The Leadership Quarterly xxx (xxxx) xxx

Contents lists available at ScienceDirect

The Leadership Quarterly

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

# The choice of control variables in empirical management research: How causal diagrams can inform the decision

Paul Hünermund [a,d], Beyers Louw [b], Mikko Rönkkö [c,*]

[a] Copenhagen Business School, Department of Strategy & Innovation, Kilevej 14A DK–2000, Frederiksberg, Denmark
[b] Erasmus University, Rotterdam School of Management, Burgemeester Oudlaan, 3062 PA Rotterdam, the Netherlands
[c] University of Jyväskylä, School of Business and Economics, Mattilanniemi 2 FI–40014, Jyväskylän yliopisto, Finland
[d] Vilnius University, Faculty of Economics and Business Administration, Sauletekio av. LT-10222, Vilnius, Lithuania

ARTICLE INFO

*Keywords:*
Control Variables
Directed Acyclic Graphs
Structural Causal Models
Regression Analysis
Causal Inference

ABSTRACT

The Leadership Quarterly and the management community more broadly prioritize identifying causal relationships to inform effective leadership practices. Despite the availability of more refined causal identification strategies, such as instrumental variables or natural experiments, control variables remain a common strategy in leadership research. The current literature generally agrees that control variables should be chosen based on theory and that these choices should be reported transparently. However, the literature provides little guidance on how specifically potential controls can be identified, how many control variables should be used, and whether a potential control variable should be included. Consequently, the current empirical literature is not fully transparent on how controls are selected and may be contaminated with bad controls that compromise causal inference. Causal diagrams provide a transparent framework to address these issues. This article introduces causal diagrams for leadership and management researchers and presents a workflow for finding an appropriate set of control variables.

## Introduction

The Leadership Quarterly and the management community more broadly prioritize identifying causal relationships to inform effective leadership practices (Antonakis et al., 2010; Bettis et al., 2014). Despite the availability of more refined causal identification strategies (Imbens, 2020), such as instrumental variables (Bastardoz et al., 2023) or natural experiments (Sieweke & Santoni, 2020), control variables remain a common strategy in leadership research (François et al., 2023; Spark et al., 2022; Stefanidis et al., 2022). Control variables are often used when randomization is not feasible and natural experiments are not available or hard to find. As such, "non-experimental designs are sometimes presented as the only feasible way to conduct research in social science" (Narita et al., 2023, p. 1). Thus, whereas there are often better strategies for causal identification—as exemplified by the increase of articles using experiments, natural experiments, and instrumental variables in Leadership Quarterly—control variables remain a fallback option for many researchers.

Control variables are not exclusive to regression but are also a fundamental element of matching strategies (Narita et al., 2023). They can also be used in instrumental variable models (Bastardoz et al., 2023). Because of the wide use of control variables, many articles discussing the selection and reporting of control variables have been published in recent years. These articles recommend that control variables be chosen based on theory (Mändli & Rönkkö, 2023). However, the current literature provides little guidance on such theory-based control variable selection, and consequently, how control variables are selected and used is rarely reported transparently. For example, consider the recent guidelines on propensity score matching by Narita et al. (2023). After emphasizing the importance of choosing the right control variables (covariates), Narita et al. (2023) simply state that "researchers must identify the appropriate covariates based on theoretical and empirical grounds" (p. 3) without providing any guidance on *how specifically this can be done*.

Causal diagrams, also called causal graphs or directed acyclic graphs (DAGs), provide a valuable tool for addressing the lack of rigor and transparency in control variable selection. Causal graphs are graphical representations of causal theory that lay between a narrative representation of a theory and a statistical model that expresses the theory as a set of statistical associations estimated from the data. Unlike algebraic

* Corresponding author.
*E-mail addresses:* phu.si@cbs.dk (P. Hünermund), louw@rsm.nl (B. Louw), mikko.ronkko@jyu.fi (M. Rönkkö).

https://doi.org/10.1016/j.leaqua.2024.101845
Received 27 April 2024; Received in revised form 11 November 2024; Accepted 14 November 2024

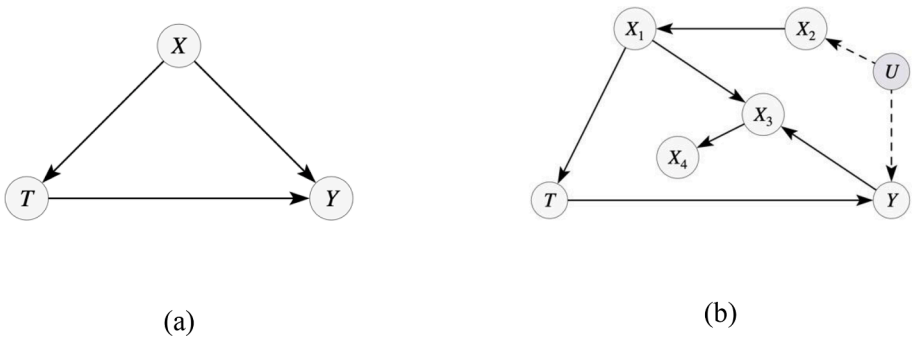



**Fig. 1.** Examples of causal diagrams.

approaches—particularly parametric structural equation models—causal graphs offer the advantage of being fully nonparametric, avoiding assumptions about functional forms or distributions. This flexibility makes them especially useful for model specification in (potentially non-linear) regression models. Graphical representation can sometimes also help in understanding the model and the constraints that it implies. Although the technical literature on causal diagrams has existed for years, management and leadership researchers have mostly ignored these models.

We begin the article with a general explanation of various strategies for causal inference and how control variables and causal diagrams position in the field. After that, we present a taxonomy of controls that sharpens the reader's intuition in analyzing causal diagrams and helps distinguish between good controls that should be used and bad controls that compromise causal inference. We follow this with a more general explanation of the backdoor criterion, the cornerstone of using causal diagrams for control variable selection. Third, we present a workflow that can be used to construct such diagrams and demonstrate this workflow by replicating a study of CEO appearance and firm performance (Hopp et al., 2023). We also discuss sensitivity analysis, which is helpful if some confounding variables in the causal diagram are unobservable.

## Theory on control variables for causal inference

Causal claims are challenging to make because causality itself is unobservable (Hitchcock, 2010). In practice, causality is inferred indirectly through three conditions: 1) association between the assumed cause and effect, 2) direction of influence, and 3) elimination of rival explanations (Antonakis et al., 2010). The third condition can be addressed in two different ways. In experimental and quasi-experimental designs, the cause is either randomized and manipulated by a researcher or occurs in a comparable way in nature, eliminating rival explanations. Even if true experiments are possible in leadership and management more broadly (Bolinger et al., 2022), they can be challenging to implement, and thus, many studies fall back to observational designs. In observational designs, rival explanations are considered by sampling or statistical modeling. These designs can be further divided into conditioning, instrumental variables, and establishing an exclusive mechanism[1] (Durand & Vaara, 2009; Morgan & Winship, 2007, p. 26).

Conditioning is a general term for procedures that eliminate the effect of a third variable by analyzing a variable or a set of variables holding the conditioned variables constant. In the research design context, conditioning is often referred to as controlling, and the variables we condition on are called control variables. Suppose we analyze whether CEO gender affects firm profitability measured by ROA. In this case, industry might cause a spurious effect (e.g., some asset-heavy industries such as mining or heavy industry might be male-dominated). When conditioning on industry, we eliminate the spurious correlation by estimating CEO gender's effect on profitability, mimicking a hypothetical scenario where all companies were in the same industry (Wooldridge, 2013, p. 76). In practice, conditioning can be done by selecting the sample so that the conditioned variable is constant (i.e., study just one industry) or by statistical adjustment (Bollen, 1989, p. 73). Matching and regression are two common strategies for statistical adjustments (Morgan & Winship, 2007, Chapters 4–5). Regardless of how conditioning is done, it assumes we know and can collect data on all possible sources of confounding, which is not always realistic. Thus, other strategies should also be considered; when studying the CEO gender effect, endogenous treatment models provide a compelling alternative (Antonakis et al., 2010, p. 1110; Clougherty et al., 2016). Causal graphs can be helpful with these models (Elwert & Winship, 2014), but these applications are beyond the scope of our article.

To be effective, conditioning requires an appropriate set of control variables. Both including too few and too many controls can lead to wrong conclusions. Econometrics texts typically focus more on the omitted variable case and how it can bias the estimates (Wooldridge, 2013, pp. 88–92), giving less attention to the issue of including too many controls or controlling for variables that should not be controlled for at all (overcontrolling) (Wooldridge, 2013, pp. 205–207). For example, Cameron and Trivedi (2005) state that in the case of including irrelevant variables, "it is straightforward to show that OLS is consistent, but there is a loss of efficiency" (p. 93). Unfortunately, including bad controls can lead to more severe problems than just loss of efficiency, as we explain later in the article.

Another problem in the econometrics literature is that it focuses on the entire model with statements such as "correlation between a single explanatory variable and the error generally results in *all* OLS estimators being biased." (Wooldridge, 2013, pp. 91–92) This claim is incorrect because omitting a relevant variable may bias some estimates but leave others unaffected. Moreover, this all-or-nothing view might make statistical control seem hopeless because the more controls we add, the more variables we have that possibly correlate with the error term. Fortunately, it is possible to consistently estimate the parameters of interest even if some of the variables in the model correlate with the error term (Hünermund & Louw, 2023). This fact is often overlooked in econometric texts.

### *Causal diagrams and causal identification*

Valid causal inferences with control variables require theoretical assumptions about the causal mechanisms that produce the data under study. The *causal diagram* is one helpful framework that allows for

[1] The strategy of establishing an exclusive mechanism is sometimes referred to as the "front-door" approach, and the conditioning strategy as the "back-door" approach (Bellemare et al., 2024; Morgan & Winship, 2007, Section 6.4.3). Despite the increased attention given to the front-door approach in the causal analysis literature, we do not focus on this approach because it is rarely used in leadership and management.

accessible communication of such assumptions (Pearl, 2009, p. 30). Causal diagrams are nonparametric versions of structural equation models. As such, they resemble path diagrams, familiar to leadership scholars. We will next introduce the concept of causal diagrams and then use them to select control variables for regression analyses.

To understand causal diagrams and causal analysis more generally, we need to start with the concept of causal identification. In econometrics, the term identification typically refers to causal identification,[2] which "is the ability to isolate the cause and effect mechanism that creates a relationship between two variables that we study—should one exist" (Shaver, 2020, p. 2). Causal identification is thus distinct from estimation or statistical inference (Morgan & Winship, 2007, Section 1.5) and concerns whether the research design and data can support causal inferences independently of the sample size or how the data are analyzed.

A causal diagram, such as the ones shown in Fig. 1, consists of circles representing the causes and effects (variables, here: $T, Y,$ and $X$; $Y$ commonly refers to the outcome, $T$ to the potential cause or treatment, and $X$ is a covariate) and arrows denoting the causal relationship between them. In the causal analysis literature, the circles are referred to as nodes, and the arrows are referred to as edges, in contrast to the applied SEM literature, where they are called circles and arrows (Kline, 2016, p. 121). Edges are *directed*, indicated by arrowheads pointing from a parent node to a child node (Fig. 1a shows three directed edges: $T\rightarrow Y$, $X\rightarrow Y$ and $X\rightarrow T$). Because causal relationships are asymmetric, causal diagrams are usually assumed to be acyclic, which means there can be no loops in the diagram.[3] These two properties are why causal diagrams are also referred to as *directed acyclic graphs*, or DAGs, in the literature.

