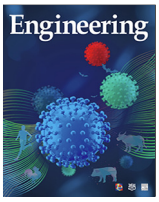
Engineering xxx (xxxx) xxx

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng



Research
Artificial Intelligence—Review

Causal Inference ☆

Kun Kuang [a,*], Lian Li [b], Zhi Geng [c], Lei Xu [d], Kun Zhang [e], Beishui Liao [f], Huaxin Huang [f], Peng Ding [g], Wang Miao [h], Zhichao Jiang [i]

[a] College of Computer Science and Technology, Zhejiang University, Hangzhou 310058, China
[b] Department of Computer Science and Technology, HeFei University of Technology, Hefei 230009, China
[c] School of Mathematical Science, Peking University, Beijing 100871, China
[d] Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[e] Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[f] School of Humanities, Zhejiang University, Hangzhou 310058, China
[g] University of California Berkeley, Berkeley, CA 94720, USA
[h] Guanghua School of Management, Peking University, Beijing 100871, China
[i] Department of Government and Department of Statistics, Harvard University, Cambridge, MA 02138, USA

ARTICLE INFO

Article history:
Received 8 May 2019
Revised 31 July 2019
Accepted 26 August 2019

Keywords:
Causal inference
Instructive variables
Negative control
Causal reasoning and explanation
Causal discovery
Counterfactual inference
Treatment effect estimation

ABSTRACT

Causal inference is a powerful modeling tool for explanatory analysis, which might enable current machine learning to become explainable. How to marry causal inference with machine learning to develop eXplainable Artificial Intelligence (XAI) algorithms is one of key steps towards to the artificial intelligence 2.0. With the aim of bringing knowledge of causal inference to scholars of machine learning and artificial intelligence, we invited researchers working on causal inference to write this survey from different aspects of causal inference. This survey includes the following sections: "Estimating average treatment effect: A brief review and beyond" from Dr. Kun Kuang, "Attribution problems in counterfactual inference" from Prof. Lian Li, "The Yule-Simpson paradox and the surrogate paradox" from Prof. Zhi Geng, "Causal potential theory" from Prof. Lei Xu, "Discovering causal information from observational data" from Prof. Kun Zhang, "Formal argumentation in causal reasoning and explanation" from Profs. Beishui Liao and Huaxin Huang, "Causal inference with complex experiments" from Prof. Peng Ding, "Instrumental variables and negative controls for observational studies" from Prof. Wang Miao, and "Causal inference with interference" from Dr. Zhichao Jiang.

1. Estimating average treatment effect: a brief review and beyond

Machine learning methods have demonstrated great success in many fields, but most lack interpretability. Causal inference is a powerful modeling tool for explanatory analysis, which might enable current machine learning to make explainable prediction. In this article, we review two classical estimators for estimating causal effect, and discuss the remaining challenges in practice. Moreover, we present a possible way to develop eXplainable Artificial Intelligence (XAI) algorithms by marrying causal inference with machine learning.

1.1. The setup

We are interested in estimating the causal effect of a binary variable based on potential outcome framework [1]. For each unit indexed by $i = 1, 2, \ldots, n$ (n denotes the sample size), we observe a treatment T_i, an outcome, and a vector of observed variables $\mathbf{X} \in R^{p\times 1}$, where p refers to the dimension of observed variables. The pair of potential outcomes for each unit i is $\{Y_i(1), Y_i(0)\}$ corresponding to its treatment assignment $T_i = 1$ (treated) or $T_i = 0$ (control). The observed outcome Y_i^{obs} is:

$$Y_i^{\text{obs}} = Y_i(T_i) = T_i \cdot Y_i(1) + (1 - T_i) \cdot Y_i(0) \quad (1)$$

Then, the average treatment effect is defined as follows:

☆ The authors contributed equally to this work. The symbol definitions and notations of each section are relatively independent.

* Corresponding author.
E-mail address: kunkuang@zju.edu.cn (K. Kuang).

https://doi.org/10.1016/j.eng.2019.08.016

$$\tau = \mathbb{E}[Y_i(1) - Y_i(0)] \tag{2}$$

where function $\mathbb{E}(\cdot)$ denotes the expectation function, and the average treatment effect for treated is defined as $\tau_t = \mathbb{E}[Y_i(1) - Y_i(0)|T_i = 1]$.

To identify τ and τ_t, we assume the un-confoundedness—that $T_i \perp ((1),\ Y_i(0))|\boldsymbol{X}_i$—and assume the overlap of the covariate distribution—that $0 < p(T_i = 1|\boldsymbol{X}_i) < 1$.

1.2. Two estimators

Here, we briefly introduce two of the most promising estimators for treatment effect estimation and discuss them for the case with many observed variables.

1.2.1. Inverse propensity weighting

In fully random experiments, the treatment is randomly assigned to units, implying that $T_i \perp \boldsymbol{X}_i$. In observational studies, however, the treatment T_i is assigned based on $\boldsymbol{X}_i$ To remove the confounding effect from $\boldsymbol{X}_i$, the propensity score, denoted as $e(\boldsymbol{X}_i) = (T_i = 1|\boldsymbol{X}_i)$, was proposed to reweight each unit *i*. Then, τ can be estimated by the following:

$$\tau = \mathbb{E}\left[\frac{Y_i^{\text{obs}} T_i}{e(\boldsymbol{X}_i)} - \frac{Y_i^{\text{obs}}(1 - T_i)}{1 - e(\boldsymbol{X}_i)}\right] \tag{3}$$

By combining propensity weighting and regression, it is also possible to estimate the treatment effect with a doubly robust method [2]. In high-dimensional settings, not all observed variables are confounders. To address this issue, Kuang et al. [3] suggest separating all observed variables into two parts: the confounders for propensity score estimation, and the adjustment variables for reducing the variance of the estimated causal effect.

1.2.2. Confounder balancing

The other promising way to remove the confounding effect is to balance the distribution of confounders between treated and control groups by sample reweighting with sample weights *W*, and to estimate τ_t as follows:

$$\tau_t = \mathbb{E}\left[Y_i^{\text{obs}}|T_i = 1\right] - \mathbb{E}\left[W_j Y_i^{\text{obs}}|T_j = 0\right] \tag{4}$$

where the sample weights *W* can be learned by confounder balancing [4] as follows:

$$W = \arg\min_{W} \|\mathbb{E}\left[Y_i^{\text{obs}}|T_i = 1\right] - \mathbb{E}\left[W_j Y_i^{\text{obs}}|T_j = 0\right]\|_{\infty}^2 \tag{5}$$

In high-dimensional settings, different confounders can contribute to different confounding biases. Thus, Kuang et al. [5] suggest jointly learning confounder weights for confounders differentiation, learning sample weights for confounder balancing, and simultaneously estimating the treatment effect with a Differentiated Confounder Balancing (DCB) algorithm.

1.3. Remaining challenges

There are now more promising methods available for estimating treatment effect in observational studies, but many challenges remain in making these methods become useful in practice. Here are some of the remaining challenges:

1.3.1. From binary to continuous

The leading estimators are designed for estimating the treatment effect of a binary variable and achieve good performance in practice. In many real applications, however, we care not only about the cause effect of a treatment, but also about the dose response functions, where the treatment dose may take on a continuum of values.

1.3.2. Interaction of treatments

In practice, the treatment can consist of multiple variables and their interactions. In social marketing, the combined causal effects of different advertising strategies may be of interest. More work is needed on the causal analyses of treatment combination.

1.3.3. Unobserved confounders

The existence of unobserved confounders is equivalent to violation of the unconfoundedness assumption and is not testable. Controlling high-dimensional variables may make unconfoundedness more plausible but poses new challenges to propensity score estimation and confounder balancing.

1.3.4. Limited on overlap

Although the overlap assumption is testable, it raises several issues in practice, including how to detect a lack of overlap in the covariate distributions, and how to deal with such a lack, especially in high-dimensional settings. Moreover, estimating the treatment effect is only possible for the region of overlap.

Recently, related works have been proposed to address the above challenges, including continuous treatment [6], the interaction of treatments [7], unobserved confounders [8], and the limits on overlap [9,10].

1.4. Toward causal and stable prediction

The lack of interpretability of most predictive algorithms makes them less attractive in many real applications, especially those requiring decision-making. Moreover, most current machine learning algorithms are correlation based, leading to instability of their performance across testing data, whose distribution might be different from that of the training data. Therefore, it can be useful to develop predictive algorithms that are interpretable for users and stable to the distribution shift from unknown testing data.

By assuming that the causal knowledge is invariant across datasets, a reasonable way to solve this problem is to explore causal knowledge for causal and stable prediction. Inspired by the confounder-balancing techniques from the literature of causal inference, Kuang et al. [11] propose a possible solution for causal and stable prediction. They propose a global variable balancing regularizer to isolate the effect of each individual variable, thus recovering the causation between each variable and response variable for stable prediction across unknown datasets.

Overall, how to deeply marry causal inference with machine learning to develop XAI algorithms is one of key steps towards to the artificial intelligence (AI) 2.0 [12,13], and remains many special issues, challenges and opportunities.

2. Attribution problems in counterfactual inference

In this section, the input variable *X* and the outcome variable *Y* are both binary.

