Ekonomi-tek, Volume 13, Number 2, 2024, 230-252

Received: August 6, 2024 *Accepted: September 1, 2024* *Review Article*

The Role of Pearl's Causal Framework in Empirical Research

Fırat BİLGEL*

Abstract

This paper underscores the necessity of formulating precise research questions that clarify causal relationships rather than simply identifying correlations and highlights the perils of relying solely on regression analysis in tackling complex causal inquiries without causal diagrams or structural causal models. It introduces Judea Pearl's causal epistemology, including causal graphs, structural causal models, and do-calculus as vital tools for estimating causal effects. It extends to the challenges of confounding and collider effects, the application of do-calculus with basic examples from Law & Economics and the advancements in causal discovery methods through constraint-based algorithms. The paper also offers a brief roadmap on best practices for identification and estimation.

Keywords: Pearl's causal epistemology, causal discovery, identification

JEL Codes: A12, C18, C51, K14

* MEF University, Department of Economics, bilgelf@mef.edu.tr, https://orcid.org/0000-0002-2585-5975

Pearl’in Nedensel Modelinin Ampirik Araştırmadaki Rolü

Öz

Bu çalışma, korelasyonları tespit etmek yerine neden-sonuç ilişkilerini açıklığa kavuşturan kesin araştırma soruları formüle etmenin gerekliliğini vurgulamakta ve karmaşık nedensel sorularla baş etmede, nedensel grafikler veya yapısal nedensel modeller olmadan sadece regresyon analizi kullanmanın tehlikelerine dikkat çekmektedir. Judea Pearl'ün nedensel epistemolojisinde kullanılan, nedensel grafikler, yapısal nedensel modeller ve do-kalkülüs gibi araçları nedensel etkileri tahmin etmek için tanıtır. Çalışma aynı zamanda karıştırıcı ve çarpışma etkileriyle ilgili zorluklara, Hukuk ve Ekonomi’den basit örneklerle, do-kalkülüs uygulamalarına ve tahdit temelli algoritmalar aracılığıyla nedensel keşif yöntemlerindeki gelişmelere değinmektedir. Makale ayrıca etki tanımlama ve tahmin konusunda en iyi uygulamalar hakkında kısa bir yol haritası sunar.

Anahtar Kelimeler: Pearl’ün nedensel epistemolojisi, nedensel keşif, etki tanımlama

JEL Kodları: A12, C18, C51, K14

1. Introduction

Strong research hinges on well-defined questions. Beyond simply identifying a topic of interest, a good research question should be free of ambiguity, address a specific and well-defined issue. In empirical research, it is crucial to distinguish between questions aimed at predicting outcomes and those seeking to establish causal relationships, as each requires different tools and approaches. For example, “what are the determinants of judicial bias?” is a valid research question if the goal is to predict judicial decisions based on a set of inputs. However, if the focus is on uncovering causal relationships, the question should be reformulated to target a specific cause-and-effect, such as “what is the causal effect of exposure to pretrial media coverage of criminal cases on sentencing outcomes?” This revised question focuses on the specific relationship as a form of judicial bias and is phrased to explore a causal relationship. Similarly, the question “what is the relation between income inequality and crime?” aims to identify a potential association but does not establish causality. Income inequality could be caused by other factors such as education, employment opportunities, or social policies that also cause crime rates. In response, one may attempt to control for these “other factors” in a regression context; however, not only does this practice fail to rectify the problem, it could make it worse if the researcher lacks causal language. A good question is “what is the impact of a progressive tax system on property crime rates?” This question is stronger because it suggests a mechanism where changes in the tax system might cause property crime by potentially altering income distribution and economic incentives.

This study underscores the critical importance of causal reasoning and advocates the adoption of Pearl’s causal framework as a complementary approach to existing tools to identify causal effects in economics and related fields. This approach not only addresses identification challenges but also enhances the rigor of empirical analysis through best practices that emphasize a solid understanding of causal relationships. The paper aims to serve as a guide and a starting point for researchers interested in understanding the basics of Pearl’s framework and employing structural causal models in empirical research. The paper also aims to equip economists with robust methodologies to discern causal mechanisms, ultimately leading to more informed policy decisions and a deeper understanding of socioeconomic phenomena.

Section 2 assesses the consequences of relying solely on regression analysis without a foundational understanding of causal relationships. It utilizes an example from sentencing and recidivism to illustrate the perils of drawing misleading inferences when conditioning on variables without considering their causal role. Section 3 discusses Pearl's causal epistemology and employs a hypothetical dataset on eyewitness identification and wrongful convictions to showcase how do-calculus can help clarify the distinction between deliberate intervention and passive observation. Section 4 discusses selected causal discovery algorithms that can be useful as a reinforcing

strategy along with causal reasoning. Section 5 provides a brief roadmap to a well-executed empirical analysis and Section 6 concludes.

2. Regression is Evil

Regression in general has, unfortunately, become a black box in empirical research. Many researchers mistakenly believe that simply applying Ordinary Least Squares (OLS) to their data can provide clear answers to complex causal questions (Sekhon, 2010). Modern causal inference methods, such as Structural Causal Models (SCMs) and Directed Acyclic Graphs (DAGs), provide a more rigorous framework for identifying and estimating causal effects. These methods help clarify the assumptions needed for causal inference and guide the proper conditioning on variables to avoid bias. Without identification, researchers may overlook important aspects of causal inference that estimation alone cannot address.

Inference based on observational data requires stronger assumptions relative to experimental data. Imagine that we are interested in the causal impact of sentencing on repeat-offending. While comparing recidivism rates for criminals with varying sentencing schemes, we make a strong and unrealistic assumption that criminals sentenced by different judges are otherwise identical in terms of the severity of crime, prior criminal record and other defendant characteristics such as age, socioeconomic status, mental health record and rehabilitation experience. We conjecture that criminals who committed serious crimes are more likely to re-offend compared to less severe crimes regardless of sentence length; those with a history of violence or repeat offenses might receive longer sentences; and age, mental health and socioeconomic status of the individual and participation in rehabilitation programs (e.g. substance abuse treatment, job training) can affect recidivism. We therefore collect data on all these observable characteristics and run a regression that looks like this:

$$Y_i = \alpha + \beta T_i + \delta' X_i + \varepsilon_i \quad (1)$$

where Y is a measure of recidivism, T is the severity of punishment, and X denotes all other observable differences mentioned above.