A path in a causal diagram is any sequence of edges connecting two nodes. These edges do not need to follow the direction of arrows. If they do, they are called *directed* paths. Fig. 1a contains two directed paths, $T\rightarrow Y$ and $X\rightarrow T\rightarrow Y$, and one undirected one, $T\leftarrow X\rightarrow Y$, between $T$ and $Y$. In the more complex causal diagram shown in Fig. 1b, there are three paths connecting $T$ and $Y$: $T\rightarrow Y$, $T\leftarrow X_1\leftarrow X_2 \dashleftarrow U \dashrightarrow Y$, and $T\leftarrow X_1\rightarrow X_3\leftarrow Y$. The variable $U$ is drawn as a shaded node with dashed edges to indicate that the variable is unobserved and not available in the data.[4] Unobservable variables, although not usable in most analyses, should be included in graphs to highlight potential threats to causal identification.

Causal diagrams are handy for causal inference because they allow us to easily check statistical independence between variables (Kline, 2016, Chapter 8; Pearl, 1988). To illustrate, consider the three simplest possible path configurations of three nodes in a diagram:

1. A *chain*:$A\rightarrow B\rightarrow C$
2. A *fork*:$A\leftarrow B\rightarrow C$
3. And a *collider*:$A\rightarrow B\leftarrow C$

In a chain and fork, the variables $A$ and $C$ are statistically dependent.[5] In a chain, $A$ has an indirect causal effect on $C$; in a fork, both $A$ and $C$ are statistically dependent because the same parent node influences both. In both cases, the dependence can be broken by conditioning on the middle variable $B$. That is, holding $B$ constant renders $A$ and $C$ *conditionally independent* (denoted by: $A\perp C|B$, read as "$A$ is independent of $C$ conditionally on $B$"). In a chain, holding $B$ constant blocks a mechanism and thus, $A$ can no longer influence $C$. For example, if strategic investment affects firm performance exclusively through sales, investments would be uncorrelated with firm performance in a sample where all firms had the same sales figures. Likewise, if the common parent $B$ is fixed in a fork, the remaining variation in $A$ and $C$ is independent. Using the example from earlier in the article, this configuration could be used to model industry ($B$) as a common cause of CEO gender ($A$) and ROA ($C$).

By contrast, the third configuration, a collider, behaves the opposite way, which is essential for understanding bad controls. Here, both $A$ and $C$ are parent nodes of $B$, but otherwise share no relation. That is why $A$ and $C$ can be expected to be independent. However, holding $B$ constant would create a correlation between $A$ and $C$, because fixing the consequence results in constraints on the values that the causes can have. A classic example involves a fictional college (Morgan & Winship, 2007, pp. 66–67), where admission ($B$) depends on SAT score ($A$) and motivation rating based on interviews ($C$). Motivation and SAT scores are independent in the population, but if we sample only admitted students (condition on admission), we find a negative correlation between SAT and motivation scores. That is because if a student in the sample has a low SAT score, their motivation must be high to be admitted and vice versa. In the causal diagram language, we say that conditioning on $B$ *unblocks* or *opens up* the path between $A$ and $C$ and renders them statistically dependent, leading to *collider bias*. This also holds when we condition on variables that are causally affected by the collider.

The property of an intermediate node blocking a path between two variables (and rendering them conditionally independent) is called *d-separation* ("d" for "directional"; Pearl, 1988, p. 117). It also holds in more complex causal diagrams. Returning to our example model in Fig. 1a, the fork path $T\leftarrow X\rightarrow Y$ can be blocked by conditioning on $X$. In Fig. 1b there are three paths connecting $T$ and $Y$: (1) $T\leftarrow X_1\leftarrow X_2 \dashleftarrow U \dashrightarrow Y$, (2) $T\leftarrow X_1\rightarrow X_3\leftarrow Y$, and (3) $T\rightarrow Y$. The first path can be blocked by $X_1$ or $X_2$. The second is already blocked due to the collider node $X_3$, but would become unblocked when conditioning on $X_3$. The third path is the directed, causal path of interest: $T\rightarrow Y$.

Before addressing the control variable selection problem with causal diagrams, we introduce one additional concept to bring some formality to the following discussion. A causal diagram is a parsimonious representation of an underlying *structural causal model* or SCM. For example, Fig. 1a represents the following structural causal model:

$$\begin{aligned} x &\leftarrow f_1(\varepsilon_1) \\ t &\leftarrow f_2(x, \varepsilon_2) \\ y &\leftarrow f_3(t, x, \varepsilon_3) \end{aligned} \tag{1}$$

We follow the common notation in the statistical and causal diagram literature (Pearl, 2009) that uses lowercase letters to denote specific values taken on by random variables, which are themselves denoted by uppercase letters. A structural causal model includes four components: 1) endogenous variables that are determined inside the model ($X$, $T$, and $Y$; or generally $V_i$); 2) exogenous background factors that are associated

[2] In sociological and psychological research on SEMs, "identification" refers to whether data can yield unique model parameter estimates, assuming access to the full population (Bollen, 1989, pp. 88–89; Kline, 2016, p. 119). This concept focuses on the adequacy of the variables for model estimation, independent of sample size and without implying causal relationships. It essentially checks if a model's free parameters can be uniquely determined rather than validating the parameters' causal significance.

[3] This means that the causal diagrams we discuss in this paper cannot be used to express simultaneity and equilibrium conditions (Imbens, 2020), even though extensions to cyclic graphs exist (Hünermund & Bareinboim, 2023). Dynamic feedback loops (needed for systems that are not at equilibrium) can be easily incorporated by introducing time subscripts, $A_t\rightarrow A_{t+1}\rightarrow A_{t+2}$. This is consistent with the idea that causality requires a time delay between the cause and effect and can thus only go in one direction.

[4] In many cases, unobserved factors are not explicitly labeled, as it can be challenging to formulate a theory for potentially numerous unobserved background variables. In such instances, bidirected edges ($\dashleftarrow\dashrightarrow$) are used to represent these unobservables. However, we emphasize that explicitly identifying the nature of omitted variables can lead to more productive discussions about endogeneity.

[5] I.e., $P(A|C) \neq P(A)$

with each variable $(\varepsilon_1, \ldots, \varepsilon_n)$[6]; 3) a set of functions, $(f_{1,}\cdots, f_n)$, that assign values to each endogenous variable taking as arguments the parent nodes and background factors to a variable; and 4) a probability distribution defined over the background factors.

A structural causal model is not a statistical model. It is a tool that can be used to express the causal assumptions on which a statistical model (e.g., structural equation model, regression, etc.) is built on. Structural causal models are purely about theory, and neither the shape of the functions $f_i$ nor the distribution of background factors $\varepsilon_i$ need to be further specified; they only require assumptions about the qualitative causal relationships in the data. In contrast, statistical models usually assume a functional form for $f_i$ and a distribution for $\varepsilon_i$.[7] Linear form and multivariate normal distribution are the most common. With this background knowledge, we can now solve the control variable selection problem with causal diagrams. The online supplement (https://osf.io/q29fb) provides more precise definitions of *causal effect* and *causal identification* (Pearl, 2009).

## A taxonomy of controls

We present a taxonomy of controls to give some intuition of the mechanics of causal diagrams (Cinelli et al., 2022) before providing a general solution to the problem of choosing control variables with graphical causal models. We will first focus on "good" controls that can reduce estimation bias. Afterward, we will discuss "bad" controls, which lead to bias when used. The general principle is that 1) controls should be used to block any non-causal path while at the same time 2) avoiding blocking causal paths or opening up new non-causal ones by controlling a collider. We demonstrate these models using datasets simulated from linear models in R. The analysis code (R and Stata) are provided in the online supplement.

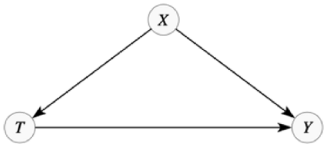
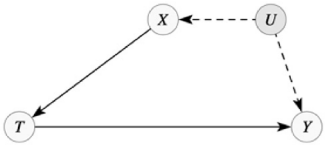
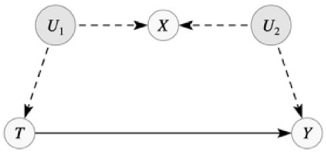
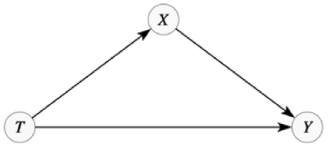
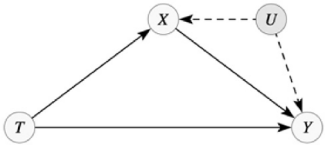
### *Good controls*

Model 1 in Table 1 shows a fully exogenous (since no arrow are pointing into it) control variable $X$. Because $X$ exerts an effect on the treatment $T$ as well as the outcome $Y$ it is a common parent of both creating a fork. Due to the fork structure, $X$ causes a spurious correlation on the non-causal path $T\leftarrow X\rightarrow Y$, and this spurious correlation contaminates the causal effect of $X$ on $Y$. The path can be blocked, however, by conditioning on $X$ to recover the true causal effect. The corresponding linear causal model is:

$$\begin{aligned} x &\leftarrow \beta_{01} + \varepsilon_1 \\ t &\leftarrow \beta_{02} + \beta_{12}x + \varepsilon_2 \\ y &\leftarrow \beta_{03} + \beta_{13}t + \underbrace{\beta_{23}x + \varepsilon_3}_{v} \end{aligned} \tag{2}$$

If $X$ is not included as a control variable in the regression of $Y$ on $T$, the resulting combined error term is equal to $v = \beta_{23}x + \varepsilon_3$, which is correlated with the regressor $T$, as $X$ is a direct causal influence factor of $T$ (eq. (2), second line). Thus, since the assumption of exogeneity is violated, $\mathrm{E}[v|t] \neq 0$, ordinary least squares (OLS) will be biased. To illustrate, we set each component of $\beta = (\beta_{01},\beta_{02},\beta_{12},\beta_{03},\beta_{13},\beta_{23})$ equal to one and sample the random background factors from a standard normal distribution, $\varepsilon_i \sim N(0,1)$, with $n = 1000$. In this setting, if we regress $Y$

**Table 1**
Taxonomy of control variables.

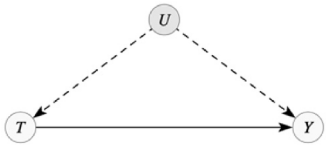
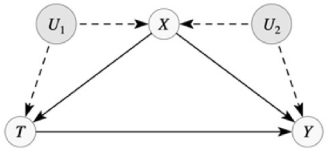
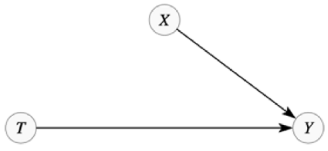
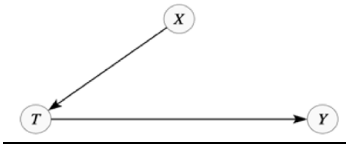
| | | |
|---|---|---|
| **Good Controls** | | |
| 1.) Exogenous |  | • $X$ creates a spurious correlation between treatment and outcome on the path $T\leftarrow X\rightarrow Y$, which can be blocked by controlling for $X$ in a multiple regression<br>• $X$ is not causally affected by any other variable and thus exogenous (standard case assumed in regression theory) |
| 2.) Endogenous |  | • $X$ is not exogenous anymore<br>• $X$ is not a causal determinant of $Y$ anymore, but only correlated with it due to the unobserved confounder $U$<br>• The backdoor path $T\leftarrow X$ <---- $U$ ----> $Y$ leads to a spurious correlation between $T$ and $Y$, which can be blocked by controlling for $X$ |
| **Bad Controls** | | |
| 3.) Collider |  | • $X$ is neither affecting the treatment nor outcome, but correlated with both due to unobserved confounders<br>• There is no spurious correlation between $T$ and $Y$, because the backdoor path $T$ <---- $U_1$ ----> $X$ <---- $U_2$ ----> $Y$ is blocked by the collider $X$<br>• Controlling for $X$ unblocks the path and leads to collider bias |
| 4.) Mediator |  | • $X$ is a post-treatment variable because it is itself causally affected by $T$<br>• The total causal effect of $T$ on $Y$ is identified without any control variables<br>• Controlling for $X$ blocks the path $T\rightarrow X\rightarrow Y$<br>• In linear models with constant treatment effects, this allows to identify the direct effect of $T\rightarrow Y$ and to disentangle causal mechanisms (Imai et al., 2010) |
| 5.) Mediator as collider |  | • $X$ is a collider on the path $T\rightarrow X$ <---- $U$ ----> $Y$ |

(*continued on next page*)

[6] Background factors in the SCM structural causal model framework are similar to "error terms" in SEM. However, this term is avoided to emphasize the causal interpretation of $\varepsilon$, in contrast to a mere statistical approximation error (Pearl, 2009, p. 162).