Counterfactual inference is an important part of causal inference. Briefly speaking, counterfactual inference is to determine the probability that the event *y* would not have occurred ($y = 0$) had the event *x* not occurred ($x = 0$), given the fact that event *x* did occur ($x = 1$) and event *y* did happen ($y = 1$), which can be represented as the following equation:

$$P(y_{x=0} = 0|x = 1,\ y = 1) \tag{6}$$

$y_{x=0}$ is a counterfactual notion, which denotes the value of *y* when the setting is $x = 0$ and the fixing effects of other variables are unchanged,

so it is different from the conditional probability $P(y|x=0)$. This formula reflects the probability that event y will not occur if event x does not occur; that is, it reflects the necessity of the causality of x and y. In social science or logical science, this is called the attribution problem. It is also known as the "but-for" criterion in jurisprudence. The attribution problem has a long history of being studied; however, previous methods used to address this problem have mostly been case studies, statistical analysis, experimental design, and so forth; one example is the influential INUS theory put forward by the Australian philosopher Mackie in the 1960s [14]. These methods are basically qualitative, relying on experience and intuition. With the emergence of big data, however, data-driven quantitative study has been developed for the attribution problem, making the inference process more scientific and reasonable.

Attribution has a twin problem, which is to determine the probability that the event y would have occurred ($y = 1$) had the event x occurred ($x = 1$), given that event x did not occur ($x = 0$) and event y did not happen ($y = 0$). Eq. (7) represents this probability.

$$P(y_{x=1} = 1 | x = 0,\ y = 0) \tag{7}$$

This equation reflects the probability that event x causes event y; that is, it reflects the sufficiency of the causality of x and y.

Counterfactual inference corresponds to human introspection, which is a key feature of human intelligence. Inference allows people to predict the outcome of performing a certain action, while introspection allows people to rethink how they could have improved the outcome, given the known effect of the action. Although introspection cannot change the existing *de facto* situation, it can be used to correct future actions. Introspection is a mathematical model that uses past knowledge to guide future action. Unless it possesses the ability of introspection, intelligence cannot be called true intelligence.

Introspection is also important in daily life. For example, suppose Ms. Jones and Mrs. Smith both had cancer surgery. Ms. Jones also had irradiation. Eventually, both recovered. Then Ms. Jones rethought whether she would have recovered had she not taken the irradiation. Obviously, we cannot infer that Ms. Jones would have recovered had she not take the irradiation, based on the fact that Mrs. Smith recovered without irradiation.

There is an enormous amount of this kind of problem in medical disputes, court trials, and so forth. What we are concerned with is what the real causality is, once a fact has occurred for a specific individual case. In these situations, general statistics data—such as the recovery rate with irradiation—cannot provide the explanation. Calculating the necessity of causality by means of introspection and attribution inference plays a key role in these areas [14].

As yet, no general calculation method exists for Eq. (6). In cases that involve solving a practical problem, researchers introduce a monotonic assumption that can be satisfied in most cases; that is:

$$y_{x=1} \geq y_{x=0}$$

The intuition of monotonicity is that the effect y of taking an action ($x = 1$) will not be worse than that of not taking the action ($x = 0$). For example, in epidemiology, the intuition of monotonicity is not true for people who are contrarily infected ($y = 0$) after being quarantined ($x = 1$), and who were uninfected ($y = 1$) before being quarantined ($x = 0$). Because of the monotonicity, Eq. (6) can be rewritten as follows:

$$\begin{aligned} P(y_{x=0} = 0 | x = 1,\ y = 1) &= \frac{P(y = 1) - P(y_{x=0} = 1)}{P(x = 1,\ y = 1)} \\ &= \frac{P(y = 1|x - 1) - P(y = 1|x = 0)}{P(y = 1|x - 1)} \\ &\quad + \frac{P(y = 1|x = 0) - P(y_{x=0} = 1)}{P(y = 1|x = 1)} \end{aligned} \tag{8}$$

Eq. (8) has two terms. The first term is named the attributable risk fraction, or the excess risk ratio, and is well known in risk statistics. This term reflects the different risk ratio conditioning on $x = 1$ and $x = 0$. The second term is the confounding factor, which should be particularly noticed. This term reflects the effect confounded by other variables. In a natural environment, a change in y could be caused by x in two different ways: First, it could be directly caused by a change in x; or, second, it could be caused by other variables. This phenomenon is called confounding. The difference $P(y = 1|x = 0) - P(y_{x=0} = 1)$ denotes the degree of confounding. In some situations, the change in x did give rise to the change in y, but x may not be the reason for the change in y (e.g., the sun rises after the cock crows). It is possible to exclude confounding by means of scientific experiments to determine the true causality of the change in y. However, scientific experiments can hardly be conducted in many social science problems, or even in some natural science problems. In such cases, only the observational data can be obtained. Thus, the question of how to recognize confounding from observational data in order to determine the true causality is a fundamental problem in artificial intelligence.

In order to explain the relationship between the attributable risk fraction and the confounding factor, and their roles in the attribution problem (i.e., the necessity of causality) more specifically, we applied the example in Ref. [15]. In this example, Mr. A goes to buy a drug to relieve his pain and dies after taking the drug. The plaintiff files a lawsuit to ask the manufacturer to take responsibility. The manufacturer and plaintiff provide the drug test results (i.e., experimental data) and survey results (i.e., nonexperimental data), respectively. The data is illustrated in Table 1, where $x = 1$ denotes taking drugs, while $y = 1$ denotes death.

The manufacturer's data comes from strict drug safety experiments, while the plaintiff's data comes from surveys among patients taking drugs by their own volition. The manufacturer claims that the drug was approved based on the drug distribution regulations. Although it causes a minor increase in death rate (from 0.014 to 0.016), this increase is acceptable compared with the analgesic effect. Based on the traditional calculation of the attributable risk fraction (excess risk ratio), the responsibility taken by the manufacturer is:

$$\frac{P(y = 1|x = 1) - P(y = 1|x = 0)}{P(y = 1|x = 1)} = \frac{0.016 - 0.014}{0.016} = 0.125 \tag{9}$$

The plaintiff argues that the drug test was conducted under experimental protocols, the subjects were chosen randomly, and the subjects did not take the drug of their own volition. Therefore,

Table 1
Experimental and non-experimental data for the example of a drug lawsuit.

Outcomes	Experimental data (number of patients)		Non-experimental data (number of patients)	
	$x = 1$	$x = 0$	$x = 1$	$x = 0$
Deaths ($y = 1$)	16	14	2	28
Survivals ($y = 0$)	984	986	998	972

Please cite this article as: K. Kuang, L. Li, Z. Geng et al., Causal Inference, Engineering, https://doi.org/10.1016/j.eng.2019.08.016

there is bias in the experiment, and the experimental setting differs from the actual situation. There is a huge difference between observational data and experimental data. Given the fact of the death of Mr. A, the calculation of the manufacturer's responsibility should obey the counterfactual equation. The result is:

$$\frac{P(y=1|x-1)-P(y=1|x=0)}{P(y=1|x-1)}+\frac{P(y=1|x=0)-P(y_{x=0}=1)}{P(y=1|x=1)}=\frac{0.002-0.028}{0.002}+\frac{0.028-0.014}{0.001}=1 \quad (10)$$

Therefore, the manufacturer should take full responsibility for the death of Mr. A.

A quick look shows that, based on the survey data, the death rates of taking and not taking the drug are 0.2% and 2.8%, respectively, which is in favor of the manufacturer. However, after careful analysis, the confounding factor is $P(y=1|x=0)-P(y_{x=0}=1)=0.014$; that is, half of the subjects died due to reasons other than not taking the drug. This part should not be attributed to the drug, so the manufacturer's responsibility increases. Of course, there is some doubt regarding whether the manufacturer should take full responsibility, as well as regarding the rationality and scientificity of the calculation [16]. Nevertheless, this example demonstrates that there are confounding factors that will disturb the discovery of true causality. The question of how to determine confounding factors is a practical problem in causal inference, naturally, and is also important in counterfactual inference.

In data science, there are simulated data and objective data, with the latter containing experimental data and observational data. Although observational data are objective, easily available, and low in cost, the confounding problems among them become an obstacle for causal inference [17]. In particular, there may be unknown variables (i.e., hidden variables) in an objective world. These variables are not observed, but may have effects on known variables—that is, the known variables should be sensitive to unmeasured confounding due to unknown variables. In this aspect, current studies on confounding are still in their infancy. Readers can refer to Ref. [18] for more detail.

3. The Yule-Simpson paradox and the surrogate paradox

An association measurement between two variables may be dramatically changed from positive to negative by omitting a third variable, Z; this is called the Yule-Simpson paradox [19,20]. The third variable, Z, is called a confounder. A numerical example is shown in Table 2. The risk difference (RD) is the difference between the proportion of lung cancer in the smoking group and that in the no-smoking group, RD = (80/200) − (100/200) = −0.10, which is negative. If the 400 persons listed in Table 2 are split into males and females, however, a dramatic change can be seen (Table 3). The RDs for both males and females are positive, at 0.10. This means that while smoking is bad for both males and females, separately, smoking is good for all of these persons.