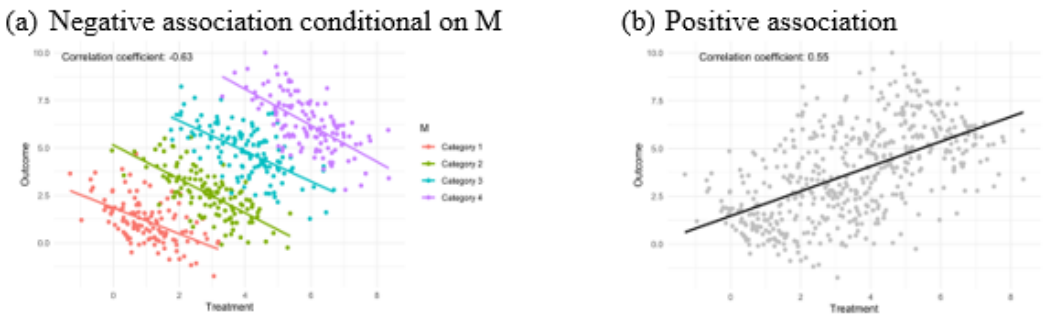
Those who receive harsher penalties might be systematically different from those who do not even after adjusting for observable confounders. For example, how about judges being more likely to impose harsher sentences on criminals that they perceive as high-risk for recidivism? Or how about criminals who are more likely to re-offend being less likely to respond positively to rehabilitation programs? Ignoring the first question leads to selection bias and the second to reverse causation. Unfortunately, this is not the

entire story. How could one possibly know that one should include every observable factor into eq. (1) the way we did without knowing how these factors interact with each other and with T and Y? As Pearl and Mackenzie (2018) note, data are agnostic to cause and effect, so are regressions.

In a randomized experiment, we could randomly assign convicted criminals to varying sentencing schemes. Although highly unethical, randomization of sentencing would rule out not only the effects of observable factors but also those that are not observable (either because data is missing or because we have not even considered they existed) and that bias the β estimate. Even if a randomized controlled trial (RCT) had been ethical, it may not be feasible for a multitude of factors.

To make my earlier point concrete, consider a hypothetical dataset that shows a mild positive correlation between the treatment (T) and the outcome (Y). Let us generate another variable, M, that is a multi-valued variable with four categories, M=1,2,3,4. Assume that we do not know anything about the type of relationship that M has vis-a-vis T and Y and simply include it as a covariate in the relationship between T and Y. The scatter plot of this situation is given in Figure 1a. For every category of M, it shows that T and Y are inversely related with a correlation coefficient of about -0.63. Then, the researcher plots another scatter diagram using the exact dataset, this time not stratified by M. This situation is given in Figure 1b. It shows a positive relationship between T and Y, with a correlation coefficient of +0.55. Which one should the researcher trust? This situation is a classical example of the Simpson's paradox and the answer depends on the nature of the triangular relationships among M, T and Y.

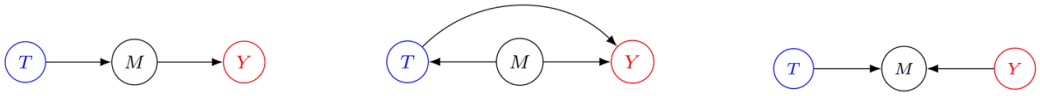
Figure 1. Scatter plot of a hypothetical dataset



If M mediates the relationship between T and Y, we have the situation in Figure 2a that shows a pure mediating effect, meaning that T would have no effect on Y in the absence of M. Controlling for M will cut off the flow of information from T to Y, hence the correct strategy is to not stratify by M (Figure 1b).

Figure 2. Three types of relationships

(a) M is a mediator (b) M is a confounder (c) M is a collider



For example, patent law grants inventors exclusive rights over their inventions (T) for a period to incentivize innovation (Y). This exclusivity allows the inventor to disclose (M) the details of their invention publicly but retain control over its commercial use. Disclosure of the invention (M) can directly stimulate further innovation (Y) by allowing other researchers and inventors to build upon the disclosed knowledge. Without proper disclosure (M), the details of the invention remain secret and a patent grant (T) might not directly lead to further innovation (Y).

If, on the other hand, M is a variable that confounds the relationship between T and Y, we have the situation in Figure 2b, showing that M is a common cause of T and Y. Here, the causal path is T → Y and T ← M → Y is a confounding path that creates a backdoor from T to Y. That is, if we follow the incoming arrow to T all the way back, we can reach Y through a path other than T → Y. If we do not adjust for M, this backdoor path will remain open (it leads us back to Y) and the effect of T on Y will be biased. It will not only contain the causal effect but also the impact of M on Y. Conditioning on M will close the backdoor T ← M → Y. Hence the correct strategy is to stratify by M (Figure 1a).

For example, patent grants (T) are often seen as a sign of a company's innovative potential, which can positively affect its stock price (Y). However, we know that high-quality inventions (M) with significant commercial potential are more likely to be patented (T) and that high-quality inventions (M) can also affect stock price (Y). Investors might see patents as a signal of innovation and future profitability, while high-quality inventions inherently hold greater commercial value.

Another most frequently encountered type of relationship is given in Figure 2c. Here, both T and Y cause M, which is known as a collider, referring to the fact that the

arrows emanating from T and Y collide on M. A stylized example from criminology is parole supervision (T), re-offending (Y) and employment (M) (Novak et al., 2023). In the population of parolees, there is no relation between parole supervision and re-offending. However, if we restrict the sample to employed individuals only, we induce an artificial correlation between parole supervision and re-offending.