[7] Modern statistics also provide techniques such as double machine learning (Huntington-Klein, 2022, p. 21.2.4) that support causal inference without requiring any functional form or distributional assumptions. Because we are aware of just one article mentioning these models in management research (Hünermund & Louw, 2023), we do not address them in our article.

**Table 1** (*continued*)

| | |
|---|---|
| | • The total causal effect of $T$ on $Y$ is identified without any control variables<br>• Controlling for $X$ introduces collider bias<br>• Direct effect of $T \rightarrow Y$ is not identifiable and disentangling of causal mechanisms fails |
| **Tricky Cases** | |
| 6.) Unobservable confounder   | • $U$ is an unobservable confounder that is jointly affecting $T$ and $Y$<br>• The backdoor path $T$ <---- $U$ ----> $Y$ cannot be blocked<br>• The causal effect of $T$ on $Y$ is not identifiable via control variables<br>• Other identification strategies that take unobservables into account (e.g., instrumental variables, difference-in-differences, regression discontinuity designs, etc.), might be possible (Antonakis et al., 2010) |
| 7.) Unidentifiable collider   | • $X$ is both a collider on the path $T$ <---- $U_1$ ----> $X$ <---- $U_2$ ----> $Y$ as well as a confounder on the path $T \leftarrow X \rightarrow Y$<br>• Controlling for $X$ reduces confounding but creates collider bias at the same time<br>• There is no solution for this problem, the causal effect of $T$ on $Y$ remains unidentifiable |
| **Unnecessary Controls** | |
| 8.) $X$ only affecting outcome   | • $X$ affects only outcome $Y$<br>• Controlling for $X$ is therefore not necessary in a regression of $Y$ on $T$<br>• However, controlling for $X$ might *reduce* estimation error and result in *higher* precision |
| 9.) $X$ only affecting treatment   | • $X$ affects only treatment $T$<br>• Controlling for $X$ is therefore not necessary in a regression of $Y$ on $T$<br>• Controlling for $X$ might *increase* estimation error and result in *lower* precision |

on $T$ without accounting for $X$, we find a bias of around 50 %. Controlling for $X$, by contrast, gives the correct causal effect (Table 2, col. 1).

The same holds if $X$ is correlated with either $T$ or $Y$, as in model 2 of Table 1. Here there is an unobserved common influence factor $U$ that affects $X$ and $Y$, which otherwise share no causal connection. The path $T \leftarrow X$ <---- $U$ ----> $Y$ is a combination of a fork and a chain and can be blocked by conditioning on $X$. For illustration, we add unobserved $U$ to eq. (2):

$$\begin{aligned} x &\leftarrow \beta_{01} + \beta_{11} u + \varepsilon_1 \\ t &\leftarrow \beta_{02} + \beta_{12} x + \varepsilon_2 \\ y &\leftarrow \beta_{03} + \beta_{13} t + \beta_{23} u + \varepsilon_3 \end{aligned} \tag{3}$$

Again, we set each element of $\beta$ to one. Failing to account for $X$ in the regression of $Y$ on $T$ leads to $\widehat{\beta}_{Y \sim T} = 1.314$, whereas its inclusion gives an estimate that is much closer to the true causal effect (Table 2, col. 2).

### *Bad controls*

Bad controls compromise causal identification either by blocking a causal path or by opening up a new non-causal one. Model 3 in Table 1 shows a graph in which $X$ exerts no causal effect, neither on $T$ nor on $Y$. $X$ is correlated with both, however, due to the presence of two unobserved confounders $U_1$ and $U_2$. Since both unobservables emit arrows that point into $X$, the node is a collider on the path $T$ <---- $U_1$ ----> $X$ <---- $U_2$ ----> $Y$. According to the d-separation criterion, $X$ thus blocks this path, which implies that the relationship between $T$ and $Y$ is currently not confounded. By contrast, if the analyst decided to include $X$ as a control in the regression; this conditioning would open up the path and lead to estimation bias.

Consider:

$$\begin{aligned} x &\leftarrow \beta_{01} + \beta_{11} u_1 + \beta_{12} u_2 + \varepsilon_1 \\ t &\leftarrow \beta_{02} + \beta_{12} u_1 + \varepsilon_2 \\ y &\leftarrow \beta_{03} + \beta_{13} t + \beta_{23} u_2 + \varepsilon_3 \end{aligned} \tag{4}$$

with all coefficients set to one, as before. Here, regressing $Y$ on $T$ gives an unbiased causal effect estimate (Table 2, col. 3). Controlling for $X$, by contrast, introduces estimation bias. Analytically, this means that instead of estimating the correct specification in line 3 of eq. (4), $X$ is added to the structural equation as:

$$y \leftarrow \beta_{03} + \beta_{13} t + \beta_{33} \underbrace{\left(\beta_{01} + \beta_{11} u_1 + \beta_{12} u_2 + \varepsilon_1\right)}_{x} + \underbrace{\beta_{23} u_2 + \varepsilon_3}_{v} \tag{5}$$

Since $U_2$ appears both in the construction of $X$ and the combined error term $v$, $\mathrm{E}[v|x] \neq 0$, this example violates the OLS assumption of no endogeneity (Wooldridge, 2013, p. 86; Assumption MLR.4), and results in biased estimates.

So far, we have assumed $X$ to be a *pre-treatment* variable, which means that it either causally precedes the treatment $T$ or is codetermined with $T$ by other factors. We now turn attention to the mediator case in model 4 of Table 1. Here, $X$ is itself causally affected by $T$, which renders it a *post-treatment* variable. The mediation path $T \rightarrow X \rightarrow Y$ is a chain according to the d-separation criterion. Controlling for $X$ would close the path and block a mechanism by which $T$ influences $Y$. This phenomenon is often described as "controlling away" part of the effect of a treatment. We simulate data from the following mediation model:

$$\begin{aligned} t &\leftarrow \beta_{01} + \varepsilon_1 \\ x &\leftarrow \beta_{02} + \beta_{12} t + \varepsilon_1 \\ y &\leftarrow \beta_{03} + \beta_{13} t + \beta_{23} x + \varepsilon_3 \end{aligned} \tag{6}$$

The total causal effect of $Y$ on $T$ is a combination of a direct effect and an indirect effect that is mediated through $X$. In our simulations, it is equal to two ($\beta_{13} + \beta_{12} \bullet \beta_{23}$) and can be correctly estimated in a simple regression of $Y$ on $T$ (Table 2, col. 4). If we control for $X$, on the other hand, the estimated effect drops to $\widehat{\beta}_{Y \sim T|X} = 1.055$. In this case, $\widehat{\beta}_{Y \sim T|X}$ corresponds to the direct effect of $T$ and not the total causal effect. The difference is important because the path $T \rightarrow Y$ is a parameter of interest only in a mediation analysis that aims to disentangle the different causal mechanisms by which an effect comes about, but it would be an incorrect parameter to estimate if the objective is estimating the total causal effect.

One interesting case where the assumptions of simple mediation

**Table 2**
Regressions related to the taxonomy of controls in Table 1.

| | Scenario | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Good controls | | Bad controls | | | Tricky cases | Unnecessary controls | | Fig. 1 (b) | |
| | Exogenous | Endogenous | Collider | Mediator | Mediator as collider | Unidentifiable collider | X only affecting outcome | X only affecting treatment | Only controlling $X_1$ | Only controlling $X_2$ |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| X is omitted as a control variable | | | | | | | | | | |
| (Intercept) | 1.011 | 0.398 | 0.984 | 1.974 | 1.960 | 0.553 | 1.980 | 1.047 | | |
| | (0.067) | (0.062) | (0.055) | (0.061) | (0.107) | (0.071) | (0.065) | (0.055) | | |
| *T* | 1.499 | 1.314 | 1.006 | 2.071 | 2.066 | 1.705 | 1.022 | 0.991 | | |
| | (0.027) | (0.024) | (0.032) | (0.044) | (0.077) | (0.022) | (0.046) | (0.022) | | |
| X is included as a control variable | | | | | | | | | | |
| (Intercept) | 1.049 | 0.600 | 0.833 | 1.049 | 0.545 | 0.833 | 1.049 | 1.049 | 0.392 | 0.555 |
| | (0.055) | (0.061) | (0.050) | (0.055) | (0.061) | (0.050) | (0.055) | (0.055) | (0.076) | (0.073) |
| *T* | 0.963 | 0.935 | 0.771 | 1.055 | 0.503 | 0.771 | 0.963 | 0.963 | 1.014 | 0.995 |
| | (0.033) | (0.039) | (0.032) | (0.047) | (0.049) | (0.032) | (0.033) | (0.033) | (0.040) | (0.027) |
| *X* | 1.055 | 0.568 | 0.419 | 0.963 | 1.485 | 1.647 | 1.018 | 0.055 | 0.306 | 0.491 |
| | (0.047) | (0.049) | (0.027) | (0.033) | (0.028) | (0.051) | (0.032) | (0.047) | (0.046) | (0.039) |

*Note:* Standard errors in parentheses. "Unobservable confounder" case is omitted because an unobserved variable cannot be controlled for.

analysis are violated is if the mediator $X$ is itself confounded by an unobserved variable $U$, as in model 5 of Table 1. Here, $X$ is a collider on the path $T \rightarrow X \dashleftarrow U \dashrightarrow Y$, which means that controlling for $X$ unblocks the path and leads to a spurious correlation between the treatment and outcome. Let

$$\begin{aligned} t &\leftarrow \beta_{01} + \varepsilon_1 \\ x &\leftarrow \beta_{02} + \beta_{12} t + \beta_{22} u + \varepsilon_1 \\ y &\leftarrow \beta_{03} + \beta_{13} t + \beta_{23} x + \underbrace{\beta_{33} u + \varepsilon_3}_{v} \end{aligned} \tag{7}$$

with all $\beta_i$ equal to one as before. A simple regression of $Y$ on $T$ gives a correct estimate (Table 2, col. 5). The unobserved confounder $U$ thus poses no threat for the causal identification of the total causal effect. However, if we control for $X$, the estimated effect is $\widehat{\beta}_{Y \sim T|X} = 0.503$, which does not correspond to the direct effect of $T$ or any other causal parameter because $X$ is associated with $U$, and thus $\mathrm{E}[v|x] \neq 0$ in the last line of eq. (6). The total effect can likewise not be correctly determined by the path-tracing rule $\widehat{\beta}_{13} + \widehat{\beta}_{12} \bullet \widehat{\beta}_{23}$ (Kline, 2016, pp. 250–252). This example shows that controlling for a mediator is tricky.

### *Tricky cases*

The correct way to deal with bad controls is to leave them out of the analysis. In the bad control cases we described so far, the (total) causal effect was already identified by regressing $Y$ and $T$, and including $X$ in the regression only made things worse (or at least more complicated). Unfortunately, things are not quite so simple for the models we will discuss now, and solutions to the causal identification problem cannot be easily found. The most obvious case is the one depicted in model 6, Table 1, where an important confounder (or a set of confounders) affecting both $T$ and $Y$ is not included in the data set and thus remains unobserved. In such a situation, which frequently arises in applied empirical work, using control variables is not sufficient for causal identification, and the analyst has to resort to other techniques, such as instrumental variables, difference-in-differences, or regression discontinuity designs, which can deal with unobservables (Antonakis et al., 2010).