The main difference between causal inference and other forms of statistical inference is whether the confounding bias induced by the confounder is considered. For experimental studies, it is possible to determine which variables affect the treatment or exposure; this is particularly true for a randomized experiment, in which the treatment or exposure is randomly assigned to individuals, as there is no confounder affecting the treatment. Thus, randomized experiments are the gold standard for causal inference. For observational studies, it is key to observe a sufficient set of confounders or an instrumental variable that is independent of all confounders. However, neither a sufficient confounder set nor an instrumental variable can be verified by observational data without manipulations.

Table 2
Smoking and lung cancer.

Condition	Number of persons		
	Cancer	No cancer	Total
Smoking	80	120	200
No smoking	100	100	200

Table 3
Smoking and lung cancer with populations stratified by gender.

Condition	Males		Females	
	Cancer	No cancer	Cancer	No cancer
Smoking	35	15	45	105
No smoking	90	60	10	40

In scientific studies, a surrogate variable (e.g., a biomarker) is often measured instead of an endpoint, due to its infeasible measurement, and; then, the causal effect of a treatment on the unmeasured endpoint is predicted by the effect on the surrogate. The surrogate paradox means that the treatment has a positive effect on the surrogate, and the surrogate has a positive effect on the endpoint, but the treatment may have a negative effect on the endpoint [21]. Numerical examples are given in Refs. [21,22]. This paradox also queries whether scientific knowledge is useful for policy analysis [23]. As a real example, doctors have the knowledge that an irregular heartbeat is a risk factor for sudden death. Several therapies can correct irregular heartbeats, but they increase mortality [24].

Yule-Simpson paradox and the surrogate paradox warn about that a conclusion obtained from data can be inverted due to unobserved confounders and emphasize the importance of using appropriate approaches to obtain data. To avoid the Yule-Simpson paradox, first, randomization is the golden standard approach for causal inference. Second, the use of an experimental approach to obtain data is expected, if randomization is prohibited, as such an approach attempts to balance all possible unobserved confounders between the two groups to be compared. Third, an encouragement-based experimental approach—in which benefits are randomly assigned to a portion of the involved persons, such that the assignment can change the probability of their exposure—can be used to design an instrumental variable. Finally, for a pure observational approach, it is necessary to verify the assumptions required for causal inference using field knowledge, and to further execute a sensitivity analysis for violations of these assumptions. The two paradoxes also point out that a syllogism and transitive reasoning may not be applicable to statistical results. Statistically speaking, smoking is good for both males and females, and the studied population consists of these males and females; however, the statistics indicate that smoking is bad for the population as a whole. Statistics may show that a new drug can correct irregular heartbeats, and it is known that a regular heartbeat can promote survival time, both statistically speaking and for individuals; however, the new drug may still shorten the survival time of these persons in terms of statistics.

4. Causal potential theory

Extensive efforts have been made to detect causal direction, evaluate causal strength, and discover causal structure from observations. Examples include not only the studies based on conditional independence and directed acyclic graphs (DAGs) by Pearl,

Please cite this article as: K. Kuang, L. Li, Z. Geng et al., Causal Inference, Engineering, https://doi.org/10.1016/j.eng.2019.08.016

Table 4
Two roads for analyzing CPT causality.

$\nabla_u E_U$	$y \to x$		$x \to y$		$x \perp\!\!\!\perp y$		$x ? y$	
	Road_A	Road_B	Road_A	Road_B	Road_A	Road_B	Road_A	Road_B
g_x	dependent of y	$\xi(x, y) + \varepsilon$	$\perp\!\!\!\perp y$	$\xi(x) + \varepsilon$	$\perp\!\!\!\perp y$	$\xi(x) + \varepsilon$	dependent of y	$\xi(x, y) + \varepsilon$
g_y	$\perp\!\!\!\perp x$	$\eta(y) + \varepsilon$	dependent of x	$\eta(x, y) + \varepsilon$	$\perp\!\!\!\perp x$	$\eta(y) + \varepsilon$	dependent of x	$\eta(x, y) + \varepsilon$

Spirtes, and many others, but also those on the Rubin causal model (RCM), structural equation model (SEM), functional causal model (FCM), additive noise model (ANM), linear non-Gaussian acyclic model (LiNGAM), post-nonlinear model (PNL), and causal generative neural networks (CGNN), as well as the studies that discovered star structure [25] and identified the so called ρ-diagram [26]. To some extent, these efforts share a similar direction of thinking. First, one presumes a causal structure (e.g., merely one direction in the simplest case, or a DAG in a sophisticated situation) for a multivariate distribution, either modeled in parametric form or partly inspected via statistics, which is subject to certain constraints. Second, one uses observational data to learn the parametric model or estimate the statistics, and then examines whether the model fits the observations and the constraints are satisfied; based on this, one verifies whether the presumed causal structure externally describes observations the well. Typically, a set of causal structures are presumed as candidates, among which the best is selected.

Causal potential theory (CPT) was recently proposed as a very different way of thinking [27]. In analogy to physics, causality is here regarded as an intrinsic kinetic nature caused by a causal potential energy. Without losing generality, this CPT is introduced by starting with the consideration of a cause-effect relation between a pair of variables, x, y,[‡] in an environment, U. Instead of presuming a causal structure (i.e., a specific direction), one estimates a nonparametric distribution $p_U(x,\ y) \triangleq p(x,\ y|U)$ from samples of x, y, and obtains the corresponding causal potential energy $E_U(x,\ y) \propto -ln p_U(x,\ y)$ in an analogy based on the Gibbs distribution. In such a perspective of causal dynamics, an event occurring at x, y is associated with $E_U(x,\ y)$ that yields a force $[g_x,\ g_y]$ to cause subsequent events by the dynamics $[\dot{x}_t,\ \dot{y}_t] \propto -[g_x,\ g_y]$, driving the information flow or causal process toward an area with the lowest energy or, equivalently, toward an area in which events have high chances to occur, using the notations $g_U \triangleq \nabla_U E_U$ and $\dot{u}_t \triangleq \frac{du}{dt}$. That is, CPT regards causality as an intrinsic nature of the dynamics $[\dot{x}_t,\ \dot{y}_t] \propto -[g_x,\ g_y]$ and discovers causality by analyzing $[g_x,\ g_y]$.

Table 4 shows two roads for analyzing CPT causality. Road_A is proceeded by testing a "Yes" or "No" answer on the mutual independence between g_y, y and on that between g_x, x, resulting in four types of Y-N combinations. The first two types indicate two types of causality. The third type, Y-Y, indicates the independence between x, y—that is, indicates that there is no relation between them. The last type, N-N, indicates "unclear ?"—that is, further study is needed to determine whether a causal relation still occurs locally, or even reciprocally, in some regions of x, y, although there is no causal relation detected globally between x, y. Road_A needs an independence test. In contrast, Road_B turns the problem into supervised learning, with x, y as inputs into a neural net to fit two gradient components $[g_x, g_y]$, each of which is fit by a different neural net, with one or both of x, y as inputs, respectively. An appropriate one is chosen according to not only fit, but also simplicity. Table 4 lists four types of outcomes based on this method [27].

It is possible to seek a certain estimator to obtain g_x, g_y directly from samples x_t, y_t, where $t = 1, \ldots, N$ and N refers to the sample size. It is also possible to obtain g_x, g_y indirectly, by estimating $p_U(x,\ y)$ first; that is, by performing a kernel estimate $p_h(x,\ y) = \frac{1}{N}\sum_{t=1}^{N} G\left(x,\ y \middle| x_t,\ y_t,\ h^2\boldsymbol{I}\right)$, where there is a Gaussian of mean m and variance σ^2. Alternatively, it is possible to obtain p_U by one presumed causal structure, and to perform CPT analyses on this p_U.

Experiments on the CauseEffectPairs (CEP) benchmark have demonstrated that a preliminary and simple implementation of CPT has achieved performances that are comparable with ones achieved by state-of-art methods.

Further development is to explore the estimation of causal structure between multiple variable distributions and multiple variables, possibly along two directions. One is simply integrating the methods in Table 4 into the famous PC algorithm [28], especially on edges that are difficultly identified by independent and conditional independent tests. The other is turning the conditions that g_y is uncorrelated (or independent) of x and that g_x is uncorrelated (or independent) of y into multivariate polynomial equations, and adding the equations into the ρ-diagram equations in [26], e.g., Eq. (29) and Eq. (33), to get an augmented group of polynomial equations. Then, the well known Wen-tsun Wu method may be adopted to check whether the equations have unique or a finite number of solutions.

5. Discovering causal information from observational data

Causality is a fundamental notion in science, and plays an important role in explanation, prediction, decision-making, and control [28,29]. There are two essential problems to address in modern causality research. One essential problem is the identification of causal effects, that is, identifying the effects of interventions, given the partially or completely known causal structure and some observed data; this is typically known as "causal inference." For advances in this research direction, readers are referred to Ref. [29] and the references therein. In causal inference, causal structure is assumed to be given in advance—but how can we find causal structure if it is not given? A traditional way to discover causal relations resorts to interventions or randomized experiments, which are too expensive or time-consuming in many cases, or may even be impossible from a practical standpoint. Therefore, the other essential causality problem, which is how to reveal causal information by analyzing purely observational data, has drawn a great deal of attention [28].

