z and W, while a one-dimensional search over the remaining free parameter was performed so that the resulting marginal correlation between Z and W attained the target value ρ. All conditional probabilities were constrained to remain strictly positive. The resulting parameter settings of $p(Z, W \mid U)$ are summarized in Table 4.

For each value of ρ, we generated 5,000 Monte Carlo replications with sample sizes $N = 1000$ and 5000, respectively, and evaluated the proposed estimators of the probabilities of potential outcome types.

Tables 13 and 14, together with Figure 5, summarize the simulation results. As the marginal correlation coefficient ρ increases from 0 to 0.3, both the bias and the sampling variability of the estimated probabilities of potential outcome types increase. The boxplots likewise show that the dispersion of the estimators becomes larger as ρ increases. The same qualitative trend is observed for both sample sizes, $N = 1000$ and 5000. These results indicate that violations of the independence assumption between the instrumental variable and the proxy variable lead to increased bias and variability of the proposed estimators.

D APPLICATION

In this section, we illustrate our results using the education-wage dataset analyzed in Card [1995]. The aim of this study is to estimate the causal effect of education on log earnings by using a binary instrumental variable indicating whether the individual grew up in an area with a nearby four-year college.

The dataset used in this section is publicly available through the CRAN package `ivmodel` (`https://cran.r-project.org/package=ivmodel`). The variables of interest are defined as follows:

X**:** subject's years of education (x_0: less than 12 years; x_1: otherwise),

Y**:** subject's log wage in 1976 (y_0: less than 6.5; y_1: otherwise),

Z**:** indicator of whether the subject grew up near a four-year college (z_1: grew up near a four-year college; z_2 otherwise),

W**:** subject's IQ-type test score collected from the high school (w_1: score < 100; w_2: $100 \leq$ score < 115; w_3: otherwise).

The sample size is 2,061[1]. We assume that the data-generating process is represented by the causal diagram shown in Figure 2. As noted in Card [1995], individuals with higher test scores (on IQ or achievement tests) tend to obtain more schooling and also tend to have higher earnings. The resulting association between schooling and earnings is often interpreted as evidence of ability bias. Here, IQ is introduced as the proxy variable W for the latent ability represented by U. Under the assumed causal model, IQ affects the outcome only through the latent ability and has no direct effect on the treatment or the outcome. The instrument is assumed to affect the outcome only through the treatment, as argued by Card [1995].

To illustrate our method in a case–control setting, we construct a control group of the same size as the case group ($Y = y_1$). The resulting dataset size is 1,514 (case: 757; control: 757).

Noting that Conditions 1, 3, and 4 of Theorem 1 appear to hold for the available data, and assuming that Condition 2 of Theorem 1 also holds, the proposed bootstrap estimates of the PN, PS, and PNS are given by 0.413, 0.433, and 0.132, respectively. The estimated PN implies that 41.3% of subjects with more than 12 years of education and a logarithmic earnings of at least 6.5 would not have attained such earnings had they not received the additional education. Similarly, the estimated PS implies that 43.3% of subjects with less than 12 years of education and a logarithmic earnings below 6.5 would have attained such earnings had they received the additional education. In contrast, the bootstrap estimates of PN, PS, and PNS based on the data without selection bias are 0.398, 0.252, and 0.088, respectively. There appears to be no substantial difference between the estimates obtained with and without a selection bias.

Table 15 shows the basic statistics of 10,000 bootstrap replications of the estimates of PN, PS, and PNS. From Table 15, the estimated PN and PS appear to be approximately symmetrically distributed, whereas the estimated PNS exibits noticeable skewness. In addition, the presence of outliers leads to a wide sample range for PN and PS, and hence the minimum and maximum values provide limited information about these quantities.

[1]The original dataset had a sample size of 3,010, but we restrict attention to the 2,061 subjects for whom IQ data are available.

Table 5: Basic statistics of estimates based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 7/16, 1/16, 7/16)$.