Much less obvious is the case presented in model 7 of Table 1. Here, $X$ is both a confounder on the path $T \leftarrow X \rightarrow Y$, as well as a collider on the path $T \dashleftarrow U_1 \dashrightarrow X \dashleftarrow U_2 \dashrightarrow Y$. This means that while we would, in principle, like to control for $X$ to close the confounding path, this automatically opens up the second path, which will in turn create collider bias. Unfortunately, there is no way out of this dilemma. When simulating data corresponding to this model, the estimate does not equal the true causal effect (equal to one) regardless of whether we control for $X$ or not (Table 2, col. 6). As in the previous case, instead of control variables, other methods for causal identification would be required.

### *Unnecessary controls*

Finally, we briefly discuss two cases in which accounting for a third variable is not necessary for causal identification but might affect estimation precision. In models 8 and 9 of Table 1, $X$ only influences the treatment or the outcome, respectively. This means that $X$ is not a confounder and does not need to be controlled. However, controlling for $X$ affects the precision of the estimates.[8] When simulating data according to model 8, we find that including $X$ as a control variable increases precision, as exemplified by the smaller standard error (Table 2, col. 7). This happens because $X$ is a part of the error term in a simple regression of $Y$ on $T$ and the precision of the estimates depends on how much unexplained variance remains in the error term (Wooldridge, 2013, pp. 50–54). By contrast, for model 9, we find that standard errors are lower when controlling for $X$ (Table 2, col. 8). This is because $X$ is a parent of $T$ and holding it constant leads to lower estimation precision. This "irrelevant regressor" case is often discussed in introductory econometric texts (Wooldridge, 2013, p. 88).

## The backdoor criterion

The taxonomy shown in Table 1 helps understand the basic principles of causal models and covers many cases that researchers might face. However, it is not exhaustive. The key to general causal identification is to block all spurious paths while keeping open genuine causal paths that transmit an effect of a treatment $T$ on an outcome $Y$. The *backdoor criterion* formalizes this notion[9]:

**Definition (Backdoor Criterion; Pearl, 2009, p. 79)** *Given a treatment variable T and outcome variable Y in a causal diagram, a set of*

[8] X that only affects the outcome affects the scaling of the coefficients of certain generalized linear models, such as logit and probit models. However, the scaling of plots that are virtually always preferred when interpreting the magnitude of the effects in management research are unaffected (Rönkkö et al., 2022). See Endnote 13 in Rönkkö et al (2022) for further explanation of the issue.

[9] The backdoor criterion is not the only way to identify results in a causal graph—it's sufficient but not essential. For more complex cases, advanced identification rules like do-calculus (Pearl, 2009) and methods beyond single-equation regression (Hünermund & Bareinboim, 2023) are required.

*variables Z satisfies the backdoor criterion* (i.e., *Z is backdoor-admissible) relative to* $(T, Y)$ *if no node in Z is a descendant of T, and Z blocks (in the d-separation sense) every path between T and Y, that contains an arrow into T.*

In non-technical terms, we should block all paths with an arrow pointing into $T$ (i.e., they enter "through the backdoor") but no paths that start from $T$ and form a causal chain. When the backdoor criterion holds for the set of variables $Z$, controlling for these variables in multiple regression, propensity score matching, inverse probability weighting, or any other conditioning technique allows to consistently estimate the causal effect of $T$ on $Y$.[10]

Given a causal diagram, finding backdoor-admissible variable sets is straightforward. Start by listing all backdoor paths connecting $T$ and $Y$ and then determine the variables that block them all while avoiding accidentally unblocking collider paths. We demonstrate this using Fig. 1b. $T$ and $Y$ are connected by three paths: the direct (causal) path $T \rightarrow Y$ and two backdoor paths $T \leftarrow X_1 \leftarrow X_2 \dashleftarrow U \dashrightarrow Y$ and $T \leftarrow X_1 \rightarrow X_3 \leftarrow Y$. The latter is blocked by the collider $X_3$ and does not induce a spurious correlation. However, controlling for $X_3$, or its descendant $X_4$, would activate the path, making both variables bad controls. The first backdoor path contains no collider and is, therefore, currently open. This path can be blocked by controlling for $X_1$ or $X_2$, which is why both variables satisfy the backdoor criterion.

The backdoor criterion leads to three important control variable principles. First, two studies sharing the same outcome do not need to use the same set of controls. This is because backdoor-admissible control variable sets are always relative to the pair of variables $(T, Y)$. Second, backdoor-admissible control variable sets need not contain all covariates. In Fig. 1b, it is an error to control for the bad controls $X_3$ and $X_4$; doing so would introduce collider bias. In addition, while both $X_1$ and $X_2$ fulfill the backdoor criterion, it is not necessary to include both in the analysis. Controlling for only one variable is beneficial for statistical precision and allows one to economize on data collection costs (Witte et al., 2020). Third, the estimated coefficients for the control variables in a regression do not always have causal interpretations (Hünermund & Louw, 2023). This is because backdoor-admissible control variable sets do not typically include all causes of $Y$ and can contain endogenous variables (see Model 2, Table 1).

To demonstrate that control variables may not have causal interpretations, we simulate data from a linear model according to Fig. 1b with $n = 1000$, all $\beta_i$ set to one, and the error terms have standard normal distributions.[11] Previously, we demonstrated how the causal effect could be estimated by controlling for either $X_1$ or $X_2$. We run the two regressions corresponding to these choices and obtain the following results:

$$\hat{y} = 0.392 + 1.014 \times t + 0.306 \times x_1 \text{ and} \tag{8}$$

$$\hat{y} = 0.555 + 0.995 \times t + 0.491 \times x_2 \tag{9}$$

In both cases, estimates for $T$ are close to the true causal effect. By contrast, the coefficients on $X_1$ and $X_2$ are positive and significant (Table 2, col. 9 and 10), although neither $X_1$ nor $X_2$ exerts a causal effect on $Y$ in the causal model depicted in Fig. 1b. The example above illustrates that the regression results for control variables in a regression are not very meaningful and should not be interpreted (Hünermund & Louw, 2023). Each variable requires its own careful causal identification argument if the corresponding regression estimates are interpreted causally. This is impractical to do for all controls.

[10] In addition, various statistical techniques come with auxiliary assumptions such as linearity in the case of OLS.

[11] The simulation equations were $X_2 = U + e_{X2}$, $X_1 = X_2 + e_{X1}$, $T = X_1 + e_T$, $Y = T + U + e_Y$, $X_3 = X_1 + Y + e_{X3}$, $X_4 = X_3 + e_{X4}$, where $U$ and all error terms $e$ were independently generated from the standard normal distribution.

**Table 3**
Longlist variables for the link between CEO appearance and performance.

| Variable | Supporting Literature |
|---|---|
| Firm Size (Total Sales) | Rule & Ambady (2011) |
| Firm Size (Total Assets) | Tosi et al. (2004); Waldman et al. (2004); Zhu et al. (2005); Nemanich & Keller (2007); Carmeli et al. (2010); Carmeli et al. (2011); Wang et al. (2011); Devine et al. (2021); Clark et al. (2014); Reina et al. (2014); Adams (2016); Georgakakis et al. (2017); Schepker et al. (2017); Zhang et al. (2017); Bechtoldt et al. (2019); Chiu & Walls (2019); Li & Patel (2019); Yang et al. (2019); Vitanova (2021); Bachrach et al. (2023); Kiss et al. (2022); Ormiston et al. (2022); Weng et al. (2022) |
| Firm Age | Carmeli et al. (2010); Carmeli et al. (2011); Wang et al. (2011); Bechtoldt et al. (2019); Yang et al. (2019); Kiss et al. (2022) |
| Leverage | Clark et al. (2014); Bechtoldt et al. (2019); Kiss et al. (2022); Chiu & Walls (2019) |
| Capital Exp. | Li & Patel (2019) |
| R&D Expense | Nemanich & Keller (2007); Li & Patel (2019); Kiss et al. (2022); |
| Leader Gender | Zhu et al. (2005); Wang, Tsui & Xin (2011); Adams (2016); Georgakakis et al. (2017); Vitanova (2021); Bachrach et al. (2023); Weng et al. (2022); Zhang et al. (2017) |
| Leader Nationality | Georgakakis et al. (2017); Yang et al. (2019) |
| Leader Educational Level | Tsui et al. (2006); Nemanich & Keller (2007); Wang et al. (2011); Reina et al. (2014); Li & Patel (2019); Zhang et al. (2017); Vitanova (2021); Kiss et al. (2022); Bachrach et al. (2023); Stefanidis et al. (2022) |
| Leader Age | Wang et al. (2011); Reina et al. (2014); Li & Patel (2019); Zhang et al. (2017); Kiss et al. (2022); Stefanidis et al. (2022) |
| Leader Tenure | Waldman et al. (2004); Nemanich & Keller (2007); Rule & Ambady (2011); Wang, Tsui & Xin (2011); Georgakakis et al. (2017); Zhang et al. (2017); Yang et al. (2019); Kiss et al. (2022); Vitanova (2021); Weng et al. (2022) |
| Founder Leader | Reina et al. (2014); Zhang et al. (2017); Vitanova (2021) |
| Leader Pay | Li & Patel (2019) |
| Board Size | Georgakakis et al. (2017); Yang et al. (2019); Adams (2016) |
| Lagged ROA | Rowe et al. (2005); Reina et al. (2014); Schepker et al. (2017); Kiss et al. (2022); Ormiston et al. (2022) |
| Growth Options | Wang et al. (2011) |
| Net Income | Rule & Ambady (2011) |
| Cost | Barrick et al. (1991); Chiu & Walls (2019); Vitanova (2021) |
| Asset Intensity | Kiss et al. (2022); |
| Org. Culture | Tsui et al. (2006); Wilderom et al. (2012); Wang et al. (2011) |
| Leader Style | Waldman et al. (2004); Wilderom et al.(2012); Banks et al. (2017) |
| Stakeholder Engagement | Treadway et al. (2009) |
| CEO Reputation | Treadway et al. (2009) |
| CEO Narcissism | Reina et al. (2014); Zhang et al. (2017) |
| Empowerment of TMT | Ling et al. (2015); Adams (2016); Schepker et al. (2017); Chiu & Walls (2019); Weng et al. (2022) |
| State-owned | Ling et al. (2015) |

*Note:* Complete references for the table are available in the online supplement.

## Illustrating the process of choosing control variables

A systematic application of causal diagrams serves three purposes: (1) It guides the choice of control variables, (2) it clarifies which variables should not be controlled (i.e., avoid bad controls), and (3) it makes the reporting of these decisions more transparent. On a high level of abstraction, choosing appropriate control variables consists of constructing a causal diagram and applying our taxonomy of controls or, more generally, applying the backdoor criterion. The first phase is much more challenging because few causal diagrams are available (Huntington-Klein, 2022, Chapter 7). The second phase is more straightforward and involves just the application of causal identification rules based on the diagram, which is a mechanical exercise that is well-explained in multiple literature sources (Pearl, 2009), and that can be automated by a computer algorithm (Textor et al., 2016).