In the last three decades, there has been a rapid spread of interest in principled methods causal discovery, which has been driven in part by technological developments. These technological developments include the ability to collect and store big data with huge numbers of variables and sample sizes, and increases in the speed of computers. In domains containing measurements such as satellite images of weather, fMRI brain imaging, gene-expression data, or single-nucleotide polymorphism (SNP) data, the number of variables can range in the millions, and there is often very limited background knowledge to reduce the space of alternative causal hypotheses. Causal discovery techniques without the aid of an automated search then appear to be hopeless. At the same time,

[‡] In this section, we reuse x, y to denote a pair variable, their relationship might be cause and effect.

the availability of faster computers with larger memories and disc space allow for practical implementations of computationally intensive automated algorithms to handle large-scale problems.

It is well known in statistics that "causation implies correlation, but correlation does not imply causation." Perhaps it is fairer to say that correlation does not *directly* imply causation; in fact, it has become clear that under suitable sets of assumptions, the causal structure (often represented by a directed graph) underlying a set of random variables can be recovered from the variables' observed data, at least to some extent. Since the 1990s, conditional independence relationships in the data have been used for the purpose of estimating the underlying causal structure. Typical (conditional independence) constraint-based methods include the Peter–Clark (PC) algorithm and fast causal inference (FCI) [28]. Under the assumption that there is no confounder (i.e., unobserved direct common cause of two measured variables), the result of PC is asymptotically correct. FCI gives asymptotically correct results even when there are confounders. These methods are widely applicable because they can handle various types of causal relations and data distributions, given reliable conditional independence testing methods. However, they may not provide all the desired causal information, because they output (independence) equivalence classes—that is, a set of causal structures with the same conditional independence relations. The PC and FCI algorithms output graphical representations of the equivalence classes. In cases without confounders, there also exist score-based algorithms that estimate causal structure by optimizing some properly defined score function. The greedy equivalence search (GES), among them, is a widely used two-phase procedure that directly searches over the space of equivalence classes.

In the past 13 years, it has been further shown that algorithms based on properly constrained functional causal models (FCMs) are able to distinguish between different causal structures in the same equivalence class, thanks to additional assumptions on the causal mechanism. An FCM represents the outcome or effect variable Y as a function of its direct causes X and some noise term E, that is, $Y = f(X, E)$, where E is independent of X. It has been shown that, without constraints on function f, for any two variables, one of them can always be expressed as a function of the other and independent noise [30]. However, if the functional classes are properly constrained, it is possible to identify the causal direction between X and Y because for wrong directions, the estimated noise and hypothetical cause cannot be independent (although they are independent for the right direction). Such FCMs include the LiNGAM [31], where causal relations are linear and noise terms are assumed to be non-Gaussian; the post-nonlinear (PNL) causal model [32], which considers nonlinear effects of causes and possible nonlinear sensor/measurement distortion in the data; and the nonlinear ANM [33,34], in which causes have nonlinear effects and noise is additive. For a review of these models and corresponding causal discovery methods, readers are referred to Ref. [30].

Causal discovery exploits observational data. The data are produced not only by the underlying causal process, but also by the sampling process. In practice, for reliable causal discovery, it is necessary to consider specific challenges posed in the causal and sampling processes, depending on the application domain. For example, for multivariate time series data such as mRNA expression series in genomics and blood-oxygenation-level-dependent (BOLD) time series in neuropsychology, finding the causal dynamics generating such data is challenging for many reasons, including nonlinear causal interactions, a much lower data-acquisition rate compared with the underlying rates of change, feedback loops in the causal model, the existence of measurement error, nonstationarity of the process, and possible unmeasured confounding causes. In clinical studies, there is often a large amount of missing data. Data collected on the Internet or in hospital often suffer from selection bias. Some datasets involve both mixed categorical and continuous variables, which may pose difficulties in conditional independence tests and in the specification of appropriate forms of the FCM. Many of these issues have recently been considered, and corresponding methods have been proposed to address them.

Causal discovery has benefited a great deal from advances in machine learning, which provide an essential tool to extract information from data. On the other hand, causal information describes properties of the process that render a set of constraints on the data distribution and is able to facilitate understanding and solve a number of learning problems involving distribution shift or concerning the relationship between different factors of the joint distribution. In particular, for learning under data heterogeneity, it is naturally helpful to learn and model the properties of data heterogeneity, which then benefit from causal modeling. Such learning problems include domain adaptation (or transfer learning) [35], semi-supervised learning, and learning with positive and unlabeled examples. Leveraging causal modeling for recommender systems and reinforcement learning is becoming an active research field in recent years.

6. Formal argumentation in causal reasoning and explanation

In this section, we sketch why and how formal argumentation can play an important role in causal reasoning and explanation. Reasoning in argumentation is realized by constructing, comparing, and evaluating arguments [36]. An argument commonly consists of a claim that may be supported by premises, which can be observations, assumptions, or intermediate conclusions of some other arguments. The claim, the premises, and the inference relation between them may be the subject of rebuttals or counter-arguments [37]. An argument can be accepted only when it survives all attacks. In AI, formal argumentation is a general formalism for modeling defeasible reasoning. It provides a natural way for justifying and explaining causation, and is complementary to machine learning approaches, for learning, reasoning, and explaining cause-and-effect relations.

6.1. Nonmonotonicity and defeasibility

Causal reasoning is the process of identifying causality, that is, the relationship between a cause and its effect, which is often defeasible and nonmonotonic. On the one hand, causal rules are typically defeasible. A causal rule may be represented in the form "c causes e" where e is some effect and c is a possible cause. The causal connective is not a material implication, but a defeasible conditional with strength or uncertainty. For example, "turning the ignition key causes the motor to start, but it does not imply it, since there are some other factors such as there being a battery, the battery not being dead, there being gas, and so on" [38]. On the other hand, causal reasoning is nonmonotonic, in the sense that causal connections can be drawn tentatively and retracted in light of further information. It is usually the case that c causes e, but c and d jointly do not cause e. For example, an agent believes that turning the ignition key causes the motor to start, but when it knows that the battery is dead, it does not believe that turning the ignition key will cause the motor to start. In AI, this is the famous qualification problem. Since the potentially relevant factors are typically uncertain, it is not cost effective to reason explicitly. So, when doing causal inference, people usually "jump" to conclusions and retract some conclusions when needed. Similarly, reasoning from evidence to cause is nonmonotonic. If an agent observes some effect e, it is allowed to hypothesize a possible cause c. The reasoning from the evidence to a cause is abductive, since for some evidence, one may accept an abductive explanation if no bet-

Please cite this article as: K. Kuang, L. Li, Z. Geng et al., Causal Inference, Engineering, https://doi.org/10.1016/j.eng.2019.08.016

ter explanation is available. However, when new explanations are generated, the old explanation might be discarded.

6.2. Efficiency and explainability

From a perspective of computation, monotonicity is a crucial property of classical logic, which means that each conclusion obtained by local computation using a subset of knowledge is equal to the one made by global computation using all the knowledge. This property does not hold in nonmonotonic reasoning and, therefore, the computation could be highly inefficient. Due to the nonmonotonicity of causal reasoning, in order to improve efficiency, formal argumentation has been evidenced to be a good candidate, by comparing it with some other nonmonotonic formalisms such as default logic and circumscription. The reason is that in formal argumentation, computational approaches may take advantage of the divide-and-conquer strategy and maximal usage of existing computational results in terms of the reachability between nodes in an argumentation graph [39]. Another important property of causal reasoning in AI is explainability. Traditional nonmonotonic formalisms are not ideal for explanation, since all the proofs are not represented in a human understandable way. Since the purpose of explanation is to let the audience understand, the cognitive process of comparing and contrasting arguments is significant [37]. Argumentation provides such a way by exchanging arguments in terms of justification and argument dialogue [40].

6.3. Connections to machine learning approaches

In explainable AI, there are two components: the explainable model and the explanation interface. The latter includes reflexive explanations that arise directly from the model and rational explanations that come from reasoning about the user's beliefs. To realize this vision, it is natural to combine argumentation and machine learning, in the sense that knowledge is obtained by machine learning approaches, while the reasoning and explanation are realized by argumentation. Since argumentation provides a general approach for various kinds of reasoning in the context of disagreement, and can be combined with some uncertainty measures, such as probability and fuzziness, it is very flexible to model the knowledge learned from data. An example is when a machine learns features and produces an explanation, such as "This face is angry, because it is similar to these examples, and dissimilar from those examples." This is an argument, which might be attacked by other arguments. And, in order to measure the uncertainty described by some words such as "angry," one may choose to use possibilistic or probabilistic argumentation [41]. Different explanations may be in conflict. For instance, there could be some cases invoking specific examples or stories that support a choice, and rejections of an alternative choice that argue against less-preferred answers based on analytics, cases, and data. By using argumentation graphs, these kinds of support-and-attack relations can be conveniently modeled and can be used to compute the status of conflicting arguments for different choices.

7. Causal inference with complex experiments

The potential outcomes framework for causal inference starts with a hypothetical experiment in which the experimenter can assign every unit to several treatment levels. Every unit has potential outcomes corresponding to these treatment levels. Causal effects are comparisons of the potential outcomes among the same set of units. This is sometimes called the experimentalist's approach to causal inference [42]. Readers are referred to Refs. [43–46], for textbook discussions.