The minority status of an individual (T) might affect the police use of force (M) due to potential racial biases in policing practices. Neighborhood crime rate (Y) could also cause police use of force (M) because areas with higher crime rates might see more frequent or aggressive police interventions. If we condition on or restrict our sample to cases in which police used force, there might appear to be a relationship between the minority status of the individual and crime rate even if they were unrelated to one another in the general population.

Hence, paths that include a collider that has not been conditioned upon are closed. Adjusting for or conditioning on a collider variable opens that path. Incorrect handling of collider variables can lead to false inference and misinterpretation of causal paths.

3. Pearl's Causal Framework

Judea Pearl's causal epistemology (Pearl, 2009; Pearl et al., 2016) offers a revolutionary framework for understanding causal relationships, particularly in situations where experimentation is not possible. By grounding causality in probability and graph theory, he developed a formal language for reasoning about interventions and counterfactuals. This framework has significantly advanced our ability to draw inferences from observational and experimental data, with profound implications in social sciences and medicine.

While Pearl's framework has undeniably transformed our approach to causal inference, it is not the only methodology available to researchers seeking to uncover causal relationships. Other causal inference methods include Structural Equation Modeling (SEM) and the Potential Outcomes (PO) Framework of Rubin (2005) and Imbens and Rubin (2010), which provide alternative, yet complementary approaches to causal inference.

SEM is a statistical approach that models relationships between variables using a system of (linear) equations. It can be viewed as a generalization of path analysis and allows researchers to specify and test complex causal relationships by incorporating both observed and unobserved variables.

Rubin's Potential Outcomes (PO) framework, also known as the Rubin Causal Model (RCM), is a foundational approach to causal inference in social sciences. The PO framework conceptualizes causality through counterfactuals. For each unit, there are potential outcomes corresponding to each possible treatment condition. The causal effect is defined as the difference between these potential outcomes.

The PO framework is widely used in economics, epidemiology, political science, and social sciences, particularly in studies where the goal is to estimate the causal effect of a treatment, intervention, or policy. It is also the basis for many modern causal inference techniques, including Difference-in-Differences (DiD), Instrumental Variables (IV), and Regression Discontinuity Design (RDD).

While the PO framework focuses on the estimation of causal effects through counterfactual reasoning and emphasizes the importance of treatment assignment and covariate adjustment, Pearl's Causal Framework provides a more formalized approach to understanding and modeling the structure of causal relationships and places more emphasis on the role of causal diagrams in identifying and clarifying assumptions about causal relationships.

Pearl proposes a hierarchical framework called the Ladder of Causation that categorizes causal knowledge into three levels: association (seeing), intervention (doing) and counterfactuals (imagining) (Pearl and Mackenzie, 2018). Association is the most basic level, where one simply observes a relationship between two variables. Intervention involves manipulating variable X to observe its effect on variable Y. Finally, counterfactuals deal with hypothetical scenarios of “What would have happened if...?” and they are crucial for causal inference. But they can only be imputed, not directly observed for they contradict what is seen. For instance, "What would have happened to the presidential race or the polling results had Donald Trump not been shot?" is a counterfactual question. Each unit has multiple potential outcomes, one for each possible treatment level. The causal effect for a unit is the difference between the potential outcome under treatment and the potential outcome under control. Unfortunately, we can only observe one potential outcome for each unit, known as the fundamental problem of causal inference (Holland, 1986). To estimate causal effects from observed data, one relies on multiple units of treatment and control.

A causal graph, or a directed acyclic graph (DAG), is a visual representation of causal relationships among variables. In a DAG, nodes represent variables and directed edges (arrows) indicate causal effects. For instance, if variable A causes B, an arrow points from A to B. A crucial aspect of DAGs is that they are acyclic, meaning that following the arrowheads, you cannot start from node A and return to itself. This ensures that causality runs in a single direction.

Directed Acyclic Graph (DAG):

A DAG is a graph $G = (V, E)$ where

1. V is a set of vertices (nodes) representing variables,
2. $E \subseteq V \times V$ is a set of directed edges (arrows) between vertices.

While DAGs offer a visual representation, Structural Causal Models (SCMs) provide a mathematical framework for specifying causal relationships (Pearl, 2009). A SCM comprises a set of variables, a set of functions that determine the value of each variable given the values of its parents in the DAG, and a probability distribution over the exogenous variables. SCMs describe how variables are generated based on their causes and encode our knowledge about causal relationships.

Structural causal model (Pearl, 2009):

A SCM consists of:

1. A set of endogenous variables $\{Y_1, Y_2, \ldots, Y_n\}$.
2. A set of exogenous variables $\{U_1, U_2, \ldots, U_n\}$.
3. A set of structural equations $\{f_1, f_2, \ldots, f_n\}$ such that each $Y_i = f_i(PA_i, U_i)$ where $PA_i \subseteq \{Y_1, Y_2, \ldots, Y_n\}$ represents the parent variables of Y_i and $U_i \subseteq \{U_1, U_2, \ldots, U_n\}$ are the exogenous variables affecting Y_i.

D-separation (Pearl, 1988):

D-separation (d stands for directional) plays a crucial role in understanding how the structure of a DAG encodes conditional independence between variables.

For all connecting paths between X and Y in the DAG, a path is considered blocked by Z if (*i*) it contains a collider where Z is not a descendant (does not have an incoming arrow to it) of the collider; and (*ii*) all other nodes along the path are not in Z. If all paths between X and Y are blocked by Z, then X and Y are d-separated given Z (Pearl, 1988). D-separation allows us to efficiently determine which variables are independent given others based solely on the DAG structure and to verify if the conditional independence assumptions encoded in the DAG structure hold.

Fourth aspect, the do-calculus is a mathematical framework within SCMs that allows us to formally analyze the effects of interventions (Pearl, 1995). It uses the concept of potential outcomes to represent the outcome that would have occurred under different treatment conditions (even if unobserved). By manipulating the SCM and applying the do-calculus, we can estimate the causal effect of an intervention on the outcome variable. The do-notation is characterized by the do-operator, denoted by do( ), used to represent an intervention. For example, $do(X = x)$ indicates setting variable X to a specific value x. After an intervention on X in causal graph G, the incoming arrow to the intervened variable

(X) can be removed, resulting in a mutilated graph $G_{\bar{X}}$. Similarly, if X has no effect on the outcome, we can express this by removing the outgoing arrow from X as $G_{\underline{X}}$.