**Table 4**
Shortlist of variables included in the causal diagram (based on Hopp et al., 2023).

| Variable | Theoretical Rationale | Empirical Support |
|---|---|---|
| ROA | Outcome variable. | |
| CEO Attractiveness | Treatment variable. | |
| Firm Size (Total Sales) | Increased sales is one possible mechanism by which CEO attractiveness affects ROA. | Rule & Ambady (2011) |
| Firm Size (Total Assets) | Asset intensity in an industry might affect the choice of CEO. Total assets are used in the calculation of ROA. | Wang et al. (2011); Schepker et al. (2017); Zhang et al. (2017); Chiu & Walls (2019); Devine et al. (2021) |
| Growth Options | Growth options in an industry affect investment cost and sales growth. Industry dynamism might affect types of CEOs chosen. | Zhang et al. (2017) |
| Asset Intensity | Asset intensity is equal to sales over total assets. | |
| Leverage | Highly leveraged industries (e.g. as oil and gas) might opt for different types of CEOs. | Chiu & Walls (2019); Devine et al. (2021) |
| Gender | Perception of CEO attractiveness varies by gender. | Zhang et al. (2017); Devine et al. (2021) |
| Cost | Costs are a plausible mechanism by which CEO attractiveness affects ROA. Costs are also affected by growth options, leverage and total assets. | Barrick et al. (1991); Rule & Ambady (2011); Devine et al. (2021) |
| Net Income | ROA is calculated by dividing net income by total assets, which means that net income is a mediator for the effect of sales on ROA. | |

*Note.* Complete references for the table are available in the online supplement.

We suggest a three-step process for drawing a causal diagram consisting of (1) longlisting variables, (2) selecting the most relevant variables on a shortlist, and (3) specifying the causal relationships between them.[12] We emphasize that drawing a causal diagram is a conceptual exercise separate from data availability; even if a relevant control cannot be measured, it should be documented. This is important when assessing the trustworthiness of the study. It is also helpful for sensitivity analysis and for future research that might be able to measure the respective variable.

To illustrate our approach, we replicate a recently published article by Hopp et al. (2023). The paper replicates a study by Rule and Ambady (2008) on the relationship between CEO appearance and company performance. They reported a positive correlation between individuals' perceptions of personality traits inferred from CEO faces and financial performance in a cross-section of Fortune 1000 firms. Hopp et al. (2023) do a conceptual replication, using return on assets (ROA) instead of profits as the primary outcome variable. They replicate the original cross-sectional results, but their extended analysis shows that the relationship between CEO appearance and performance disappears after adding fixed effects to the regression. In the following, we will use this context (more specifically, models 1–3 in their Table 3) as an example for creating a causal diagram and discuss why some of the previously used controls in this literature might be problematic. The online supplement (https://osf.io/q29fb/) contains R and Stata code for the replication.

### *Longlisting variables based on prior literature*

Developing a causal model starts with a literature review to produce a longlist of variables that will be considered for the causal model. The review should ideally follow established guidelines for systematic reviews (Ferguson et al., 2020; Siddaway et al., 2019) and be done in a study's conceptual phases instead of deferring it until the data are collected. The review should focus on (1) the theorized causal mechanisms and (2) the variables that other researchers have considered when studying the phenomenon empirically. Of particular interest are variables that causally affect either the treatment or the outcome (or both), even if only indirectly, because they are the ones that lie on potential backdoor and mediating paths in the causal diagram.

The list of control variables used in a study does not need to follow prior research and can also include variables not studied before. First, the justification of controls in prior studies may be weak; therefore, some of these controls might be unnecessary or even bad controls (e.g., Becker et al., 2016). Even so, longlisting all prior controls is essential because it allows documenting unnecessary or bad controls so that other researchers know to exclude them. Second, the prior literature may have ignored important controls. For this reason, the variables identified from the literature can and should be complemented by the author's intuition (e.g., if the relevant literature studying a phenomenon is still in its infancy).

Correlation tables should also be inspected—albeit with care. A correlation does not imply causation, and conversely, a lack of correlation does not mean a lack of a causal relationship. Thus, correlations should not be used to directly infer causal relationships or a lack thereof.[13] In particular, weak correlations may not be interpreted as evidence of causal independence because there are many reasons why a study might estimate a correlation incorrectly, such as weak measures and small or biased samples. However, if there is strong evidence that two variables have a non-trivial correlation, it should be accounted for in the causal diagram. According to the d-separation criterion, if two variables are (unconditionally) associated, they must be linked by an open path. This path can either be a directed (and therefore causal) or a backdoor path, in which case a common ancestor node produces the association. Even if such common causes cannot be named, they should be indicated as unobserved nodes in the causal diagram. Thus, constructing a causal diagram forces a researcher to ask why correlations exist.

To provide an example of a longlist, we reviewed strategy and leadership articles that used firm performance as an outcome variable. Our search criteria included the terms "Firm performance", "Return on Assets", or "ROA". We focused on performance rather than CEO appearance because we are unaware of other studies besides Hopp et al. (2023) that would have used this variable. Table 3 contains our longlist, which mainly consists of variables used as control variables in the reviewed studies. All these variables share a plausible connection to CEO appearance via a potential direct causal mechanism (e.g., age) or the selection of better-looking managers into leadership roles (e.g., in some industries).

### *Shortlisting variables for inclusion*

The second step is shortlisting the variables. A shortlist is based on the previous longlist, and here, researchers must decide which variables to include and which ones to exclude from the final causal diagram. While the longlist contains potential influence factors that are either

[12] For an alternative approach relying on domain experts and brainstorming, see Rodrigues et al (2022).

[13] Technically, inferring causal information from associational information requires an assumption of *causal faithfulness*, which rules out pathological cases of two causal mechanisms completely canceling each other out and is frequently invoked in the literature on so-called causal discovery (Peters et al., 2017).

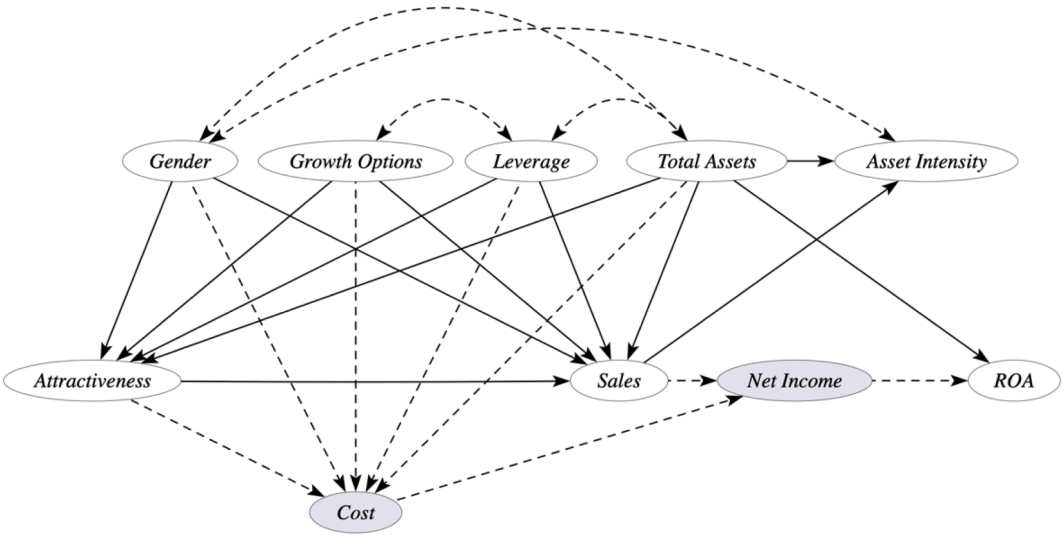



**Fig. 2.** Causal model for the link between CEO appearance and firm performance (based on Hopp et al., 2023).

related to the treatment or the outcome, only variables that affect both at the same time should be shortlisted. The choice should be guided by (1) how compelling the theoretical rationale for a causal relationship is, and (2) how strong the existing empirical evidence in prior literature is. It is unavoidable that the theoretical rationale depends on the researcher's judgment and contextual understanding.

Table 4 shows an example of such a shortlist for the context of CEO appearance. We focus on a CEO's attractiveness as the treatment variable for simplicity since the graph would otherwise become very complex.[14] The relevant confounders for the other dimensions of CEO appearance would be either equivalent or at least very similar. For ease of exposition, we keep the firm fixed effect in Hopp et al.'s (2023) analysis implicit and include only time-variant variables in the causal diagram. Furthermore, we did not include lagged *ROA* for two important reasons. First, Hopp et al. used lagged variables "to control for events that took place in preceding years" (p. 6), which brings us to an essential difference between a statistical model and a causal model. While it is indeed possible to use a lagged dependent variable to control for some unobserved causes in a statistical model (Wooldridge, 2013, p. 313), in this case, lagged *ROA* itself is not assumed to be a cause of current *ROA* because it is not the profits but their determinants that persist over time.[15] Second, including a lagged outcome variable in a fixed effects model leads to dynamic panel bias in traditional panel data regressions (Dishop & DeShon, 2022).[16] Therefore, to avoid dynamic panel bias and to keep our example as simple as possible, we do not include lagged values of *ROA* in the shortlist.

### *Drawing the causal diagram*

The final step is to connect the variables in the causal diagram based on the shortlist. Draw the included variables as circles (nodes) and connect them with arrows (edges) following the assumed causal relationships. It is essential to justify why each edge is included and other possible edges are excluded. Excluding an edge from the diagram is a stronger assumption than including one because it means that the causal effect is precisely zero. In contrast, including an edge means just a possible causal association that does not need to be large.

Fig. 2 shows the causal diagram. Cost and net income are marked as unobserved shaded nodes in the diagram because the data from Hopp et al. (2023) does not include these variables. The graph also includes bidirected edges without labels. These are shorthand, sometimes used to indicate the presence of unobserved common parents. That is, $X_2 \dashleftarrow\dashrightarrow Y$ thus serves as a shortcut notation for $X_2 \dashleftarrow U \dashrightarrow Y$ (with at least one unobservable $U$).

### *Choosing the controls based on the causal diagram*

We will now use Fig. 2 to select control variables. The application of the backdoor criterion leads to two important differences in control choices, as Table 5 shows. Hopp et al. (2023) report that "we controlled for the size of the company during each year using total sales (as the natural logarithm) and the book value of total assets (as the natural logarithm)." (p. 6) Controlling for sales is problematic. *ROA* is defined as $\frac{Netincome}{Totalassets}$ and net income is sales subtracted by all expenses and taxes. This means that increasing sales is one of the mechanisms by which CEO appearance would affect performance. This turns sales into a mediator variable, making it a bad control (case 4 in the taxonomy), as discussed earlier.

Hopp et al. (2023) control for asset intensity in their regressions. While the authors might have theoretical reasons to expect it to be related to the other variables, the article does not explain the reasoning for inclusion. Asset intensity (commonly defined as $\frac{Totalassets}{Sales}$) is affected by sales and total assets by definition.[17] However, we do not see why asset intensity should affect return on assets in the same period. Therefore, a respective edge is missing in Fig. 2. This makes asset intensity an unnecessary control, as it does not lie on a backdoor path connecting CEO appearance and performance (the situation is akin to case 9 in the taxonomy, although asset intensity is causally affected by the treatment).

[14] In complex applications, causal graphs can be drawn and parsed with the help of a computer (Textor et al., 2016).

[15] For a discussion on how lagged dependent variables can be used to identify causal effects, see Section 6.4.1 in Morgan and Winship (2007).

[16] The firm-specific fixed effect exerts an influence on the outcome variable in every time period: $ROA_t, ROA_{t-1}, ROA_{t-2}, \cdots$. At the same time, the fixed effect is part of the error term in a within-variation specification. Lagged outcome variables are thus correlated with the error term, which renders them endogenous.

[17] We note that Hopp et al. (2023) used a more focused definition: "We also included the asset intensity of the firm using the ratio of property, plant, and equipment to sales." (p. 5) We use the more general definition because it simplifies the causal diagram. Using the Hopp et al. definition would not change the conclusion.