7.1. Randomized factorial experiments

Neyman [47] first formally discussed the following randomization model. In an experiment with n units, the experimenter randomly assigns $(n_1, \ldots, n_J)$ units to treatment levels $(1, \ldots, J)$, where $n = \sum_{j=1}^{J} n_j$. Unit i has potential outcomes $\{Y_i(1), \cdots, Y_i(J)\}$, with $Y_i(j)$ being the hypothetical outcome if unit i receives treatment level j. With potential outcomes, we can define causal effects; for example, the comparison between treatment levels j and j' as $\tau\left(j, j'\right) = n^{-1}\sum_{i=1}^{n}\left\{Y_i(j) - Y_i\left(j'\right)\right\}$. Let $T_i(j)$ be the indicator if unit i actually receives treatment level j. Let $Y_i = \sum_{j=1}^{J} T_i(j)Y_i(j)$ be the observed outcome of unit i. With observed data $\{T_i(1), \cdots, T_i(J), Y_i\}_{i=1}^{n}$, Neyman [47] proposed to use $\widehat{\tau}\left(j, j'\right) = n_j^{-1}\sum_{i=1}^{n} T_i(j)Y_i - n_{j'}^{-1}\sum_{i=1}^{n} T_i\left(j'\right)Y_i$ as an estimator for $\tau\left(j, j'\right)$. He showed that $\widehat{\tau}\left(j, j'\right)$ is unbiased with variance $\frac{S^2(j)}{n_j} + \frac{S^2\left(j'\right)}{n_{j'}} - \frac{S^2\left(j-j'\right)}{n}$, where $S^2(j)$, $S^2\left(j'\right)$ and $S^2\left(j - j'\right)$ are the sample variances of $Y_i(j)$, $Y_i\left(j'\right)$ and $Y_i(j) - Y_i\left(j'\right)$. Note that the randomness comes from the treatment indicators with all the potential outcomes fixed. Neyman [47] further discussed variance estimation and the large-sample confidence interval.

We can extend the framework from Ref. [47] to a general causal effect defined as $\tau = n^{-1}\sum_{i=1}^{n}\tau_i$ where $\tau_i = \sum_{j=1}^{J} c_j Y_i(j)$ is the individual effect and the c_j are contrast matrices with $\sum_{j=1}^{J} c_j = 0$. With appropriately chosen contrast matrices, the special cases include analysis of variance [48] and factorial experiments [49,50]. Furthermore, with an appropriately chosen subset of units, the special cases include subgroup analysis, post-stratification [51], and peer effects [52]. Ref. [53] provides the general forms of central limit theorems under this setting for asymptotic inference. Ref. [54] discusses split-plot designs, and Ref. [55] discusses general designs.

7.2. The role of covariates in the analysis of experiments

Neyman's [47] randomization model also allows for the use of covariates to improve efficiency without strong modeling assumptions. In the case with a binary treatment, for unit i, let $\{Y(1), Y(0)\}$ be the potential outcomes, T_i be the binary treatment indicator, and x_i be pretreatment covariates. The average causal effect $\tau = n^{-1}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\}$ has an unbiased estimator $\widehat{\tau} = n_1^{-1}\sum_{i=1}^{n} T_i Y_i - n_0^{-1}\sum_{i=1}^{n}(1 - T_i)Y_i$. Fisher [56] suggested using the analysis of covariance to improve efficiency; that is, running a least squares fit of Y_i on T_i and x_i and using the coefficient of T_i to estimate τ. Ref. [57] uses the model from Ref. [47] to show that Fisher's analysis of the covariance estimator is inferior because it can be even less efficient than $\widehat{\tau}$ and the ordinary least squares can give an inconsistent variance estimate. Ref. [58] proposes a simple correction: First, center covariates to have mean $\bar{x} = 0$; second, run a least squares fit of Y_i on $(T_i, x_i, T_i \times x_i)$ and use the coefficient of T_i to estimate τ, and third, use the Eicker-Huber-White variance estimator [59–61]. With large samples, the estimator from Ref. [58] is at least as efficient as $\widehat{\tau}$, and that researcher's variance estimate is consistent for the true variance of $\widehat{\tau}$.

Ref. [62] extends to the setting with high-dimensional covariates and replaces the least squares fit by the least absolute shrinkage and selection operator (LASSO) [63]. Ref. [64] examines the theoretical boundary of the estimator from Ref. [58], allowing for a diverging number of covariates. Ref. [65] investigates treatment effect heterogeneity using the least squares fit of Y_i on $(T_i, x_i, T_i \times x_i)$. Ref. [66] discusses covariate adjustment in a facto-

rial experiment, and Ref. [67] discusses covariate adjustment in general designs.

7.3. The role of covariates in the design of experiments

An analyzer can use covariates to improve the estimation efficiency. As a dual, a designer can use covariates to improve the covariate balance and consequently improve the estimation efficiency. Ref. [68] hints at the idea of re-randomization—that is, only accepting random allocation that ensures covariate balance. In particular, we accept a random allocation $(T_1, \ldots, T_n)$ if and only if $\hat{\tau}'_x \left\{ \frac{nS_x^2}{(n_1 n_0)} \right\}^{-1} \hat{\tau}_x \leq a$, where $\hat{\tau}_x = n_1^{-1} \sum_{i=1}^{n} T_i x_i - n_0^{-1} \sum_{i=1}^{n} (1 - T_i) x_i$, $S_x^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \hat{x})(x_i - \hat{x})'$ and $a > 0$ is a predetermined constant. Ref. [69] formally discusses its statistical properties under the constant treatment effect model with equal group sizes and Gaussian covariates. Ref. [70] develops its asymptotic theory without these assumptions. In particular, Ref. [70] shows that $\hat{\tau}$ has a non-Gaussian limiting distribution and is more concentrated at τ under re-randomization than under complete randomization. A consequence of the result from Ref. [70] is that when $a \approx 0$, the asymptotic variance of $\hat{\tau}$ under re-randomization is identical to the estimator from Ref. [58] under complete randomization. Therefore, we can view re-randomization as the dual of regression adjustment.

Ref. [71] proposes a re-randomization scheme that allows for tiers of covariates, and Ref. [70] derives its asymptotic properties. Refs. [72,73] extend re-randomization to factorial experiments, and Ref. [74] proposes sequential re-randomization.

7.4. Final remarks

Following Ref. [47], I have focused on the repeated sampling properties of estimators with randomized experiments. Alternatively, Fisher randomization tests are finite-sample exact for any test statistics and for any designs, under the sharp null hypothesis that $Y_i(1) = \cdots = Y_i(J)$ for all units $i = 1, \ldots, J$ [46,75,76]. Refs. [77,78] propose the use of covariate adjustment in randomization tests, and Ref. [69] proposes the use of randomization tests to analyze re-randomization. Refs. [79–81] apply randomization tests to experiments with interference. Refs. [48,50,82] discuss the properties of randomization tests for weak null hypotheses. Refs. [83–85] invert randomization tests to construct exact confidence intervals. Finally, Ref. [86] discusses different inferential frameworks from the missing data perspective.

8. Instrumental variables and negative controls for observational studies

In a great deal of scientific research, the ultimate goal is to evaluate the causal effect of a given treatment or exposure on a given outcome or response variable. Since the work published in Ref. [75], randomized experiments have become a powerful and influential tool for the evaluation of causal effects; however, they are not feasible in many situations due to ethical issues, expensive cost, or imperfect compliance. In contrast, observational studies offer an important source of data for scientific research. However, causal inference with observational studies is challenging, because confounding may arise. Confounders are covariates that affect both the primary exposure and the outcome. In the presence of unmeasured confounders, statistical association does not imply causation, and vice versa, which is known as the Yule-Simpson paradox [19,20]. Refs. [87,88] review the concepts of confounding, and Refs. [2,89,90] discuss methods for the adjustment of observed confounders, such as regressing analysis, propensity score, and inverse probability weighting, as well as doubly robust methods. Here, we review two methods for the adjustment of unmeasured confounding: the instrumental variable approach and the negative control approach.

Throughout, we let X and Y denote the exposure and outcome of interest, respectively, and we let U[§] denote an unmeasured confounder; for simplicity, we omit observed confounders, which can be incorporated in the following by simply conditioning on them. We use lowercase letters to denote realized values of random variables—for example, y for a realized value of Y.

The instrumental variable approach, which was first proposed in econometrics literature in the 1920s [91,92], has become a popular method in observational studies to mitigate the problem of unobserved confounding. In addition to the primary treatment and outcome, this approach involves an instrumental variable Z that satisfies three core assumptions:

(i) It has no direct effect on the outcome, that is, $Z \perp Y | (X, U)$ (exclusion restriction);
(ii) It is independent of the unobserved confounder, that is, $Z \perp U$ (independence);
(iii) It is associated with the exposure, that is, $Z \not\perp X$ (relevance).