	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.000	0.001	0.007	0.024	0.006	0.143	0.157	0.375
1st Quantile	0.037	0.042	0.046	0.055	0.322	0.390	0.411	0.428
Median	0.059	0.056	0.057	0.061	0.416	0.430	0.437	0.438
Mean	0.064	0.057	0.057	0.061	0.405	0.424	0.435	0.438
3rd Quantile	0.085	0.070	0.069	0.067	0.495	0.466	0.461	0.448
Maximum	0.222	0.249	0.151	0.091	0.745	0.596	0.553	0.507
s.e.	0.038	0.022	0.017	0.010	0.122	0.063	0.041	0.015
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.002	0.005	0.006	0.018	0.008	0.244	0.276	0.359
1st Quantile	0.046	0.051	0.052	0.055	0.353	0.402	0.410	0.426
Median	0.065	0.065	0.066	0.064	0.452	0.447	0.440	0.439
Mean	0.067	0.066	0.065	0.062	0.467	0.452	0.442	0.439
3rd Quantile	0.087	0.080	0.078	0.071	0.567	0.496	0.472	0.452
Maximum	0.279	0.216	0.134	0.104	0.935	0.738	0.789	0.517
s.e.	0.030	0.022	0.020	0.013	0.152	0.075	0.050	0.019

Table 6: Basic statistics of estimates based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 1/16, 1/16, 13/16)$.

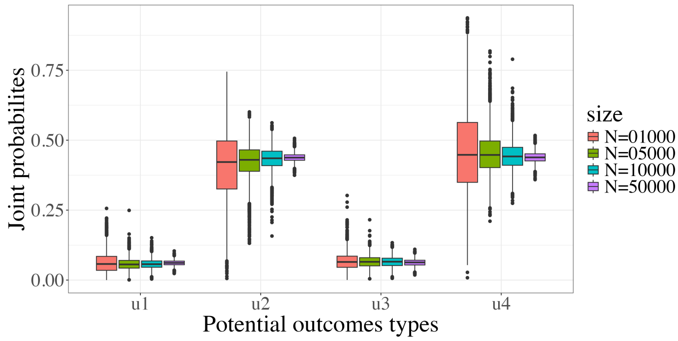
	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.000	0.000	0.000	0.001	0.002	0.002	0.002	0.001
1st Quantile	0.069	0.036	0.034	0.039	0.098	0.052	0.045	0.047
Median	0.136	0.074	0.062	0.057	0.150	0.084	0.071	0.064
Mean	0.162	0.083	0.067	0.057	0.154	0.089	0.076	0.064
3rd Quantile	0.229	0.116	0.095	0.074	0.205	0.120	0.103	0.081
Maximum	0.686	0.417	0.344	0.160	0.454	0.342	0.254	0.190
s.e.	0.124	0.060	0.044	0.026	0.075	0.048	0.040	0.025
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.001	0.004	0.002	0.022	0.003	0.008	0.393	0.614
1st Quantile	0.059	0.051	0.051	0.052	0.419	0.665	0.712	0.773
Median	0.095	0.072	0.068	0.062	0.603	0.760	0.794	0.812
Mean	0.112	0.081	0.074	0.064	0.570	0.745	0.783	0.813
3rd Quantile	0.148	0.101	0.090	0.074	0.748	0.848	0.865	0.856
Maximum	0.446	0.410	0.306	0.158	0.974	0.987	0.981	0.959
s.e.	0.074	0.043	0.033	0.018	0.222	0.134	0.103	0.061

Table 7: Basic statistics of estimates based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (5/16, 5/16, 5/16, 1/16)$.

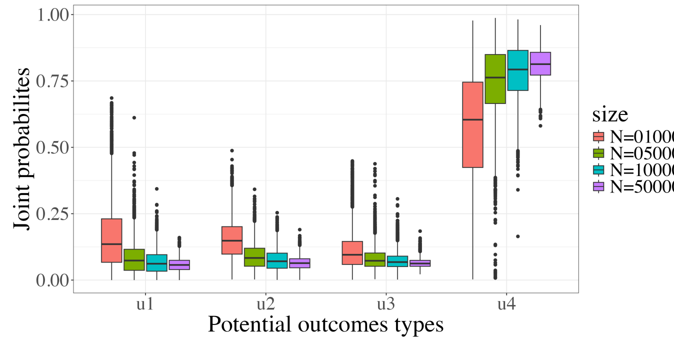
	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.013	0.111	0.187	0.199	0.077	0.123	0.140	0.153
1st Quantile	0.239	0.285	0.294	0.305	0.257	0.279	0.289	0.305
Median	0.296	0.308	0.310	0.312	0.295	0.302	0.306	0.311
Mean	0.300	0.309	0.311	0.312	0.289	0.294	0.299	0.310
3rd Quantile	0.356	0.332	0.327	0.319	0.327	0.319	0.319	0.317
Maximum	0.650	0.532	0.454	0.358	0.475	0.401	0.363	0.350
s.e.	0.093	0.040	0.027	0.012	0.055	0.037	0.031	0.011
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.035	0.096	0.145	0.160	0.000	0.008	0.012	0.034
1st Quantile	0.262	0.287	0.293	0.304	0.054	0.057	0.058	0.059
Median	0.312	0.310	0.308	0.311	0.091	0.081	0.075	0.067
Mean	0.308	0.305	0.305	0.308	0.103	0.091	0.084	0.070
3rd Quantile	0.357	0.329	0.323	0.318	0.137	0.114	0.101	0.076
Maximum	0.504	0.460	0.397	0.356	0.533	0.305	0.259	0.219
s.e.	0.071	0.039	0.029	0.018	0.066	0.046	0.039	0.019