**Table 5**
Applying the backdoor criterion rules.

| Variable | Descendant of Attractiveness | Blocked backdoor path | Used as control |
|---|---|---|---|
| Gender | No | Attr. ← Gender → Sales | Yes |
| | | ---→ NI → ROA Attr. ← Gender | |
| | | ---→ Cost | |
| Growth Options | No | ---→ NI → ROA Attr. ← GO → Sales | Yes |
| | | ---→ NI → ROA Attr. ← GO | |
| | | ---→ Cost | |
| Leverage | No | ---→ NI → ROA Attr. ← GO | Yes |
| | | ←-------→ Leverage | |
| | | ---→ Cost | |
| | | ---→ NI → ROA Attr. ← GO | |
| | | ←-------→ Leverage → Sales | |
| | | ---→ NI → ROA Attr. ← Leverage | |
| | | ←-------→ GO → Sales | |
| Total Assets | No | ---→ NI → ROA Attr. ← Gender | Yes |
| | | ←-------→ TA | |
| | | ---→ Cost | |
| | | ---→ NI → ROA Attr. ← Gender | |
| | | ←-------→ TA → ROA Attr. ← Gender | |
| | | ←-------→ TA → Sales | |
| | | ---→ NI → ROA | |
| Asset Intensity | Yes | | No |
| Sales | Yes | | No |

*Note:* Backdoor paths blocked by another control variable are omitted for simplicity. Attr. = Attractiveness, NI = Net Income, TA = Total Assets, GO = Growth Options.

*Replication results*

Table 6 shows the replication results.[18] The first model replicates the original model 3a results from Hopp et al. (2023) exactly. The following three models omit sales, asset intensity, and all sales and asset-related variables. All three models show non-significant effects for attractiveness. In this case, omitting a possible mediator (sales) or an irrelevant control (asset intensity) has no effect because there was no effect to be mediated to start with.

Controlling for a mediator has severe consequences if a causal effect exists in the data. To demonstrate this, we created an alternative version of the Hopp et al. (2023) dataset where attractiveness strongly affected ROA. This effect is mediated through sales, which was already strongly associated with ROA in the original data. We adjusted the sales variable so that it depends on all leadership variables. Then, we used our original model to predict new fitted values for ROA using this new dataset and added the original residuals. Estimating this model produced the same coefficients as the original regression, shown in our model 5.[19] This is the expected outcome because while the causal path *Leadership→Sales→ROA* now exists in the data, it is blocked by controlling sales (controlling for mediator). When we drop sales as a control in model 6, the path opens, and we get a large and highly significant regression coefficient. This example illustrates how controlling for a mediator would have produced a very misleading result if a causal effect had existed in the data.

*Testing the causal model against data*

Causal diagrams can often be tested empirically. Once a causal diagram has been specified, the d-separation relationships implied by the graph may be used to check whether the hypothesized model is compatible with the data (Kline, 2016, Chapter 11). For example, Fig. 1b implies the following d-separation relationships:

$$\begin{aligned} &T \perp X_2|X_1;\ T \perp X_3|X_1, Y;\ T \perp X_4|X_3;\ X_1 \perp X_4|X_3 \\ &X_2 \perp X_3|X_1, Y;\ X_2 \perp X_4|X_3;\ Y \perp X_1|T, X_2;\ Y \perp X_4|X_3. \end{aligned} \tag{10}$$

If any of these conditional independencies does not hold in the data, the hypothesized model should be revised. With linear models and jointly normal error terms, this would involve regressing one of the conditionally independent variables on another, controlling for the variable(s) that the independence depends on. For example, to test $T \perp X_2|X_1$, we would regress $T$ on $X_1$ and $X_2$ and check the statistical significance of the regression coefficient of $X_2$.[20] Compared to global goodness of fit tests, such as the $\chi^2$ statistic from the SEM literature (Kline, 2016, Chapter 12), assessing the model based on conditional independencies has the advantage of providing concrete clues about where the graph is incompatible with the data. Instead of simply rejecting the entire model, the researcher only needs to adjust certain parts of the model locally, which allows for a more targeted approach.

Our causal diagram model for the CEO appearance context in Fig. 2 gives rise to the following conditional independencies:

1. *Attractiveness* ⊥ *Asset Intensity* | *Gender*, *Growth Options*, *Leverage*, *Sales*, *Total Assets*(p = 0.212)
2. *ROA* ⊥ *Asset Intensity* | *Gender*, *Growth Options*, *Leverage*, *Sales*, *Total Assets*(p = 0.810)
3. *Growth Options* ⊥ *Gender*(p = 0.598)
4. *Growth Options* ⊥ *Total Assets*(p = 0.558)
5. *Leverage* ⊥ *Gender*(p = 0.248)

The p-values in parentheses provide the corresponding conditional independency test based on fixed effects regressions using the plm package in R and estimated with cluster-robust standard errors. We find

[18] The R code for producing this table is included in an online supplement.

[19] The $R^2$ is larger because the independent variables correlate more strongly than in the original model.

[20] Such a regression tests uncorrelatedness, not statistical independence. In general, uncorrelatedness does not imply independence. Conditional independence testing for non-Gaussian data is a complex topic that goes beyond the scope of this paper. The interested reader is referred to the literature on constrained-based *causal discovery* methods (e.g., Shah & Peters, 2020)(e.g., Shah and Peters, 2020). It is also possible that a conditional independence holds for just a subset of the data. We refer the reader to the literature on context-specific independence (CSI) for discussion of this scenario (Corander et al., 2019).

**Table 6**
Results of replicating Hopp et al. (2023).

| | Model 3a | Without Sales | Without Asset Intensity | Without anything assets related | Model 3a, simulated | Without Sales, simulated | Model 3d |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *Dependent variable* | **ROA** | **ROA** | **ROA** | **ROA** | **ROA** | **ROA** | **CEO Compensation** |
| Leadership | −0.074 | 0.0003 | −0.070 | −0.039 | −0.074 | 0.941*** | 0.035 |
| | (0.772) | (1.000) | (0.784) | (0.891) | (0.800) | (0.001) | (0.433) |
| Competence | 0.261 | 0.255 | 0.253 | 0.315 | 0.261 | 1.206*** | 0.047 |
| | (0.530) | (0.550) | (0.546) | (0.471) | (0.530) | (0.005) | (0.537) |
| Dominance | −0.188 | −0.216 | −0.187 | −0.155 | −0.188 | 0.739* | −0.019 |
| | (0.599) | (0.575) | (0.601) | (0.687) | (0.627) | (0.052) | (0.716) |
| Likeability | 0.008 | 0.046 | 0.014 | 0.032 | 0.008 | 0.992*** | −0.012 |
| | (0.981) | (0.893) | (0.967) | (0.928) | (0.982) | (0.004) | (0.879) |
| Maturity | 0.192 | 0.180 | 0.188 | 0.233 | 0.192 | 1.132*** | −0.002 |
| | (0.471) | (0.571) | (0.480) | (0.471) | (0.485) | (0.000) | (0.976) |
| Trust | −0.575* | −0.561 | −0.574* | −0.533 | −0.575 | 0.387 | −0.024 |
| | (0.096) | (0.144) | (0.097) | (0.190) | (0.119) | (0.303) | (0.775) |
| Attractiveness | −0.028 | −0.076 | −0.033 | −0.004 | −0.028 | 0.880*** | 0.080** |
| | (0.894) | (0.735) | (0.877) | (0.986) | (0.909) | (0.0001) | (0.036) |
| Affect | 0.096 | 0.039 | 0.095 | 0.027 | 0.096 | 0.997*** | −0.034 |
| | (0.577) | (0.831) | (0.579) | (0.887) | (0.653) | (0.000) | (0.329) |
| Gender | 0.186 | 0.294 | 0.202 | −0.070 | 0.186 | 0.281 | −0.017 |
| | (0.742) | (0.641) | (0.725) | (0.917) | (0.742) | (0.651) | (0.834) |
| Total Sales (log) | 6.192*** | | 6.099*** | | 6.192*** | | 0.025 |
| | (0.000) | | (0.000) | | (0.000) | | (0.827) |
| Total Assets (log) | −5.073*** | −1.260*** | −5.011*** | | −5.073*** | −1.740*** | 0.266** |
| | (0.000) | (0.002) | (0.000) | | (0.000) | (0.000) | (0.023) |
| Growth Options | 1.589** | 1.853** | 1.580** | | 1.589** | 1.820** | 0.003 |
| | (0.049) | (0.041) | (0.048) | | (0.049) | (0.041) | (0.974) |
| Asset Intensity | 0.153 | −2.054*** | | | 0.153 | −1.776*** | −0.097 |
| | (0.774) | (0.002) | | | (0.774) | (0.005) | (0.196) |
| Leverage | −1.994* | −1.922** | −2.015* | | −1.994* | −1.931** | −0.238 |
| | (0.055) | (0.046) | (0.053) | | (0.055) | (0.045) | (0.158) |
| ROA | | | | | | | 0.019*** |
| | | | | | | | (0.006) |
| Firm-fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| P > chi-square | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Overall R2 | 0.213 | 0.135 | 0.214 | 0.010 | 0.304 | 0.373 | 0.341 |
| Within R2 | 0.236 | 0.104 | 0.236 | 0.037 | 0.469 | 0.398 | 0.286 |
| Between R2 | 0.259 | 0.159 | 0.259 | 0.000 | 0.308 | 0.372 | 0.410 |
| Observations | 3,237 | 3,237 | 3,237 | 3,237 | 3,237 | 3,237 | 3,183 |

*Note:* These results in this table are estimated and reported in a similar fashion to columns (3a −3d) in Hopp et al. (2023). We use a fixed-effects regression panel regression with return on assets (column 1–6) and natural log of CEO compensation (column 7) as dependent variable. The first line reports the unstandardized coefficients and the second line reports the p-value estimated with cluster robust standard error. *p < 0.1; **p < 0.05; ***p < 0.01.

that none of these d-separation relationships can be rejected at a 5 % significance level. Therefore, the model is compatible with the data from Hopp et al. (2023).[21]

*Sensitivity analysis*

Sensitivity analysis is applicable when an unobserved confounding cannot be ruled out (as in model 6 of Table 1). We are aware of two general approaches to sensitivity analysis. The first starts with the estimated effect and asks how much confounding there would need to be for the inference on the existence of an effect to be compromised. The second focuses on the point estimate (i.e., regression coefficient) and how much it would change with different levels of confounding (Kiviet, 2020; Kripfganz & Kiviet, 2021). This result is then either summarized by plotting the estimate as a function of confounding or used to adjust confidence intervals to be wider to take the uncertainty due to confounding into account. These techniques require specifying the confounder effects, and are thus useful for testing specific assumptions such as a lack of correlation between error terms in mediation models (Imai et al., 2010) or the exclusion restriction of instrumental variable models (Conley et al., 2010). We focus on the first kind of sensitivity analysis because it is more commonly discussed in this context and also has gained popularity in management research recently (Busenbark et al., 2022). However, regardless of how sensitivity analysis is carried out, it does not absolve researchers from finding identification strategies—such as natural experiments (Jacquart et al., 2024)—ex-ante, as Lonati and Wulff (2024) explicitly point out.

We start our discussion on sensitivity analysis by focusing on the ITCV technique. This technique is inferior to more modern alternatives and we thus recommend against its use. However, we present it first because it may be familiar to readers (Busenbark et al., 2022) and serves as a starting point from which to understand more sophisticated techniques. Consider the regression equation:

$$y = \beta_0 + \beta_1 t + \beta_2 x + \beta_3 z + \varepsilon \tag{11}$$

with outcome $Y$, treatment $T$, error term $\varepsilon$, an observed confounder

[21] Some of the bidirected dashed edges in Fig. 2, indicating unobserved common parents, were chosen to make the causal graph compatible with the data in Hopp et al. (2023).

$X$, and an unobserved confounder $Z$.[22] Frank (2000) defines the *impact* of the unobserved confounder $Z$ as $R_{Y\sim Z|X}R_{T\sim Z|X}$. This corresponds to the product of the two path coefficients on the path $T\leftarrow Z\rightarrow Y$, i.e., the partial correlations of $Z$ with the treatment and outcome. Based on this definition, Frank calculates a threshold (the so-called *impact threshold for a confounding variable*, ITCV) at which the impact of $Z$ would be large enough to make the estimate $\widehat{\beta_1}$ statistically indistinguishable from zero. To compute the impact threshold, only information on the estimated coefficient, its standard error, and the degrees of freedom is necessary, which can be obtained from published studies (Rosenberg et al., 2022).

