24. From DAGs to linear path models

Ian Lundberg
Cornell Info 6751: Causal Inference in Observational Settings
Fall 2022

15 Nov 2022

Learning goals for today

At the end of class, you will be able to:

1. Recognize differences between DAGs and linear path models
2. Determine the covariance between standardized variables using a linear path model
3. Quantify the biases of coefficient estimators using path models
4. Be equipped to read Elwert & Pfeffer (2022)

We have been working with DAGs

$$\begin{tikzcd}
A \arrow[r] & Y \\
L \arrow[u] \arrow[ur] &
\end{tikzcd}$$

- Each node is a variable
- Each edge is a direct causal effect
- The DAG is nonparametric
  - Effect of A may be nonlinear
  - Effect of A may be heterogeneous across L

$$\mathrm{E}(Y \mid A, L) = f(A, L) \qquad \leftarrow \text{arbitrarily complex } f()$$

Today we will use linear path models

Today we will use linear path models

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Today we will use linear path models

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}$$

Adds a **parametric assumption:**
Each output is a linear, additive function of its inputs

Today we will use linear path models

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

Adds a **parametric assumption:**
Each output is a linear, additive function of its inputs

$$\mathrm{E}(Y \mid A, L) = \alpha + bA + cL$$

Today we will use linear path models

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}$$

Adds a **parametric assumption:**
Each output is a linear, additive function of its inputs

$$E(Y \mid A, L) = \alpha + bA + cL$$

- Homogeneous causal effects

Today we will use linear path models

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Adds a **parametric assumption:**
Each output is a linear, additive function of its inputs

$$\mathrm{E}(Y \mid A, L) = \alpha + bA + cL$$

- Homogeneous causal effects
- Linear causal effects

Today we will use linear path models

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

Adds a **parametric assumption:**
Each output is a linear, additive function of its inputs

$$\mathrm{E}(Y \mid A, L) = \alpha + bA + cL$$

- Homogeneous causal effects
- Linear causal effects

We will also assume all variables are **standardized**
to mean 0 and variance 1

CORRELATION AND CAUSATION

By SEWALL WRIGHT

Senior Animal Husbandman in Animal Genetics, Bureau of Animal Industry, United States Department of Agriculture

PART I. METHOD OF PATH COEFFICIENTS

INTRODUCTION

The ideal method of science is the study of the direct influence of one condition on another in experiments in which all other possible causes of variation are eliminated. Unfortunately, causes of variation often seem to be beyond control. In the biological sciences, especially, one often has to deal with a group of characteristics or conditions which are correlated because of a complex of interacting, uncontrollable, and often obscure causes. The degree of correlation between two variables can be calculated by well-known methods, but when it is found it gives merely the resultant of all connecting paths of influence.

The present paper is an attempt to present a method of measuring the direct influence along each separate path in such a system and thus of finding the degree to which variation of a given effect is determined by each particular cause. The method depends on the combination of knowledge of the degrees of correlation among the variables in a system with such knowledge as may be possessed of the causal relations. In cases in which the causal relations are uncertain the method can be used to find the logical consequences of any particular hypothesis in regard to them.

CORRELATION

Relations between variables which can be measured quantitatively are usually expressed in terms of Galton's (*4*)[1] coefficient of correlation, $r_{XY}=\frac{\Sigma X'Y'}{n\sigma_X\sigma_Y}$ (the ratio of the average product of deviations of X and Y to the product of their standard deviations), or of Pearson's (*7*) correlation ratio, $\eta_{X\cdot Y}=\frac{\sigma\left({}_{Y}{}^{M}{}_{X}\right)}{\sigma_X}$ (the ratio of the standard deviation of the mean values of X for each value of Y to the total standard deviation of X), the standard deviation being the square root of the mean square deviation.

Use of the coefficient of correlation (r) assumes that there is a linear relation between the two variables—that is, that a given change in one variable always involves a certain constant change in the corresponding average value of the other. The value of the coefficient can never exceed

[1] Reference is made by number (italic) to "Literature cited," p. 585.