The three rules of do-calculus are intervention rule, composition rule and reduction rule.

Intervention rule (insertion/deletion of observations):

The first rule of do-calculus iterates d-separation and states that we can ignore a node Z if it does not affect the outcome Y. As long as the path between Y and Z is blocked, conditional on X and W, the node Z can be removed from the do-expression:

$$Pr(Y|do(X = x), Z, W) = Pr(Y|do(X = x), W) \text{ if } Y \perp Z|X, W \text{ in } G_{\bar{X}} \quad (2)$$

which means "Y is independent of Z, given W and X" in the mutilated graph.

Composition rule (action/observation exchange):

The second rule of do-calculus allows us to combine multiple interventions. It states that the effect of an intervention $do(Z = z)$ has the same effect as observation Z = z if Y and Z are independent, conditional on X and W. For this, we need to remove all outgoing arrows from Z to reflect the premise that an intervened Z has the same effect as a passively observed Z:

$$Pr(Y|do(X = x), do(Z = z), W) = Pr(Y|do(X = x), Z, W) \text{ if } Y \perp Z|X, W \text{ in } G_{\bar{X}\underline{Z}} \quad (3)$$

Reduction rule (insertion/deletion of actions):

The third rule of do-calculus allows us to remove an intervention ($do(Z = z)$) from the expression completely if Z does not have a causal effect on the outcome Y, either directly or indirectly through any other variables in the model. Intuitively, if Z has no influence on Y, then intervening on it will not change the outcome distribution, and we can remove it from the expression:

$$Pr(Y|do(X = x), do(Z = z), W) = Pr(Y|do(X = x), W) \text{ if } Y \perp Z|X, W \text{ in } G_{\overline{XZ(W)}} \quad (4)$$

where Z(W) denotes any Z node that is not an ancestor or parent of W.

These rules are shown to be sufficient and complete in the sense that any acyclic causal relationship can be captured as long as a causal question can be represented by a DAG and an associated SCM, and that all causal questions can be answered by these rules (Huang and Valtorta, 2012).

Backdoor criterion (Pearl, 1995):

The backdoor criterion is a method for identifying a set of variables to adjust for or condition on when estimating a causal effect. A set of variables, Z, satisfies the backdoor criterion relative to an ordered pair of variables (X, Y) in a DAG if (*i*) no node in Z is a descendant of X and (*ii*) Z blocks every path between X and Y that contains an arrow into X (Pearl, 1995). In simpler terms, the backdoor criterion requires that the set Z does not itself be affected by the treatment (X) and that they block all alternative paths between the treatment and outcome that could confound the relationship. By conditioning on the variables in Z, we can control for confounding and estimate the causal effect of X on Y.

Frontdoor criterion (Pearl, 1995):

The frontdoor criterion is an alternative approach to estimating causal effects when the backdoor criterion cannot be satisfied. It involves a mediator variable, M, that lies on the causal pathway between the treatment, X, and the outcome, Y.

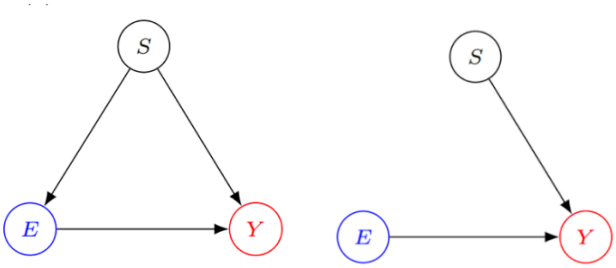
The conditions for the frontdoor criterion are (*i*) all directed paths from X to Y are blocked by M, (*ii*) there are no unblocked paths from X to M given W, and (*iii*) all backdoor paths from M to Y are blocked by X and W (Pearl, 1995). If these conditions hold, the causal effect of X on Y can be estimated by first estimating the effect of X on M, then the effect of M on Y while controlling for X and W, and finally multiplying these two estimates. The frontdoor criterion can be useful when it is difficult to measure or control for all confounders.[1]

Let us take on a stylized example from criminology in which wrongful conviction (Y) is a function of eyewitness identification (E) and stressful interview techniques used by police (S) only. Eyewitness identification is a common method used in criminal investigations. However, eyewitness memory can be unreliable, especially when influenced by factors like stress or leading questions. Hence, it can directly contribute to wrongful conviction if the eyewitness misidentifies an innocent person. On the other hand, stressful interview techniques used by police (S) can act as a confounder in the relationship between

[1] There are other core elements such as necessary and sufficient causation, counterfactuals, mediation, selection, transportability and Z-identifiability, which are beyond the scope of this paper. For a review of Pearl's causal framework, see Bareinboim and Pearl (2016); Bareinboim et al. (2022) and Hünermund and Bareinboim (2023).

E and Y. We conjecture that stressful interview techniques can pressure suspects into making false confessions or providing inaccurate identifications of others. A simple DAG of this situation is given in Figure 3.

Figure 3. Eyewitness Identification and Wrongful Conviction



(a) Causal graph G (b) Mutilated graph $G_{\bar{E}}$

Assume all three variables are binary and consider an individual-level dataset with 1000 conviction cases whose tabulation is given in Table 1. We are interested in the causal effect of eyewitness identification (E) on wrongful convictions (Y). We can express this causal query using the do-operator as $Pr\big(Y = 1|do(E = 1)\big)$. Intervening on E means that we are performing a surgery in Figure 3a by removing all arrows that goes into E. The resulting graph, called the mutilated graph, is given in Figure 3b.