Table 8: Basic statistics of estimates based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/4, 1/4, 1/4, 1/4)$.

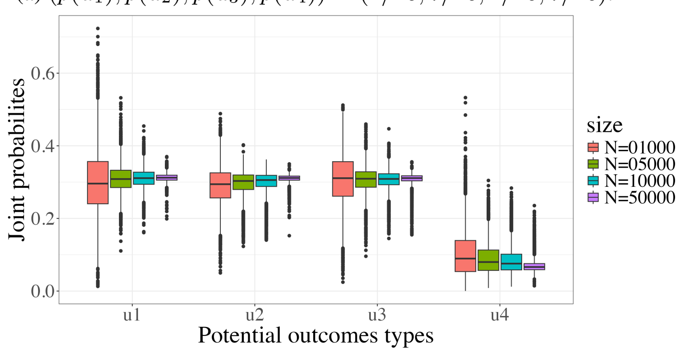
	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.000	0.108	0.124	0.169	0.016	0.060	0.104	0.214
1st Quantile	0.169	0.217	0.229	0.241	0.206	0.226	0.235	0.244
Median	0.233	0.241	0.247	0.249	0.250	0.250	0.251	0.250
Mean	0.239	0.242	0.247	0.250	0.243	0.245	0.248	0.250
3rd Quantile	0.300	0.266	0.266	0.258	0.287	0.268	0.265	0.256
Maximum	0.668	0.445	0.385	0.312	0.469	0.357	0.317	0.303
s.e.	0.103	0.043	0.029	0.013	0.066	0.036	0.025	0.009
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.006	0.071	0.121	0.121	0.000	0.035	0.114	0.176
1st Quantile	0.220	0.232	0.233	0.242	0.171	0.218	0.227	0.240
Median	0.255	0.251	0.249	0.250	0.253	0.259	0.258	0.252
Mean	0.255	0.250	0.247	0.248	0.268	0.263	0.259	0.252
3rd Quantile	0.290	0.270	0.264	0.257	0.348	0.304	0.288	0.264
Maximum	0.504	0.428	0.397	0.287	0.767	0.562	0.516	0.336
s.e.	0.062	0.034	0.027	0.014	0.137	0.068	0.046	0.019



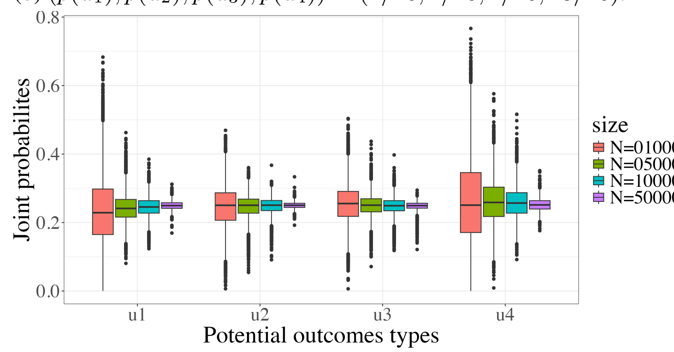
(a) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 7/16, 1/16, 7/16)$.



(b) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 1/16, 1/16, 13/16)$.



(c) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (5/16, 5/16, 5/16, 1/16)$.



(d) $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/4, 1/4, 1/4, 1/4)$.

Figure 4: Boxplots of estimates based on the proposed method.

Table 9: Basic statistics of estimates for "probabilities of causation" based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 7/16, 1/16, 7/16)$.