Under these three assumptions, only certain upper and lower bounds of causal effects can be derived [93,94], and extra model assumptions are required to achieve identification. The SEM [91,95] and structural mean model [96] are commonly used models, which in fact can achieve identification by assuming effect homogeneity (see Section 16 of Ref. [97]). One such example is the linear regression model $E(Y|X, U) = \alpha + \beta X + U$, which encodes a constant causal effect in the regression coefficient β and yields the well-known instrumental variable identification $\beta_{\text{iv}} = \frac{\sigma_{zy}}{\sigma_{xz}}$. Alternatively, in certain situations, especially when Z is a binary treatment assignment that occurs before X, it is sometimes reasonable to assume effect monotonicity: The effect of Z on X is monotone, that is, $X_{Z=1} \geq X_{Z=0}$, which means that no one accepts the opposite treatment of this assignment. The monotonicity assumption leads to identification of the complier average causal effect (CACE) = $E(Y_1 - Y_0 \mid X_1 = 1, X_0 = 0)$, as shown in Ref. [98]. As an extension of the single instrument case, Refs. [99,100] consider variable selection and estimation with high-dimensional instrumental variables.

However, in practice, the instrumental variable assumptions may not be met, and the approach is highly sensitive to the violation of any of them. Validity checking and violation detection of these assumptions are important before applying the instrumental variable approach, and have been attracting researchers' attention [94,101]. In case of a violation of the core assumptions, identification of the causal effect is often impossible, and bounding and sensitivity analysis methods [102,103] have been proposed for causal inference.

Alternatively, we have formally established the double negative control method [104–106] for the adjustment of unmeasured confounding. The negative control approach we have proposed also offers a promising mitigation tool for invalid instrumental variables. Negative control variables are classified into two classes: negative control outcome $W : W \perp X | U$, $W \not\perp U$ and negative control exposure $Z : Z \perp Y | (U, X), Z \perp W | (U, X)$. The negative control exposure Z can be viewed as a generalization of an instrumental variable that fails to be independent of the unmeasured confounder, and the negative control outcome W is used to eliminate

§ In this section, we reuse U to denote the unmeasured confounders. Please note that, U was used to denote an environment in Section 4.

the bias. Given both a negative control exposure and outcome, Refs. [104,106] show that the average causal effect is nonparametrically identified under certain regularity conditions. For illustration, consider again the regression model $E(Y|X, U) = \alpha + \beta X + U$, and assume that $E(W|U)$ also follows a linear model; then, β can be identified by the following:

$$\beta_{\text{nc}} = \frac{\sigma_{xw}\sigma_{zy} - \sigma_{xy}\sigma_{zw}}{\sigma_{xw}\sigma_{xz} - \sigma_{xx}\sigma_{zw}}$$

This formula does apply to a valid instrumental variable; in which case, $Z \perp U$, and thus, $\sigma_{zw} = 0$, according to the negative control outcome assumption. Therefore, the instrumental variable identification can be viewed as a special case of the negative control approach. However, in contrast to the instrumental variable, negative controls require weak assumptions that are more likely to hold in practice. Refs. [107,108] provide elegant surveys on the existence of negative controls in observational studies. Refs. [105,109] point out that negative controls are widely available in time series studies, as long as no feedback effect is present, such as studies about air pollution and public health.

Refs. [107,109,110] examine the use of negative controls for confounding detection or bias reduction when a solely negative control exposure or outcome is available but are unable to achieve identification. Refs. [111,112] propose the use of multiple negative control outcomes to remove confounding in statistical genetics but must rest on a factor analysis model.

9. Causal inference with interference

The stable unit treatment value assumption plays an important role in the classical potential outcomes framework. It assumes that there is no interference between units [76]. However, interference is likely to be present in many experimental and observational studies, where units socially or physically interact with each other. For example, in educational or social sciences, people enrolled in a tutoring or training program may have an effect on those not enrolled due to the transmission of knowledge [113,114]. In epidemiology, the prevention measures for infectious diseases may benefit unprotected people by reducing the probability of contagion [115,116]. In these studies, one unit's treatment can have a direct effect on its own outcome as well as a spillover effect on the outcome of other units. The direct and spillover effects are of scientific or societal interest in real problems; they enable an understanding of the mechanism of a treatment effect, and provide guidance for policy making and implementation.

In the presence of interference, the number of potential outcomes of a unit grows exponentially with the number of units.[**] As a result, it is intractable to estimate the direct and spillover effects without restriction in the literature on the estimation of treatment effects with interference structure. There has been a rapidly growing interest in interference (see Ref. [117] for a recent review). A significant direction of work focuses on limited interference within non-overlapping clusters and assumes that there is no interference between clusters [52,114,118–122]. This is referred to as the partial interference assumption [114]. Recently, several researchers have considered the relaxation of the partial interference assumption to account for a more general structure of interference [e.g., 123–126]. The variance estimation is more complicated under interference. As pointed out in Ref. [118], it is difficult to calculate the variances for the direct and spillover effects even under partial interference. In model-free settings, a typical assumption for obtaining valid variance estimation is that the outcome of a unit depends on the treatments of other units only through a function of the treatments. Ref. [118] provides a variance estimator under the stratified interference assumption, and Ref. [124] generalizes it under a weaker assumption.

Another direction of work targets new designs to estimate treatment effects based on the interference structure. Under the partial interference assumption, Ref. [118] proposes the two-stage randomized experiment as a general experimental solution to the estimation of the direct and spillover effects. In more complex structures such as social networks, researchers have proposed several designs for the point and variance estimation of the treatment effects [127–129].

For the inference under interference, Refs. [130,131] rely on models for the potential outcomes. Ref. [79] develops a conditional randomization test for the null hypothesis of no spillover effect. Ref. [80] extends this test to a larger class of hypotheses restricted to a subset of units, known as focal units. Building on this work, Ref. [132] provides a general procedure for obtaining powerful conditional tests.

Interference brings up new challenges. First, the asymptotic properties require advanced techniques deriving. Ref. [133] investigates the consistency of the difference in the means estimator when the number of the units that can be interfered with does not grow as quickly as the sample size. Ref. [134] develops the central limit theorem for direct and spillover effects under partial interference and stratified interference. Ref. [52] provides the central limit theorem for a peer effect under partial interference and stratified interference. However, under general interference, the asymptotic properties remain unsolved—even for the simplest difference in the means estimator. Second, interference becomes even harder to deal with when data complications are present. Refs. [120,121,135,136] consider noncompliance in an interference setting. Ref. [137] examines the censoring of time-to-event data in the presence of interference. However, for other data complications such as missing data and measurement error, no methods are yet available. Third, most of the literature focuses on the direct effect and the spillover effect. However, interference may be present in other settings, such as mediation analysis (see Ref. [138] for a mediation analysis under interference) and longitudinal studies, where different quantities are of interest. As a result, it is necessary to generalize the commonly used methods in these settings to account for the interference between units.

Compliance with ethics guidelines

Kun Kuang, Lian Li, Zhi Geng, Lei Xu, Kun Zhang, Beishui Liao, Huaxin Huang, Peng Ding, Wang Miao, and Zhichao Jiang declare that they have no conflict of interest or financial conflicts to disclose.

References

[1] Imbens GW, Rubin DB. Causal inference for statistics, social, and biomedical sciences. New York: Cambridge University Press; 2015.
[2] Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. Biometrics 2005;61(4):962–73.
[3] Kuang K, Cui P, Li B, Jiang M, Yang S, Wang F. Treatment effect estimation with data-driven variable decomposition. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence; 2017 Feb 4–9; San Francisco, CA, USA, 2017.
[4] Athey S, Imbens GW, Wager S. Approximate residual balancing: debiased inference of average treatment effects in high dimensions. J R Stat Soc Ser B (Stat Methodol) 2018;80(4):597–623.
[5] Kuang K, Cui P, Li B, Jiang M, Yang S. Estimating treatment effect in the wild via differentiated confounder balancing. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2017 Aug 13–17; Halifax, NS, Canada. p. 265–74.
[6] Imai K, Van Dyk DA. Causal inference with general treatment regimes: generalizing the propensity score. J Am Stat Assoc 2004;99(467):854–66.
[7] Egami N, Imai K. Causal interaction in factorial experiments: application to conjoint analysis. J Am Stat Assoc 2019;114(526):529–40.

[**] If the total number of units is N, then there are 2^N potential outcomes for each unit.