Table 1. A hypothetical dataset on convictions, N = 1000

	Treatment (E = 1)			Treatment (E = 0)		
	Y = 1	**Y = 0**	**%Y = 1**	**Y = 1**	**Y = 0**	**%Y = 1**
Stressful techniques used (S = 1) $\frac{269}{100} = 0.27$	100	21	$\frac{100}{121} = 0.83$	132	16	$\frac{132}{148} = 0.89$
Stressful techniques not used (S = 0) $\frac{731}{100} = 0.73$	402	91	$\frac{402}{493} = 0.82$	200	38	$\frac{200}{238} = 0.84$
Total	502	112	$\frac{502}{614} = 0.82$	332	54	$\frac{332}{386} = 0.86$

Let us elaborate the causal query to obtain a do-free expression using the rules of do-calculus:

$$Pr\big(Y=1|do(E=1)\big)=\sum_S Pr\big(Y=1,S=s,do(E=1)\big) \quad (5)$$

$$=\sum_S Pr\big(Y=1,S=s,do(E=1)\big)Pr\big(S=s,do(E=1)\big)$$

$$=\sum_S Pr(Y=1,S=s,E=1)Pr(S=s)$$

$$=Pr(Y=1|S=1,E=1)Pr(S=1)+Pr(Y=1|S=0,E=1)Pr(S=0)$$

$$=(0.83\times 0.27)+(0.82\times 0.73)=0.134$$

The first expression is eq. (2) states that the probability of Y=1 given that we intervene to set E=1 is the sum over the joint distribution of Y=1 and all possible values of confounder S. The second expression applies the chain rule decomposing the joint probability into a product of conditional probabilities. The third expression uses the third rule of do calculus (insertion/deletion of observations), stating that S is not affected by the intervention on E. Hence $Pr\big(S=s,do(E=1)\big)=Pr(S=s)$. The fourth expression uses the second rule of do-calculus (action/observation exchange), stating that if S blocks all backdoor paths between E and Y, then conditioning on S is the same as intervening on E. Hence $=Pr\big(Y=1,S=s,do(E=1)\big)=Pr(Y=1,S=s,E=1)$. Plugging in the values from Table 1, the causal effect of eyewitness identification on wrongful convictions is 0.134. The association between E and Y on the other hand is $Pr(Y=1|E=1)=0.82\neq Pr\big(Y=1|do(E=1)\big)=0.134$. The discrepancy between the two estimates is due to confounding bias induced by ignoring S.

4. Causal Discovery

Causal reasoning and causal discovery, while intricately linked, address different aspects. Causal reasoning uses existing knowledge or a defined causal structure via SCMs or DAGs to draw conclusions about cause and effect. It analyzes the consequences of interventions or counterfactual scenarios. Causal reasoning therefore runs from a causal model and observational data to outcomes, changes and interventions. On the other hand, causal discovery runs in the opposite direction. It aims to identify the underlying causal structure between a set of variables from data.

Table 2 shows the number of DAGs that can be built for a given number of nodes, without incorporating any prior domain knowledge or restriction. For example, with 2 nodes, say A and B, either A causes B, or B causes A or neither causes the other (i.e. unrelated), giving rise to a total of 3 unique DAGs. Notice that the series is explosive,

reaching about 3.8 million unique DAGs with only 6 nodes. However, once prior knowledge and assumptions based on a given theory are embedded via required or forbidden edges between any two nodes and the implied conditional dependencies are conformable to the data, this number will reduce considerably. Even so, it is not possible to obtain a single DAG for a given problem with observational data, nor is it possible to ensure that prior knowledge and conditional dependencies will be one-on-one mapped with a causal model. A very large number of causal diagrams may be consistent with our background knowledge and the implied conditional dependencies, but a subset of these causal diagrams may lead to different inferences than others (Scheines et al., 1998). However, it is possible to learn a completed partially DAG (CPDAG) from data, which is an equivalence class of DAGs that contains some undirected edges.

Table 2. Number of DAGs as a function of nodes

nodes (j)	Number of DAGs with j nodes
2	3
3	25
4	543
5	29281
6	3781503
7	1138779265
8	783702329343
9	1213442454842881
10	4175098976430598143
11	3160345939641891760742 5
12	521939651343829405020504063
13	186766007444320351866648169267 21
14	143942814104439833494179071983 9535103
15	237725265553410354992180218286376719253505
16	837566707737333202876993030479964122352231383 03
17	627079211969238898994464526024949219069635514826752 01
18	994211953221595158952289145923545245165550268785883050147 83
19	332771901227107591736177573311261125883583076258421902583546773505
20	234488045105108898815255985522909918889908119223429129879580323606849126 3

Source: https://oeis.org/A003024/b003024.txt

Constraint-based methods use a set of logical constraints or rules to identify possible causal relationships. They focus on identifying constraints that must be satisfied by the causal structure consistent with the observed data. This section focuses on causal discovery algorithms that allow for latent confounding. In order to understand how these algorithms work, a good starting point is the Peter-Clark (PC) algorithm of Spirtes and Glymour (1991), one of the algorithms that uses a constraint-based approach. PC algorithm relies upon the assumptions of causal Markow condition (two variables that are d-separated are probabilistically independent), faithfulness (conditional independences that hold in the data implies absence of a direct causal relationship between them), absence of latent confounding (there are no unobservable common causes) and acyclicity (there are no cycles in the resulting graph, which is a CPDAG).

The PC algorithm consists of two phases. In the adjacency phase, it starts with a complete, undirected graph where every variable is connected to every other variable. The algorithm iteratively removes edges between variables based on d-separation. If finding a set of conditioning variables (Z) renders two variables (X and Y) independent, the undirected edge between X and Y is removed. This process continues until no more edges can be removed based on d-separation. In the orientation phase, it uses conditioning sets from the adjacency phase and determines the direction of as many edges as possible based on the information from the conditioning and specific collider and non-collider structures.