An important limitation of the ITCV approach is that there is only an indirect relationship between the omitted variable bias in a linear model and the impact of an unobserved confounder as defined by Frank (Cinelli & Hazlett, 2020; Lonati & Wulff, 2024). Oster (2019) provides a formal causal identification result that allows assessing the magnitude of the omitted variable bias under realistic scenarios and benchmark the strength of $Z$ with that of the observed control variables $X$, which is a useful basis for comparison. To do so, she defines the two quantities $R_{\max}$ and $\delta$. The former corresponds to the maximum explanatory power (in terms of $R^2$) that can be attained in the regression in (12). The latter denotes "a value for the relative degree of selection on observed and unobserved variables" (Oster, 2019, p. 188). A formal sensitivity analysis then requires the analyst to specify plausible ranges for these two parameters to assess the potential bias in $\widehat{\beta_1}$.

Cinelli and Hazlett (2020) improve on both previous approaches. They likewise derive a formal causal identification result to bound the effect of an unobserved confounder that is assumed to be '$k$ times' as strong as the observed control variable $X$ (with $k$ being specified by the analyst) on the coefficient estimate $\widehat{\beta_1}$. Compared to Oster's (2019) $\delta$ parameter, which has a rather complicated statistical interpretation (see section 6.3 in Cinelli & Hazlett, 2020 for more details), their approach requires the analyst to reason about the explanatory power that the unobserved confounder $Z$ would add if it were included in a treatment and outcome regression, i.e., ${R^2}_{Y\sim Z|X,T}$ and ${R^2}_{T\sim Z|X}$.[23] This straightforward interpretation makes it easy for applied researchers to bring in substance knowledge from prior research to judge the robustness of their results. An open-source software implementation of their approach is available in R, Python, and Stata (Cinelli et al., 2020), which additionally provides useful graphical tools for sensitivity analysis.

While the replication by Hopp et al. (2023) suggests that there is no effect of CEO appearance on performance, their analysis shows a significant positive relationship between CEO attractiveness and compensation (with a one standard deviation change in mean attractiveness ratings resulting in a 5.1 % change in annual executive compensation, Model 3d). We test the sensitivity of this finding with the procedure proposed by Cinelli and Hazlett using the "sensemakr" package provided in R (Cinelli et al., 2020). Results indicate that an unobserved confounder which explains 5.65 % of the residual variance in the outcome regression (${R^2}_{Y\sim Z|X,T}$) and 5.65 % of the residual variance in the treatment regression (${R^2}_{T\sim Z|X}$) would drive the effect of attractiveness on compensation to zero. To put these numbers into perspective, other regressors can be used as a benchmark. For example, the time-varying control variables used by Hopp et al. (2023) and depicted in Fig. 2 jointly explain 3.4 % of the residual variance in the outcome regression and 7.08 % in the treatment regression. Thus, the unobserved confounder would need to attain an explanatory power of 166 % (= 5.65/3.4) of the rest of the outcome model (excl. fixed effects) to nullify the measured effect size, which makes their result appear quite robust.[24]

## Discussion

Causal diagrams may be helpful, but they are not a silver bullet for making causal claims with observational data. We introduced causal diagrams as a tool to increase rigor in both control choices and reporting practices, highlighting that this tool can be used to determine when a variable should *not* be controlled for, which the previous guidelines on control variables do not address. Yet, we recognize that adopting causal diagrams as a standard practice may face significant barriers because graphs can be difficult to construct. A key challenge in leadership and management is that compared to, e.g., economics, our theories are often weak, tested just once, and relying on research designs that do not allow for drawing robust causal conclusions (Antonakis, 2017). Nevertheless, even if we cannot know for sure whether a causal diagram is correct, their use presents an improvement over the current ad hoc approaches to control variable selection.

Causal diagrams provide a helpful tool for more rigorous and transparent control variable selection and reporting but have been only passingly discussed in journal editorials (e.g., Maula & Stam, 2019). Similarly, articles applying causal diagrams have only recently started to emerge (Frake et al., 2024; Lee & Bettis, 2022). One possible reason for the slow adoption is that most writings on causal diagrams are somewhat technical, lacking links to the kind of applied research management researchers do. However, this does not explain why the uptake of causal graphs has been slow in more technical disciplines like economics. Imbens (2020) notes two likely reasons why economists have not embraced causal graphs so far. First, causal graphs do not add much to the relatively small set of causal identification strategies frequently used by economists, such as regression discontinuity designs or difference-in-differences, which have been mostly developed within the potential outcomes framework. Second, articles advocating for causal graphs have so far failed to present convincing, realistic examples of the methods used, and instead have often adhered to "toy examples" that do not intend to approximate a real-world application. In more realistic settings, models will be necessarily more complex, and specifying all causal relationships between many variables is a difficult task that requires a high degree of domain knowledge about the phenomenon under investigation.

The lack of real-world examples is an issue that will be resolved over time as causal diagrams are taken into use in empirical research. Indeed, our article demonstrates what a realistic application might look like. To facilitate going beyond "toy models", we have proposed a three-step procedure involving (1) longlisting of potential variables, (2) shortlisting them, and (3) specifying the causal relationships between the variables. This process should be guided by existing theory and empirical evidence. However, this requires a sound evidence base to draw on, which some fields might not have yet produced. As such, we see this process as a collective task for an entire research community, as discussed in the next section. Until sufficient high-quality evidence is accumulated, the causal assumptions encoded in causal diagrams—and any causal conclusion drawn using the diagrams—must be regarded as preliminary. Still, even if imperfect, causal diagrams can improve the rigor and transparency of control variable selection over the current ad

[22] For ease of exposition, we will assume $Z$ to be a single omitted variable. Extensions to vector-valued confounders are straightforward and are provided by each of the papers we discuss.

[23] Note that in contrast to Frank (2000), here ${R^2}_{Y\ Z|X,T}$ is also conditional on $T$ (see section 6.2 in Cinelli & Hazlett, 2020).

[24] If statistical significance is used as a criterion, an unobserved confounder would need to account for a smaller portion of the residual variance. The sensemakr package features testing the robustness of t-statistics; however, the version employed in this paper (0.1.4) does not yet support cluster-robust standard errors, which are utilized in Hopp et al. (2023).

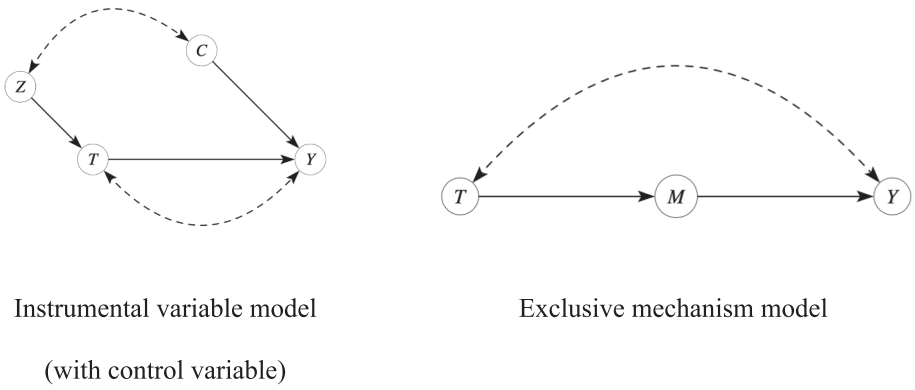



**Fig. 3.** Causal diagrams for instrumental variable model and exclusive mechanism model.

hoc approaches.

Instead of focusing on causal diagrams, Imbens (2020) recommends focusing on a few well-understood identification strategies, often involving quasi-experimental (as-if) randomization, like instrumental variable estimation and regression discontinuity designs, or model-based assumptions in difference-in-differences and synthetic control designs. We fully agree with this recommendation; these strategies should be understood by every researcher who uses observational designs. However, some of these methods still depend on control variables for various justifications. Consequently, even with quasi-experimental designs, selecting valid control variables and avoiding collider bias remains crucial and this is where causal diagrams are helpful.

*Extensions beyond selection on observables designs*

Causal diagrams and the backdoor criterion extend beyond statistical control and observable selection designs, such as regression with control variables or various matching techniques (Narita et al., 2023). For example, causal diagrams can be used to express instrumental variable designs and designs where causal identification is accomplished by establishing an exclusive mechanism. This can be useful in teaching and when explaining the assumptions behind an empirical study because "Compared to the traditional econometrics setup where the critical assumptions are expressed in terms of the correlation between residuals and instruments, […] the DAGs are superior in clarity" (Imbens, 2020, p. 1138).

Since key identifying assumptions are relatively easily recognizable in a causal diagram—once familiar with the methodology—their use can clarify explanations of identification strategies other than control variables. This is especially important when combining multiple identification strategies. Consider, for example, the recent article on assumptions of instrumental variable models (Bastardoz et al., 2023). The article mentions that combining instrumental and control variables is possible, noting that if a control correlates with (shares variance with) an instrument, it must be controlled for. This can be easily seen in the instrumental variable causal diagram in Fig. 3: Unless controlled for, $C$ opens up a backdoor path from $Z$ to $Y$, compromising causal inference.[25] Presenting the models and their assumptions as causal diagrams provides an alternative to equations that might be easier for some readers to understand.

Causal models have also been applied to various other designs to eliminate unobserved confounders. Examples include our replication of Hopp et al. (2023), who used fixed effects to address unobserved confounding in CEO appearance and firm performance studies; Hünermund and Bareinboim (2023), who reviewed $z$-identification, a generalization of instrumental variable estimation using surrogate experiments; Steiner et al. (2017), who discussed regression discontinuity designs in causal diagrams, closely aligning with IV designs near the discontinuity threshold; and Zhang et al. (2021), who developed a framework for incorporating equality constraints in graphical causal models, essential for capturing assumptions like parallel trends in difference-in-differences designs.

*Future prospects of causal diagrams in leadership and management*

A key challenge in applying causal diagrams is that formulating a valid causal model and thoroughly testing it may be too much to ask from one single paper. Instead, we see this as a gradual effort where causal diagrams build on and refine previously presented diagrams. As such, the first diagrams can be simpler and incomplete, and later, more refined versions can be presented. The approach to causal diagrams thus needs to be programmatic. This can occur either on an individual researcher level, where a researcher who specializes in, e.g., CEO appearance builds and develops a causal diagram about this variable, or it can occur on an academic community level, where researchers come together and engage in a process of cumulative theory building, embedded in a larger research program (Shaver, 2020). Initial contributions might be more explorative and involve, e.g., qualitative research designs. In more mature stages, theory testing will then become increasingly relevant.

An important question for causal modeling is how to deal with complexity and scope. The fields of management and leadership show a tendency towards exploration and developing new theories instead of probing and pruning existing causal hypotheses from the literature (Antonakis, 2017; Bettis et al., 2016). This results in a complex theoretical landscape that can be hard to capture adequately in a causal diagram. For example, across all fields of management, there are probably hundreds or thousands of variables that have been used to explain ROA. Going through all these variables and thinking if they might relate to, e.g., CEO appearance or other variables already in the graphs, is an impossible task. Modeling, therefore, requires a fine Occam's razor, which focuses on the main mechanisms that are relevant in a particular context. After all, a model must always remain an abstraction of reality to be useful (Box, 1976). Preference should, therefore, be given to well-established causal relationships and constructs supported by a substantial body of literature over time.