	PN				PS			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.052	0.286	0.576	0.767	0.000	0.017	0.099	0.355
1st Quantile	0.831	0.844	0.847	0.852	0.327	0.421	0.454	0.482
Median	0.882	0.886	0.884	0.877	0.464	0.485	0.496	0.500
Mean	0.876	0.885	0.886	0.879	0.454	0.473	0.489	0.499
3rd Quantile	0.933	0.930	0.924	0.902	0.588	0.541	0.533	0.517
Maximum	1.000	1.000	1.000	0.998	0.993	0.763	0.731	0.614
s.e.	0.074	0.060	0.053	0.038	0.190	0.105	0.067	0.027

Table 10: Basic statistics of estimates for "probabilities of causation" based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/16, 1/16, 1/16, 13/16)$.

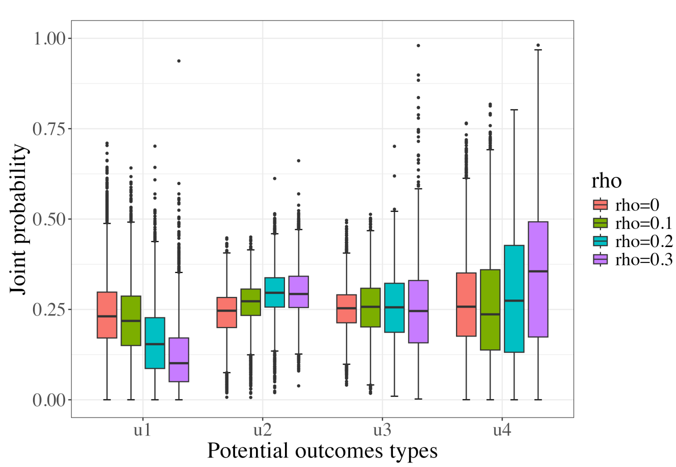
	PN				PS			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.048	0.044	0.005	0.365	0.000	0.000	0.000	0.000
1st Quantile	0.478	0.466	0.463	0.457	0.076	0.044	0.041	0.049
Median	0.547	0.519	0.515	0.506	0.157	0.087	0.076	0.070
Mean	0.567	0.546	0.543	0.536	0.193	0.100	0.084	0.071
3rd Quantile	0.639	0.600	0.587	0.578	0.268	0.144	0.119	0.092
Maximum	0.999	1.000	0.998	1.000	0.991	0.510	0.350	0.214
s.e.	0.139	0.123	0.119	0.114	0.157	0.071	0.056	0.032

Table 11: Basic statistics of estimates for "probabilities of causation" based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (5/16, 5/16, 5/16, 1/16)$.

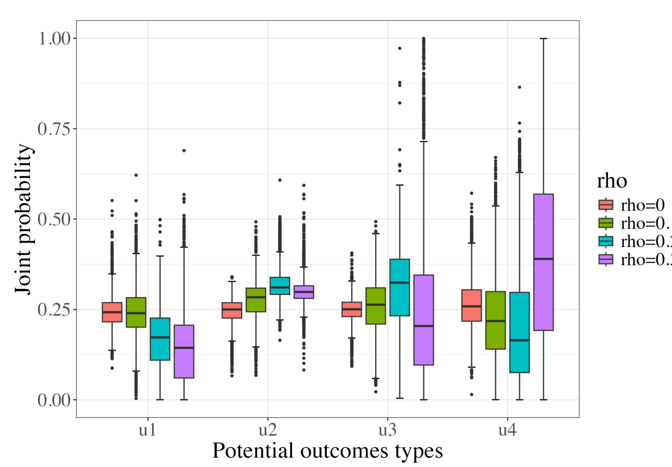
	PN				PS			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.007	0.080	0.265	0.415	0.002	0.000	0.002	0.061
1st Quantile	0.441	0.469	0.477	0.488	0.615	0.676	0.711	0.797
Median	0.494	0.497	0.499	0.500	0.810	0.811	0.819	0.835
Mean	0.494	0.500	0.502	0.500	0.728	0.748	0.764	0.820
3rd Quantile	0.544	0.527	0.521	0.511	0.905	0.891	0.883	0.863
Maximum	1.000	0.983	0.942	0.873	1.000	1.000	1.000	0.995
s.e.	0.095	0.059	0.046	0.024	0.238	0.206	0.182	0.075

Table 12: Basic statistics of estimates for "probabilities of causation" based on the proposed method in the case where $(p(u_1), p(u_2), p(u_3), p(u_4)) = (1/4, 1/4, 1/4, 1/4)$.