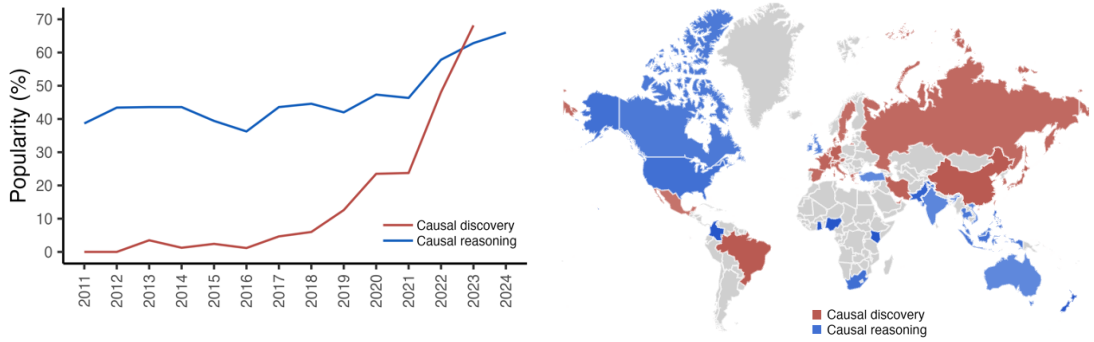
The Fast Causal Inference (FCI) algorithm, developed by Spirtes et al. (1993, 1999) and Spirtes et al. (2000) is a constraint-based algorithm that accepts sample data and background knowledge as inputs and produces an equivalence class of Causal Bayesian Networks (CBNs) that reflect the set of conditional independence relationships present in the population. It uses the adjacency and the orientation phases as in the PC algorithm. The resulting graph is a partial ancestral graph (PAG) in which all edges are directed but the pairwise relationships may contain latent confounders. Relative to PC, the FCI algorithm is computationally intensive due to the additional tests and rules for handling latent variables, making it suitable for more complex datasets.

Two notable variants of the FCI that allow for latent confounding whose resulting graph is also a PAG are the greedy FCI (GFCI) and really FCI (RFCI). The GFCI is a hybrid algorithm that combines constraint-based (FCI) and score-based (fast greedy search or FGES of Meek (1997) and Chickering (2002)) methods. RFCI, which is an improved version of FCI developed by Colombo et al. (2012), runs faster to mitigate computational inefficiencies of the FCI, making it suitable for larger datasets.[2]

[2] A comprehensive overview of causal discovery methods is given in Nogueira et al. (2022).

Figure 4 displays the worldwide popularity of causal reasoning and causal discovery over the last 14 years using data from Google Trends along with a world map showing countries where each topic is dominantly popular. Causal discovery has witnessed a significant surge in popularity in the last 5 years, recently surpassing that of causal reasoning. One particular reason of this surge lies in the popularity of Large Language Models (LLMs) that integrate causal reasoning capabilities and Causal AI, a branch of artificial intelligence that focuses on understanding and modeling cause-and-effect within data. Causal reasoning is the dominant web search topic, clustered in North America, Oceania, India, United Kingdom and South Africa whereas causal discovery is more popular in western Europe, Scandinavia, Asia and South America.

Figure 4. Trends in causal reasoning vs. causal discovery



Source: Google Trends under the topics “Causal discovery” and “Causal reasoning”.

5. A Roadmap

The roadmap to a well-executed empirical analysis consists of several critical steps. It begins with acquiring domain knowledge and making explicit assumptions, followed by building a causal graph to visualize relationships. The causal graph may be reinforced using causal discovery techniques to validate existing patterns or to uncover potential unforeseen causal links. The next step is refutation or validation of the causal model to ensure that the implied conditional independencies are satisfied. Subsequently, the causal query and the corresponding causal estimand are identified. Only then statistical methods should be employed to estimate the causal effect, followed by a series of falsification tests to verify its credibility.

5.1. The Causal Problem

"Give me a causal graph and I will tell you whether and if so how, a causal query can be answered." Although simplified, Pearl's framework is about an epistemological problem. "How a causal graph is constructed?" on the other hand is an ontological problem. For example, Pearl's framework cannot tell you what kind of relationships exist between income equality, crime and a plethora of surrounding factors. As Pearl notes on many occasions, this requires expert knowledge. The accuracy of a causal diagram can be seen as an ontological concern, as it reflects how well it represents the "true" causal structure of the world. However, Pearl's framework does not provide definitive answers about the "true" ontological nature of the system being studied. It offers a flexible tool to model causal relationships based on available knowledge and data.

The causal relationship we are trying to understand should be clearly articulated. What is the treatment and the outcome of interest? We need to list all the variables we believe might influence the relationship between the treatment and outcome, irrespective of whether they are measured or not. A DAG should not only consider measured variables but also acknowledge the existence of unmeasured variables. These can be factors we are aware of but have not been included in our data collection (known unknowns). Even more challenging are the variables we might not even be aware of that could be influencing the relationships in our model (unknown unknowns). Economic theory, existing literature, and domain knowledge should be considered here. Arranging variables/nodes in a chronological order based on when they are measured or occur helps determine the direction of the arrows. For each pair of nodes, we need to consider the underlying causal mechanisms and assess whether one variable directly causes the other and/or whether there are any measured or unmeasured mediators along the way or measured or unmeasured confounding factors influencing both.

You should resist the urge to look at the data before building a causal diagram for it can introduce a number of biases, even if you plan on using causal discovery algorithms along the way. Peeking at the data before constructing the DAG might lead to (*i*) favor evidence that confirms our initial hunch about the causal relationships and interpret patterns in a way that reinforces pre-existing ideas, potentially overlooking alternative causal structures that the data might also support; (*ii*) decide about which variables to include or exclude in the DAG, or how to connect them; (*iii*) or choose to focus on specific subsets or exclude outliers, leading to a biased representation of the causal relationships. You cannot possibly include every factor into a causal diagram, and no one expects you to do so. Essentially, what you need to incorporate are the ones that are highly-plausible to exist but also identification-altering. For example, if your DAG does not incorporate a known-unknown treatment-outcome confounder that you think should have been there but could

not incorporate because of data unavailability, this will change identification and consequently inference because it will introduce a backdoor that cannot be closed.

There are several packages that will help you construct and analyze causal diagrams using R statistical software (R Core Team, 2024): **DiagrammeR** (Iannone and Roy, 2024), **causaloptim** (Sachs et al., 2023), **causaleffect** (Tikka and Karvanen, 2017), **dagitty** (Textor et al., 2016), **MRPC** (Badsha et al., 2021), **ggdag** (Barrett, 2024) and **dosearch** (Tikka et al., 2021). Beyond these, there is an online software, **Causal Fusion** (Bareinboim and Pearl, 2016) that incorporates Pearl's framework, developed by Elias Bareinboim, Juan, D. Correa and Chris Jeong.[3] For Python users, the **DoWhy** library for causal inference would be an excellent alternative.[4] Unfortunately, as far as the identification and causal problems are concerned, Stata users are empty-handed.