[25] Bastardoz et al. (2023) further note that "If controls are themselves endogenous, they may only be added when they are uncorrelated with the IV z; otherwise, they—against all intentions of IVE—will reintroduce bias into the estimated effect of x on y" (p 3). Why this is the case becomes apparent if one were to draw it as a causal diagram; there are multiple possible ways of doing so. For example, if $C$ and $Y$ in Fig. 3 are affected by an unobserved confounder, and $C$ correlates with $Z$, we have a case of an unidentifiable collider (Case 7 in Table 1). In this case, omitting $C$ leads to omitted variable bias, but controlling for $C$ leads to collider bias, making valid causal inference impossible.

## Conclusion

We introduced causal diagrams and proposed a workflow that can be implemented in leadership and management research. Causal diagrams provide a helpful framework for finding controls that will help make causal claims between the variable of interest and the outcome variable. In contrast to prior guidelines, causal diagrams are also helpful in identifying bad controls that may compromise causal inference. However, causal diagrams are not the be and end-all of causal inference. In the words of Imbens (2020) "Should it be the framework of choice for all causal questions, everywhere, or at least in the social sciences, as [Pearl's The Book of Why] argues? […] In my view the answer […] is no." (p. 1172) Control variables, and more generally conditioning, is not the only possible causal identification strategy. If there is no suitable adjustment set in which all control variables are observed, alternative solutions would be to employ quasi-experimental research designs (Sieweke & Santoni, 2020) or even true experiments (Bolinger et al., 2022). While these strategies come with their own sets of assumptions, they are often preferred over statistical controls due to their ability to address unobserved confounders.

## CRediT authorship contribution statement

**Paul Hünermund:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Beyers Louw:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Mikko Rönkkö:** Writing – review & editing, Methodology, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5–21. https://doi.org/10.1016/j.leaqua.2017.01.006

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Bastardoz, N., Matthews, M. J., Sajons, G. B., Ransom, T., Kelemen, T. K., & Matthews, S. H. (2023). Instrumental variables estimation: Assumptions, pitfalls, and guidelines. *The Leadership Quarterly, 34*(1), Article 101673. https://doi.org/10.1016/j.leaqua.2022.101673

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior, 37*(2), 157–167. https://doi.org/10.1002/job.2053

Bellemare, M. F., Bloem, J. R., & Wexler, N. (2024). The paper of how: Estimating treatment effects using the front-door criterion. *Oxford Bulletin of Economics and Statistics*. https://doi.org/10.1111/obes.12598

Bettis, R. A., Ethiraj, S., Gambardella, A., Helfat, C., & Mitchell, W. (2016). Creating repeatable cumulative knowledge in strategic management. *Strategic Management Journal, 37*(2), 257–261. https://doi.org/10.1002/smj.2477

Bettis, R. A., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal, 35*(7), 949–953. https://doi.org/10.1002/smj.2278

Bolinger, M. T., Josefy, M. A., Stevenson, R., & Hitt, M. A. (2022). Experiments in strategy research: A critical review and future research opportunities. *Journal of Management, 48*(1), 77–113. https://doi.org/10.1177/01492063211044416

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Son Inc.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association, 71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949.

Busenbark, J. R., Yoon, H. (Elle), Gamache, D. L., & Withers, M. C. (2022). Omitted variable bias: Examining management research with the impact threshold of a confounding variable (ITCV). *Journal of Management*, *48*(1), 17–48. https://doi.org/10.1177/01492063211006458.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

Cinelli, C., Ferwerda, J., & Hazlett, C. (2020). *Sensemakr: Sensitivity analysis tools for OLS in R and Stata* (SSRN Scholarly Paper No. 3588978). https://doi.org/10.2139/ssrn.3588978.

Cinelli, C., Forney, A., & Pearl, J. (2022). A crash course in good and bad controls. *Sociological Methods & Research*. https://doi.org/10.1177/00491241221099552

Cinelli, C., & 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. https://doi.org/10.1111/rssb.12348

Clougherty, J. A., Duso, T., & Muck, J. (2016). Correcting for self-selection based endogeneity in management research: Review, recommendations and simulations. *Organizational Research Methods, 19*(2), 286–347. https://doi.org/10.1177/1094428115619013

Conley, T. G., Hansen, C. B., & Rossi, P. E. (2010). Plausibly exogenous. *The Review of Economics and Statistics, 94*(1), 260–272. https://doi.org/10.1162/REST_a_00139

Corander, J., Hyttinen, A., Kontinen, J., Pensar, J., & Väänänen, J. (2019). A logical approach to context-specific independence. *Annals of Pure and Applied Logic, 170*(9), 975–992. https://doi.org/10.1016/j.apal.2019.04.004

Dishop, C. R., & DeShon, R. P. (2022). A tutorial on Bollen and Brand's approach to modeling dynamics while attending to dynamic panel bias. *Psychological Methods, 27*(6), 1089–1107. https://doi.org/10.1037/met0000333

Durand, R., & Vaara, E. (2009). Causation, counterfactuals, and competitive advantage. *Strategic Management Journal, 30*(12), 1245–1264. https://doi.org/10.1002/smj.793

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology, 40*(1), 31–53. https://doi.org/10.1146/annurev-soc-071913-043455

Ferguson, K. D., McCann, M., Katikireddi, S. V., Thomson, H., Green, M. J., Smith, D. J., & Lewsey, J. D. (2020). Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): A novel and systematic method for building directed acyclic graphs. *International Journal of Epidemiology, 49*(1), 322–329. https://doi.org/10.1093/ije/dyz150

Frake, J., Hagemann, A., & Uribe, J. (2024). Collider bias in strategy and management research: An illustration using women CEO's effect on other women's career outcomes. *Strategic Management Journal*. https://doi.org/10.1002/smj.3588

François, A., Panel, S., & Weill, L. (2023). Dictators' facial characteristics and foreign direct investment. *The Leadership Quarterly, 34*(2), Article 101644. https://doi.org/10.1016/j.leaqua.2022.101644

Frank, K. A. (2000). Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research, 29*(2), 147–194. https://doi.org/10.1177/0049124100029002001

Hitchcock, C. (2010). Causation. In S. Psillos & M. Curd (Eds.), *The Routledge companion to philosophy of science* (1. publ. in paperback, pp. 317–326). Routledge.

Hopp, C., Wentzel, D., & Rose, S. (2023). Chief executive officers' appearance predicts company performance, or does it? A replication study and extension focusing on CEO successions. *The Leadership Quarterly, 34*(4), Article 101437. https://doi.org/10.1016/j.leaqua.2020.101437

Hünermund, P., & Bareinboim, E. (2023). Causal inference and data fusion in econometrics. *The Econometrics Journal*, utad008. https://doi.org/10.1093/ectj/utad008.

Hünermund, P., & Louw, B. (2023). On the nuisance of control variables in causal regression analysis. *Organizational Research Methods*. https://doi.org/10.1177/10944281231219274

Huntington-Klein, N. (2022). *The effect: An introduction to research design and causality*. https://theeffectbook.net/index.html.

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science, 25*(1), 51–71. https://doi.org/10.1214/10-STS321

Imbens, G. W. (2020). Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature, 58*(4), 1129–1179.

Jacquart, P., Santoni, S., Schudy, S., Sieweke, J., & Withers, M. (2024). Exogenous shocks: Definitions, types, and causal identification issues. *The Leadership Quarterly, 35*(5), Article 101823. https://doi.org/10.1016/j.leaqua.2024.101823

Kiviet, J. F. (2020). Testing the impossible: Identifying exclusion restrictions. *Journal of Econometrics, 218*(2), 294–316. https://doi.org/10.1016/j.jeconom.2020.04.018

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kripfganz, S., & Kiviet, J. F. (2021). kinkyreg: Instrument-free inference for linear regression models with endogenous regressors. *The Stata Journal, 21*(3), 772–813. https://doi.org/10.1177/1536867X211045575

Lee, G. K., & Bettis, R. (2022). Structural causal modeling of managerial interventions: What if managers had not intervened by doing this? *Strategy Science*, stsc.2022.0169. https://doi.org/10.1287/stsc.2022.0169.

Lonati, S., & Wulff, J. (2024). Hic Sunt Dracones: On the risks of comparing the ITCV with control variable correlations. *Journal of Management.*

Mändli, F., & Rönkkö, M. (2023). To omit or to include? Integrating the frugal and prolific perspectives on control variable use. *Organizational Research Methods*. https://doi.org/10.1177/10944281231221703.

Maula, M., & Stam, W. (2019). Enhancing rigor in quantitative entrepreneurship research. *Entrepreneurship Theory and Practice, 44*(6), 1059–1090. https://doi.org/10.1177/1042258719891388

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge University Press.

Narita, K., Tena, J. D., & Detotto, C. (2023). Causal inference with observational data: A tutorial on propensity score analysis. *The Leadership Quarterly, 34*(3), Article 101678. https://doi.org/10.1016/j.leaqua.2023.101678

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics, 37*(2), 187–204. https://doi.org/10.1080/07350015.2016.1227711

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. Morgan Kaufmann.

Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press.

Rodrigues, D., Kreif, N., Lawrence-Jones, A., Barahona, M., & Mayer, E. (2022). Reflection on modern methods: Constructing directed acyclic graphs (DAGs) with domain experts for health services research. *International Journal of Epidemiology, 51*(4), 1339–1348. https://doi.org/10.1093/ije/dyac135

Rönkkö, M., Aalto, E., Tenhunen, H., & Aguirre-Urreta, M. I. (2022). Eight simple guidelines for improved understanding of transformations and nonlinear effects. *Organizational Research Methods, 25*(1), 48–87. https://doi.org/10.1177/1094428121991907

Rosenberg, J. M., Xu, R., & Frank, K. A. (2022). *Konfound-It!: Quantify the robustness of causal inferences*. http://konfound-it.com.

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science, 19*(2), 109–111. https://doi.org/10.1111/j.1467-9280.2008.02054.x

Shah, R. D., & Peters, J. (2020). The hardness of conditional independence testing and the generalised covariance measure. *The Annals of Statistics, 48*(3), 1514–1538. https://doi.org/10.1214/19-AOS1857

Shaver, J. M. (2020). Causal identification through a cumulative body of research in the study of strategy and organizations. *Journal of Management, 46*(7), 1244–1256. https://doi.org/10.1177/0149206319846272

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology, 70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Sieweke, J., & Santoni, S. (2020). Natural experiments in leadership research: An introduction, review, and guidelines. *The Leadership Quarterly, 31*(1), Article 101338. https://doi.org/10.1016/j.leaqua.2019.101338

Spark, A., O'Connor, P. J., Jimmieson, N. L., & Niessen, C. (2022). Is the transition to formal leadership caused by trait extraversion? A counterfactual hazard analysis using two large panel datasets. *The Leadership Quarterly, 33*(2), Article 101565. https://doi.org/10.1016/j.leaqua.2021.101565

Stefanidis, D., Nicolaou, N., Charitonos, S. P., Pallis, G., & Dikaiakos, M. (2022). What's in a face? Facial appearance associated with emergence but not success in entrepreneurship. *The Leadership Quarterly, 33*(2), Article 101597. https://doi.org/10.1016/j.leaqua.2021.101597

Steiner, P. M., Kim, Y., Hall, C. E., & Su, D. (2017). Graphical models for quasi-experimental designs. *Sociological Methods & Research, 46*(2), 155–188. https://doi.org/10.1177/0049124115582272

Textor, J., van der Zander, B., Gilthorpe, M. S., Li{\acute{s}}kiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: The R package 'dagitty'. *International Journal of Epidemiology, 45*(6), 1887–1894. https://doi.org/10.1093/ije/dyw341

Witte, J., Henckel, L., Maathuis, M. H., & Didelez, V. (2020). On efficient adjustment in causal graphs. *The Journal of Machine Learnign Research, 21*(1), 9956–10000.

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach* (5th ed). South-Western Cengage Learning.

Zhang, C., Cinelli, C., Chen, B., & Pearl, J. (2021). Exploiting equality constraints in causal inference. *International Conference on Artificial Intelligence and Statistics*, 1630–1638. http://proceedings.mlr.press/v130/zhang21c.html.