[8] Louizos C, Shalit U, Mooij JM, Sontag D, Zemel R, Welling M. Causal effect inference with deep latent-variable models. In: Proceedings of Advances in Neural Information Processing Systems 30; 2017 Dec 4–9; Long Beach, CA, USA. p. 6446–56.
[9] Crump RK, Hotz VJ, Imbens GW, Mitnik OA. Dealing with limited overlap in estimation of average treatment effects. Biometrika 2009;96(1):187–99.
[10] Li F, Thomas LE, Li F. Addressing extreme propensity scores via the overlap weights. Am J Epidemiol 2019;188(1):250–7.
[11] Kuang K, Cui P, Athey S, Xiong R, Li B. Stable prediction across unknown environments. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; 2018 Aug 19–23; London, UK. p. 1617–26.
[12] Zhuang Y, Wu F, Chen C, Pan Y. Challenges and opportunities from big data to Knowledge in AI2.0. Front Inf Technol Elec Eng 2017;18(1):3–14.
[13] Pan Y. 2018 special issue on artificial intelligence 2.0: theories and applications. Front Inf Technol Elec Eng 2018;19(1).
[14] Hoerl C, McCormack T, Beck SR, editors. Understanding counterfactuals, understanding causation: issues in philosophy and psychology. New York: Oxford University Press; 2011.
[15] Pearl J, Glymour M, Jewell NP. Causal inference in statistics: a primer. Hoboken: John Wiley & Sons; 2016.
[16] Daniel RM, De Stavola BL, Vansteelandt S. Commentary: the formal approach to quantitative causal inference in epidemiology: misguided or misrepresented? Int J Epidemiol 2016;45(6):1817–29.
[17] Pearl J. Causal and counterfactual inference. Forthcoming section in the handbook of rationality. Cambridge: MIT press; 2018.
[18] Goldfeld K. Considering sensitivity to unmeasured confounding: part 1 [Internet]. New York: Keith Golgfeld; 2019. Jan 2 [cited 2019 Jun 1]. Available from:https://www.rdatagen.net/post/what-does-it-mean-if-findings-are-sensitive-to-unmeasured-confounding/.
[19] Yule GU. Notes on the theory of association of attributes in statistics. Biometrika 1903;2(2):121–34.
[20] Simpson EH. The interpretation of interaction in contingency tables. J R Stat Soc B 1951;13(2):238–41.
[21] Chen H, Geng Z, Jia J. Criteria for surrogate end points. J R Stat Soc Series B Stat Methodol 2007;69(5):919–32.
[22] Geng Z, Liu Y, Liu C, Miao W. Evaluation of causal effects and local structure learning of causal networks. Annu Rev Stat Appl 2019;6(1):103–24.
[23] Pearl J. Is scientific knowledge useful for policy analysis? A peculiar theorem says: no. J Causal Infer 2014;2(1):109–12.
[24] Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med 1996;125(7):605–13.
[25] Xu L, Pearl J. Structuring causal tree models with continuous variables. in Proc. 3rd Annu. Conf. Uncertainty in Artificial Intelligence, Seattle, USA, pp. 170-179, 1987.
[26] Xu L. Deep bidirectional intelligence: alphazero, deep IA-search, deep IA-infer, and TPC causal learning. Appl Inf 2018;5(5).
[27] Xu L. Machine learning and causal analyses for modeling financial and economic data. Appl Inf 2018;5(11).
[28] Spirtes P, Glymour C, Scheines R. Causation, prediction, and search. 2nd ed. Cambridge: MIT Press; 2001.
[29] Pearl J. Causality: models, reasoning, and inference. Cambridge: Cambridge University Press; 2000.
[30] Spirtes P, Zhang K. Causal discovery and inference: concepts and recent methodological advances. Appl Inform 2016;3(1):3.
[31] Shimizu S, Hoyer PO, Hyvärinen A, Kerminen A. A linear non-gaussian acyclic model for causal discovery. J Mach Learn Res 2006;7:2003–30.
[32] Zhang K, Hyvärinen A. On the identifiability of the post-nonlinear causal model. In: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence; 2009 Jun 18–21; Montreal, QC, Canada. Arlington: AUAI Press; 2019. p. 647–55.
[33] Hoyer PO, Janzing D, Mooij JM, Peters J, Scholkopf B. Nonlinear causal discovery with additive noise models. In: Proceedings of International Conference on Neural Information Processing Systems; 2008 Dec 8–13; Vancouver, BC, Canada. p. 689–96.
[34] Zhang K, Hyvärinen A. Causality discovery with additive disturbances: an information-theoretical perspective. In: Buntine W, Grobelnik M, Mladenić D, Shawe-Taylor J, editors. Machine learning and knowledge discovery in databases. Berlin: Springer; 2009. p. 570–85.
[35] Zhang K, Schölkopf B, Muandet K, Wang Z. Domain adaptation under target and conditional shift. In: Proceedings of the 30th International Conference on Machine Learning; 2013 Jun 16–21; Atlanta, GA, USA. p. 819–27.
[36] Baroni P, Gabbay DM, Giacomin M, Van der Torre L. Handbook of formal argumentation. London: College Publications; 2018.
[37] Osborne J. Arguing to learn in science: the role of collaborative, critical discourse. Science 2010;328(5977):463–6.
[38] Shoham Y. Nonmonotonic reasoning and causation. Cogn Sci 1990;14(2):213–52.
[39] Liao B, Jin L, Koons RC. Dynamics of argumentation systems: a division-based method. Artif Intell 2011;175(11):1790–814.
[40] Sklar EI, Azhar MQ. Explanation through argumentation. In: Proceedings of the 6th International Conference on Human-Agent Interaction; 2018 Dec 15–18; Southampton, UK. p. 277–85.
[41] Fazzinga B, Flesca S, Furfaro F. Complexity of fundamental problems in probabilistic abstract argumentation: beyond independence. Artif Intell 2019;268:1–29.
[42] Pearl J. On a class of bias-amplifying variables that endanger effect estimates. In: Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence; 2010 Jul 8–11; Catalina Island, CA, USA. p. 425–32.
[43] Kempthorne O. The design and analysis of experiments. New York: Wiley; 1952.
[44] Scheffe H. The analysis of variance. New York: John Wiley & Sons; 1959.
[45] Hinkelmann K, Kempthorne O. Design and analysis of experiments: volume 1: introduction to experimental design. 2nd ed. New York: John Wiley & Sons; 2007.
[46] Imbens GW, Rubin DB. Causal inference for statistics, social, and biomedical sciences: an introduction. New York: Cambridge University Press; 2015.
[47] Splawa-Neyman J. On the application of probability theory to agricultural experiments: essay on principles. Section 9. Stat Sci 1990;5(4):465–72.
[48] Ding P, Dasgupta T. A randomization-based perspective on analysis of variance: a test statistic robust to treatment effect heterogeneity. Biometrika 2018;105(1):45–56.
[49] Dasgupta T, Pillai NS, Rubin DB. Causal inference from $2^{??}$ factorial designs by using potential outcomes. J R Stat Soc Series B Stat Methodol 2015;77(4):727–53.
[50] Wu J, Ding P. Randomization tests for weak null hypotheses. 2018. arXiv:1809.07419.
[51] Miratrix LW, Sekhon JS, Yu B. Adjusting treatment effect estimates by post-stratification in randomized experiments. J R Stat Soc Series B Stat Methodol 2013;75(2):369–96.
[52] Li X, Ding P, Lin Q, Yang D, Liu JS. Randomization inference for peer effects. J Am Stat Assoc 2019:1–31.
[53] Li X, Ding P. General forms of finite population central limit theorems with applications to causal inference. J Am Stat Assoc 2017;112(520):1759–69.
[54] Zhao A, Ding P, Mukerjee R, Dasgupta T. Randomization-based causal inference from split-plot designs. Ann Stat 2018;46(5):1876–903.
[55] Mukerjee R, Dasgupta T, Rubin DB. Using standard tools from finite population sampling to improve causal inference for complex experiments. J Am Stat Assoc 2018;113(522):868–81.
[56] Fisher R. Statistical methods for research workers. Edinburgh: Oliver and Boyd; 1925.
[57] Freedman DA. On regression adjustments to experimental data. Adv Appl Math 2008;40(2):180–93.
[58] Lin W. Agnostic notes on regression adjustments to experimental data: reexamining Freedman's critique. Ann Appl Stat 2013;7(1):295–318.
[59] Eicker F. Limit theorems for regressions with unequal and dependent errors. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1967 Jun 21–Jul 18; Berkeley, CA, USA. Berkeley: University of California Press; 1967. p. 59–82.
[60] Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1967 Jun 21–Jul 18; Berkeley, CA, USA. p. 221–33.
[61] White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980;48(4):817–38.
[62] Bloniarz A, Liu H, Zhang CH, Sekhon JS, Yu B. Lasso adjustments of treatment effect estimates in randomized experiments. Proc Natl Acad Sci USA 2016;113(27):7383–90.
[63] Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1996;58(1):267–88.
[64] Lei L, Ding P. Regression adjustment in completely randomized experiments with a diverging number of covariates. 2018. arXiv:1806.07585.
[65] Ding P, Feller A, Miratrix L. Decomposing treatment effect variation. J Am Stat Assoc 2019;114(525):304–17.
[66] Lu J. Covariate adjustment in randomization-based causal inference for $2^{??}$ factorial designs. Stat Probab Lett 2016;119:11–20.
[67] Middleton JA. A unified theory of regression adjustment for design-based inference. 2018. arXiv:1803.06011.
[68] Cox DR. Randomization and concomitant variables in the design of experiments. In: Anderson TW, Styan GHP, Kallianpur GG, Krishnaiah PR, Ghosh JK, editors. Statistics and probability: essays in honor of CR Rao. Amsterdam: North-Holland; 1982. p. 197–202.
[69] Morgan KL, Rubin DB. Rerandomization to improve covariate balance in experiments. Ann Stat 2012;40(2):1263–82.
[70] Li X, Ding P, Rubin DB. Asymptotic theory of rerandomization in treatment-control experiments. Proc Natl Acad Sci USA 2018;115(37):9157–62.
[71] Morgan KL, Rubin DB. Rerandomization to balance tiers of covariates. J Am Stat Assoc 2015;110(512):1412–21.
[72] Branson Z, Dasgupta T, Rubin DB. Improving covariate balance in $2^{??}$ factorial designs via rerandomization with an application to a New York City department of education high school study. Ann Appl Stat 2016;10(4):1958–76.
[73] Li X, Ding P, Rubin DB. Rerandomization in 2?? factorial experiments. 2018. arXiv:1812.10911.
[74] Zhou Q, Ernst PA, Morgan KL, Rubin DB, Zhang A. Sequential rerandomization. Biometrika 2018;105(3):745–52.
[75] Fisher RA. The design of experiments. Edinburgh: Oliver and Boyd; 1935.
[76] Rubin DB. Comment on "randomization analysis of experimental data: the Fisher randomization test". J Am Stat Assoc 1980;75(371):591–3.
[77] Tukey JW. Tightening the clinical trial. Control Clin Trials 1993;14(4):266–85.
[78] Rosenbaum PR. Covariance adjustment in randomized experiments and observational studies. Stat Sci 2002;17(3):286–327.