Once all prior knowledge and assumptions about the causal question are embedded in the causal diagram, the conditional independencies implied by the causal graph should be checked using the data. If at least one conditional independency is not met by the data, the DAG should be revised. What makes the construction of a causal graph extremely challenging and intriguing are the unknown unknowns. Causal reasoning has little to offer in that respect. You cannot possibly incorporate a node that you are not aware of. Fortunately, there is at least a partial remedy via causal discovery, which is very useful once a reasonably accurate DAG is obtained. Causal discovery using observational data is not infallible as it can be sensitive to various assumptions and limitations of the data. Therefore, our background knowledge should steer the wheel.

There are several packages that will help you perform causal discovery in R: **pcalg** (Kalisch et al., 2012), **bnlearn** (Scutari et al., 2024), **causalDisco** (Petersen, 2022) and **tetrad** (Scheines et al., 1998; Ramsey and Andrews, 2023).[5]

5.2. The Statistical Problem

Once identification is complete, the next step is estimation or the statistical problem. There are many options to choose from, depending on the context, the nature of the data and most importantly, whether and if so, how identification can be achieved. It is crucial to ensure that the estimated effects are not driven by biases, confounders, or artifacts of the data. Falsification tests provide a rigorous approach to assess the robustness of causal estimates by introducing hypothetical scenarios or alternative models. One such test involves introducing a random confounder, an independent random variable added to the

[3] Available at: causalfusion.net (login required)
[4] Available at: https://www.pywhy.org/dowhy/v0.11.1/
[5] Tetrad is available at: https://github.com/cmu-phil/py-tetrad/tree/main/pytetrad/R

model as a common cause. If the causal estimate remains stable, this suggests that the original estimate is not driven by unmeasured confounding. A significant change in the estimate indicates potential vulnerability to confounding and raises concerns about the validity of the causal inference.

Further falsification strategies include placebo tests. A placebo treatment test examines whether replacing the actual treatment variable with an independent random variable significantly changes the estimate. The expectation is that the estimated effect should become indistinguishable from zero to reinforce the idea that the original treatment variable was indeed causally related to the outcome. Similarly, a placebo outcome test replaces the actual outcome variable with another, known to be unrelated to the treatment and other observables. Again, the estimated effect should not be significantly different from zero at conventional test levels.

To ensure that the estimates are consistent across different subsamples of data, data subset validation and bootstrap validation can be performed. In data subset validation, the dataset is replaced by random subsets of the same data. The estimated effects should be similar to the original estimate. Bootstrap validation involves replacing the dataset with bootstrap samples, where the estimated effects should also remain stable. Together, these tests assess the internal consistency and generalizability of the causal estimate, ensuring that the findings are not artifacts of a specific dataset.

Sensitivity analysis complements falsification tests by evaluating how sensitive the causal estimates are to various assumptions and potential deviations. A leave-one-out (LOO) analysis examines the stability of the estimates by iteratively excluding each unit from the sample. If the estimates change significantly with the exclusion of any single unit, it may indicate that the findings are overly dependent on particular observations, which could undermine the generalizability of the results.

Another critical aspect of sensitivity analysis is assessing unobserved confounding. This involves introducing a hypothetical unobserved confounder that is correlated with both the treatment and the outcome. By observing how the estimates change, researchers can gauge how robust their results are to potential biases from unmeasured variables. If the causal estimate remains relatively stable despite the introduction of a latent confounder, it suggests that the original estimate is likely to be robust. Similarly, one can assess the strength of a latent confounding that would explain away the estimate of the impact of treatment on the outcome via formal sensitivity analysis of Cinelli and Hazlett (2020). In the context of instrumental variables (IV), the identification of the average causal effect relies upon the instrument being (1) relevant (the instrument has a clear effect on treatment), (2) clean (the instrument affects the outcome only and only through the treatment) and (3) excludable (there is no latent confounding between the instrument and the outcome). While

(1) and (2) can be assessed, (3) is untestable. Therefore, researchers may be concerned with the possibility that the instrument may not satisfy the exclusion restriction and therefore may be invalid. One can conduct a sensitivity analysis to invalid IV of Wang et al. (2018) for this purpose.

6. Conclusion

Pearl's causal epistemology provides a robust framework for understanding and analyzing causal relationships in empirical research. Through the use of causal graphs, structural causal models, and do-calculus, researchers can move beyond mere correlations. The distinction between causal reasoning and causal discovery highlights the different methodologies employed to derive causal insights. While causal reasoning allows us to draw conclusions based on established causal structures, causal discovery seeks to identify these structures from the data. The feedback between the two approaches can help researchers better navigate the complexities of causal relationships and avoid pitfalls associated with regression analysis.

Graphical approaches, particularly DAGs, are powerful tools for identifying causal relationships, especially in microeconomic analysis where individual-level data often allow for a clear delineation of causal pathways. However, these methods are not without limitations. One significant challenge arises when applying DAGs to aggregated data, common in macroeconomic analysis. In such contexts, simultaneity, unmeasured confounders, and the nature of aggregated variables can obscure the causal relationships that DAGs are designed to clarify. The assumptions required for DAGs to provide valid causal insights may not hold as readily in aggregated datasets, where complex feedback loops and interdependencies are more prevalent. Acknowledging these challenges is important for a balanced understanding of the applicability of graphical causal models across different levels of economic analysis. While DAGs offer clear advantages in microeconomic contexts, their utility in macroeconomic analysis may be limited without careful consideration of the underlying data structure and potential biases introduced by aggregation.

References

Badsha, M. B., Martin, E. A., and Fu, A. Q. (2021). MRPC: An R package for inference of causal graphs. *Frontiers in Genetics*, 12:651812.