	PN				PS			
	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$	$N = 1000$	$N = 5000$	$N = 10000$	$N = 50000$
Minimum	0.002	0.080	0.289	0.399	0.000	0.001	0.001	0.295
1st Quantile	0.457	0.466	0.469	0.482	0.324	0.394	0.432	0.475
Median	0.513	0.506	0.502	0.500	0.510	0.495	0.501	0.502
Mean	0.522	0.514	0.510	0.501	0.495	0.479	0.488	0.500
3rd Quantile	0.575	0.549	0.539	0.518	0.660	0.575	0.557	0.527
Maximum	0.999	0.991	0.999	0.950	1.000	0.982	0.819	0.665
s.e.	0.116	0.077	0.068	0.032	0.229	0.143	0.103	0.040



(a) $N = 1000$.



(b) $N = 5000$.

Figure 5: Boxplots for the sensitivity analysis.

Table 13: Basic statistics for the sensitivity analysis in the case where $N = 1000$.

	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$
Minimum	0.000	0.000	0.000	0.000	0.007	0.007	0.020	0.039
1st Quantile	0.171	0.150	0.087	0.050	0.200	0.233	0.256	0.255
Median	0.231	0.218	0.154	0.101	0.246	0.273	0.296	0.293
Mean	0.240	0.222	0.163	0.118	0.237	0.266	0.296	0.302
3rd Quantile	0.298	0.287	0.227	0.171	0.283	0.306	0.338	0.342
Maximum	0.710	0.641	0.702	0.937	0.448	0.450	0.612	0.661
s.e.	0.102	0.104	0.099	0.088	0.065	0.060	0.066	0.070
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$
Minimum	0.041	0.018	0.010	0.002	0.000	0.000	0.000	0.000
1st Quantile	0.213	0.202	0.187	0.158	0.176	0.138	0.131	0.174
Median	0.253	0.257	0.256	0.246	0.258	0.236	0.274	0.355
Mean	0.252	0.256	0.255	0.247	0.271	0.257	0.286	0.340
3rd Quantile	0.290	0.309	0.322	0.330	0.351	0.360	0.427	0.492
Maximum	0.496	0.513	0.701	0.980	0.765	0.818	0.802	0.981
s.e.	0.062	0.080	0.096	0.119	0.132	0.155	0.179	0.193

Table 14: Basic statistics for the sensitivity analysis in the case where $N = 5000$.

	$\widehat{p}(u_1)$				$\widehat{p}(u_2)$			
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$
Minimum	0.087	0.004	0.000	0.000	0.066	0.068	0.165	0.083
1st Quantile	0.216	0.201	0.110	0.060	0.226	0.243	0.291	0.281
Median	0.242	0.240	0.173	0.144	0.250	0.284	0.311	0.298
Mean	0.244	0.243	0.168	0.142	0.244	0.274	0.325	0.300
3rd Quantile	0.269	0.283	0.226	0.206	0.269	0.309	0.339	0.315
Maximum	0.551	0.621	0.499	0.689	0.341	0.492	0.608	0.593
s.e.	0.044	0.066	0.079	0.093	0.036	0.051	0.054	0.036
	$\widehat{p}(u_3)$				$\widehat{p}(u_4)$			
	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$	$\rho = 0$	$\rho = 0.1$	$\rho = 0.2$	$\rho = 0.3$
Minimum	0.093	0.022	0.004	0.000	0.015	0.000	0.000	0.000
1st Quantile	0.230	0.209	0.232	0.096	0.218	0.140	0.075	0.192
Median	0.250	0.263	0.324	0.204	0.258	0.218	0.165	0.390
Mean	0.249	0.259	0.306	0.236	0.263	0.224	0.203	0.390
3rd Quantile	0.270	0.310	0.389	0.345	0.304	0.299	0.297	0.569
Maximum	0.406	0.493	0.972	1.000	0.572	0.670	0.865	0.999
s.e.	0.035	0.073	0.111	0.180	0.068	0.114	0.160	0.234

Table 15: Basic statistics of bootstrapped estimates for "probabilities of causation" from the Card data.

	without selection bias			with selection bias		
	PN	PS	PNS	PN	PS	PNS
Minimum	0.001	0.000	0.000	0.000	0.003	0.000
1st Quantile	0.159	0.053	0.018	0.177	0.175	0.037
Median	0.335	0.138	0.048	0.358	0.422	0.079
Mean	0.399	0.252	0.088	0.413	0.433	0.132
3rd Quantile	0.631	0.380	0.110	0.646	0.672	0.170
Maximum	0.996	0.999	0.837	0.996	0.996	0.846
s.e.	0.283	0.259	0.113	0.281	0.291	0.145