Please cite this article as: K. Kuang, L. Li, Z. Geng et al., Causal Inference, Engineering, https://doi.org/10.1016/j.eng.2019.08.016

[79] Aronow PM. A general method for detecting interference between units in randomized experiments. Sociol Methods Res 2012;41(1):3–16.
[80] Athey S, Eckles D, Imbens GW. Exact *p*-values for network interference. J Am Stat Assoc 2018;113(521):230–40.
[81] Basse G, Feller A, Toulis P. Exact tests for two-stage randomized designs in the presence of interference. 2017. arXiv:1709.08036.
[82] Ding P. A paradox from randomization-based causal inference. Stat Sci 2017;32(3):331–45.
[83] Rosenbaum PR. Exact confidence intervals for nonconstant effects by inverting the signed rank test. Am Stat 2003;57(2):132–8.
[84] Rigdon J, Hudgens MG. Randomization inference for treatment effects on a binary outcome. Stat Med 2015;34(6):924–35.
[85] Li X, Ding P. Exact confidence intervals for the average causal effect on a binary outcome. Stat Med 2016;35(6):957–60.
[86] Ding P, Li F. Causal inference: a missing data perspective. Stat Sci 2018;33 (2):214–37.
[87] Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. Stat. Sci 1999;14:29–46.
[88] Greenland S, Pearl J. Adjustments and their consequences—collapsibility analysis using graphical models. Int Stat Rev 2011;79(3):401–26.
[89] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41–55.
[90] Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. J Am Stat Assoc 1952;47(260):663–85.
[91] Wright PG. Tariff on animal and vegetable oils. New York: Macmillan; 1928.
[92] Heckman J. Instrumental variables: a study of implicit behavioral assumptions used in making program evaluations. J Hum Resour 1997;32 (3):441–62.
[93] Manski CF. Nonparametric bounds on treatment effects. Am Econ Rev 1990;80(2):319–23.
[94] Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. J Am Stat Assoc 1997;92(439):1171–6.
[95] Goldberger AS. Structural equation methods in the social sciences. Econometrica 1972;40(6):979–1001.
[96] Robins JM. Correcting for non-compliance in randomized trials using structural nested mean models. Commun Stat Theory Method 1994;23 (8):2379–412.
[97] Hernán MA, Robins JM. Causal inference. Boca Raton: Chapman & Hall; 2011.
[98] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996;91(434):444–55.
[99] Lin W, Feng R, Li H. Regularization methods for high-dimensional instrumental variables regression with an application to genetical genomics. J Am Stat Assoc 2015;110(509):270–88.
[100] Kang H, Zhang A, Cai TT, Small DS. Instrumental variables estimation with some invalid instruments and its application to mendelian randomization. J Am Stat Assoc 2016;111(513):132–44.
[101] Wang L, Robins JM, Richardson TS. On falsification of the binary instrumental variable model. Biometrika 2017;104(1):229–36.
[102] Manski CF, Pepper JV. Monotone instrumental variables: with an application to the returns to schooling. Econometrica 2000;68(4):997–1010.
[103] Small DS. Sensitivity analysis for instrumental variables regression with overidentifying restrictions. J Am Stat Assoc 2007;102(479):1049–58.
[104] Miao W, Geng Z, Tchetgen Tchetgen EJ. Identifying causal effects with proxy variables of an unmeasured confounder. Biometrika 2018;105 (4):987–93.
[105] Miao W, Tchetgen Tchetgen E. Invited commentary: bias attenuation and identification of causal effects with multiple negative controls. Am J Epidemiol 2017;185(10):950–3.
[106] Miao, W., Shi, X., Tchetgen E., Tchetgen. A. confounding cridge approach for couble negative control inference on causal effects 2018. arXiv:1808.04945.
[107] Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology 2010;21(3):383–8.
[108] Smith GD. Negative control exposures in epidemiologic studies. Epidemiology 2012;23(2):350–1.
[109] Flanders WD, Strickland MJ, Klein M. A new method for partial correction of residual confounding in time-series and other observational studies. Am J Epidemiol 2017;185(10):941–9.
[110] Rosenbaum PR. The role of known effects in observational studies. Biometrics 1989;45(2):557–69.
[111] Wang J, Zhao Q, Hastie T, Owen AB. Confounder adjustment in multiple hypothesis testing. Ann Stat 2017;45(5):1863–94.
[112] Gagnon-Bartsch JA, Speed TP. Using control genes to correct for unwanted variation in microarray data. Biostatistics 2012;13(3):539–52.
[113] Hong G, Raudenbush SW. Evaluating kindergarten retention policy: a case study of causal inference for multilevel observational data. J Am Stat Assoc 2006;101(475):901–10.
[114] Sobel ME. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. J Am Stat Assoc 2006;101 (476):1398–407.
[115] Halloran ME, Struchiner CJ. Causal inference in infectious diseases. Epidemiology 1995;6(2):142–51.
[116] Halloran ME, Struchiner CJ. Study designs for dependent happenings. Epidemiology 1991;2(5):331–8.
[117] Halloran ME, Hudgens MG. Dependent happenings: a recent methodological review. Curr Epidemiol Rep 2016;3(4):297–305.
[118] Hudgens MG, Halloran ME. Toward causal inference with interference. J Am Stat Assoc 2008;103(482):832–42.
[119] Basse G, Feller A. Analyzing two-stage experiments in the presence of interference. J Am Stat Assoc 2018;113(521):41–55.
[120] Forastiere L, Mealli F, VanderWeele TJ. Identification and estimation of causal mechanisms in clustered encouragement designs: disentangling bed nets using bayesian principal stratification. J Am Stat Assoc 2016;111 (514):510–25.
[121] Kang H, Imbens G. Peer encouragement designs in causal inference with partial interference and identification of local average network effects. 2016. arXiv:1609.04464.
[122] Rigdon J, Hudgens MG. Exact confidence intervals in the presence of interference. Stat Probab Lett 2015;105:130–5.
[123] Aronow PM, Samii C, 2013. Estimating average causal effects under interference between units. arXiv preprint arXiv:1305.61563, 16.
[124] Aronow PM, Samii C. Estimating average causal effects under general interference, with application to a social network experiment. Ann Appl Stat 2017;11(4):1912–47.
[125] Choi D. Estimation of monotone treatment effects in network experiments. J Am Stat Assoc 2017;112(519):1147–55.
[126] Forastiere L, Airoldi EM, Mealli F. Identification and estimation of treatment and interference effects in observational studies on networks. 2016. arXiv:1609.06245.
[127] Eckles D, Karrer B, Ugander J. Design and analysis of experiments in networks: reducing bias from interference. J Causal Inference 2017;5(1).
[128] Eckles D, Kizilcec RF, Bakshy E. Estimating peer effects in networks with peer encouragement designs. Proc Natl Acad Sci USA 2016;113(27):7316–22.
[129] Jagadeesan R, Pillai N, Volfovsky A. Designs for estimating the treatment effect in networks with interference. 2017. arXiv:1705.08524.
[130] Bowers J, Fredrickson MM, Panagopoulos C. Reasoning about interference between units: a general framework. Polit Anal 2013;21(1):97–124.
[131] Toulis P, Kao E. Estimation of causal peer influence effects. In: Proceedings of 30th International Conference on Machine Learning; 2013 Jun 16–21; Atlanta, GA, USA. p. 1489–97.
[132] Basse GW, Feller A, Toulis P. Randomization tests of causal effects under interference. Biometrika 2019;106(2):487–94.
[133] Sävje F, Aronow PM, Hudgens MG. Average treatment effects in the presence of unknown interference. 2017. arXiv:1711.06399.
[134] Liu L, Hudgens MG. Large sample randomization inference of causal effects in the presence of interference. J Am Stat Assoc 2014;109(505):288–301.
[135] Imai K, Jiang Z, Malani A. Causal inference with interference and noncompliance in two-stage randomized experiments. Princeton: Technical report Princeton University; 2018.
[136] Kang H, Keele L. Spillover effects in cluster randomized trials with noncompliance. 2018. arXiv:1808.06418.
[137] Loh WW, Hudgens MG, Clemens JD, Ali M, Emch ME. Randomization inference with general interference and censoring. 2018. arXiv:1803.02302.
[138] Vanderweele TJ, Hong G, Jones SM, Brown JL. Mediation and spillover effects in group-randomized trials: a case study of the 4Rs educational intervention. J Am Stat Assoc 2013;108(502):469–82.

Please cite this article as: K. Kuang, L. Li, Z. Geng et al., Causal Inference, Engineering, https://doi.org/10.1016/j.eng.2019.08.016