Bareinboim, E., Correa, J. D., Ibeling, D., and Icard, T. (2022). In H. Geffner, R. Dechter, and J. Y. Halpern (Eds.). *Probabilistic and Causal Inference: The Works of Judea Pearl* (1st. ed., pp. 507–556). New York: ACM Books, Vol. 36. Association for Computing Machinery.

Bareinboim, E. and Pearl, J. (2016). Causal inference and the data-fusion problem. In Shiffrin, R. M. (Ed.). *Proceedings of the National Academy of Sciences*, 113, 7345–7352.

Barrett, M. (2024). ggdag: Analyze and Create Elegant Directed Acyclic Graphs. R package version 0.2.12, https://r-causal.github.io/ggdag/.

Chickering, D. M. (2002). Optimal structure identification with greedy search. *Journal of machine learning research*, 3(Nov), 507–554.

Cinelli, C. and Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1), 39–67.

Colombo, D., Maathuis, M. H., Kalisch, M., and Richardson, T. S. (2012). Learning high-dimensional directed acyclic graphs with latent and selection variables. *The Annals of Statistics*, 40 (1), 294–321.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945–960.

Huang, Y. and Valtorta, M. (2012). Pearl's calculus of intervention is complete. arXiv preprint arXiv:1206.6831.

Hünermund, P. and Bareinboim, E. (2023). Causal inference and data fusion in econometrics. *The Econometrics Journal*, page utad008.

Iannone, R. and Roy, O. (2024). DiagrammeR: Graph/Network Visualization. R package version 1.0.11.9000, https://github.com/rich-iannone/DiagrammeR.

Imbens, G.W., Rubin, D.B. (2010). Rubin Causal Model. In: Durlauf, S.N., Blume, L.E. (Eds.) Microeconometrics. (pp. 229–241). London: The New Palgrave Economics Collection. Palgrave Macmillan.

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., and Bühlmann, P. (2012). Causal inference using graphical models with the R package pcalg. *Journal of Statistical Software*, 47, 1–26.

Meek, C. (1997). Graphical Models: Selecting causal and statistical models. Ph.d thesis, Carnegie Mellon University.

Nogueira, A. R., Pugnana, A., Ruggieri, S., Pedreschi, D., and Gama, J. (2022). Methods and tools for causal discovery and causal inference. Wiley interdisciplinary reviews: data mining and knowledge discovery, 12(2): e1449.

Novak, A., Boutwell, B. B., and Smith, T. B. (2023). Taking the problem of colliders seriously in the study of crime: A research note. *Journal of Experimental Criminology*, 1–10.

Pearl, J. (1988). Probabilistic reasoning in intelligent systems. San Mateo, CA: Morgan Kaufmann.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4), 669–688.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference* (2nd ed.). New York, NY: Cambridge University Press.

Pearl, J., Glymour, M., and Jewell, N. P. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons.

Pearl, J. and Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic Books.

Petersen, A. H. (2022). causalDisco: Tools for Causal Discovery on Observational Data. R package version 0.9.1.

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramsey, J. and Andrews, B. (2023). Py-Tetrad and RPy-Tetrad: A new Python interface with R support for tetrad causal search. *Proceedings of Machine Learning Research (PMLR), Causal Analysis Workshop Series*, 223, 40-51.

Rubin, D. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association,* 100 (469), 322–331.

Sachs, M. C., Jonzon, G., Sjölander, A., and Gabriel, E. E. (2023). A general method for deriving tight symbolic bounds on causal effects. *Journal of Computational and Graphical Statistics*, 32(2), 567–576.

Scheines, R., Spirtes, P., Glymour, C., Meek, C., and Richardson, T. (1998). The tetrad project: Constraint based aids to causal model specification. *Multivariate Behavioral Research*, 33(1), 65–117.

Scutari, M., Silander, T., and Ness, R. (2024). bnlearn: Bayesian Network Structure Learning, Parameter Learning and Inference. R package version 4.9.4.

Sekhon, J. S. (2010). The Statistics of Causal Inference in the Social Sciences. Political Science 236A and Statistics 239A Lecture Notes, Berkeley Political Science.

Spirtes, P. and Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. *Social science computer review*, 9(1), 62–72.

Spirtes, P., Glymour, C., and Scheines, R. (1993). Causation, Prediction, and Search. Lecture notes in Statistics. Springer.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, prediction, and search*. MIT Press Books, 1.

Spirtes, P., Meek, C., and Richardson, T. (1999). An algorithm for causal inference in the presence of latent variables and selection bias. *Computation, causation, and discovery*, 21, 1–252

Textor, J., Van der Zander, B., Gilthorpe, M., Liśkiewicz, M., and Ellison, G. (2016). Robust causal inference using directed acyclic graphs: the R package dagitty. *International journal of epidemiology*, 45(6), 1887–1894.

Tikka, S., Hyttinen, A., and Karvanen, J. (2021). Causal effect identification from multiple incomplete data sources: A general search-based approach. *Journal of Statistical Software*, 99, 1–40.

Tikka, S. and Karvanen, J. (2017). Identifying causal effects with the R package causaleffect. *Journal of Statistical Software*, 76(12), 1–30.

Wang, X., Jiang, Y., Zhang, N., and Small, D. (2018). Sensitivity analysis and power for instrumental variable studies. *Biometrics*, 74(4), 1150–1160.

BEYANLAR:
Araştırma ve Yayın Etiği Beyanı: Bu çalışma bilimsel araştırma ve yayın etiği kurallarına uygun olarak hazırlanmıştır.
Yazarların Makaleye Katkı Oranları: Birinci yazarın makaleye katkısı %100'dür.
Çıkar Beyanı: Yazarlar açısından ya da üçüncü taraflar açısından çalışmadan kaynaklı çıkar çatışması bulunmamaktadır.
Araştırma Desteği ve Teşekkür: Bu araştırma herhangi bir kurum tarafından desteklenmemiştir.
Etik Kurul Onayı Bilgileri: Makalede açıklanan araştırmada insan denekleri kullanılmadığı için etik kurul onayı alınmamıştır.