Journal of Agricultural Research, Washington, D. C.
wh
17777°—21——4

Vol. XX, No. 7
Jan. 3, 1921
Key No. A-55

Wright's (1921) path rule:
Connecting causal paths to statistical associations

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

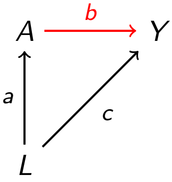
$$\mathrm{Cov}(A, \tilde{Y}) =$$

Wright's (1921) path rule:
Connecting causal paths to statistical associations

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

$$\text{Cov}(A, \tilde{Y}) =$$

Wright's (1921) path rule: Connecting causal paths to statistical associations



$$\mathrm{Cov}(A, \tilde{Y}) = b$$

Wright's (1921) path rule: Connecting causal paths to statistical associations

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

$$\mathrm{Cov}(A, \tilde{Y}) = b$$

Wright's (1921) path rule:
Connecting causal paths to statistical associations

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}$$

$$\mathrm{Cov}(A, \tilde{Y}) = b + ac$$

Wright's (1921) path rule: Connecting causal paths to statistical associations

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

$$\text{Cov}(A, \tilde{Y}) = b + ac$$

Wright's rule:
When all variables are standardized to variance 1,
the covariance between two variables
is the sum over unblocked paths
of the product of coefficients on that path

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths
of the product of coefficients on each path

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\mathrm{Cov}(Z, A) = ?$$
$$\mathrm{Cov}(Z, Y) = ?$$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths
of the product of coefficients on each path

$$
\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}
$$

$$\mathrm{Cov}(Z, A) = a$$
$$\mathrm{Cov}(Z, Y) = ?$$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths of the product of coefficients on each path

$$\begin{tikzcd} Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\ & U \arrow[u, "c"] \arrow[ur, "d"'] & \end{tikzcd}$$

$$\mathrm{Cov}(Z, A) = a$$
$$\mathrm{Cov}(Z, Y) = ab$$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths
of the product of coefficients on each path

$$
\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}
$$

$$\mathrm{Cov}(Z, A) = a$$
$$\mathrm{Cov}(Z, Y) = ab$$

Note. $\frac{\mathrm{Cov}(Z,Y)}{\mathrm{Cov}(Z,A)}$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths of the product of coefficients on each path

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\mathrm{Cov}(Z, A) = a$$
$$\mathrm{Cov}(Z, Y) = ab$$

Note. $\frac{\mathrm{Cov}(Z,Y)}{\mathrm{Cov}(Z,A)} = \frac{ab}{a}$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths of the product of coefficients on each path

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\text{Cov}(Z, A) = a$$
$$\text{Cov}(Z, Y) = ab$$

Note. $\frac{\text{Cov}(Z,Y)}{\text{Cov}(Z,A)} = \frac{ab}{a} = b$

Quick practice

Wright's rule:
The covariance is the sum over unblocked paths of the product of coefficients on each path

$$\begin{tikzcd} Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\ & U \arrow[u, "c"] \arrow[ur, "d"'] & \end{tikzcd}$$

$$\mathrm{Cov}(Z, A) = a$$
$$\mathrm{Cov}(Z, Y) = ab$$

Note. $\frac{\mathrm{Cov}(Z,Y)}{\mathrm{Cov}(Z,A)} = \frac{ab}{a} = b$

Stay tuned...future class!

When is Wright's rule helpful? Reasoning about biases

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Suppose we don't observe U.

When is Wright's rule helpful? Reasoning about biases

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Suppose we don't observe U. We regress Y on A.

$$\mathsf{E}(Y \mid A) = \alpha + \beta A$$

When is Wright's rule helpful? Reasoning about biases

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Suppose we don't observe U. We regress Y on A.

$$\mathrm{E}(Y \mid A) = \alpha + \beta A$$

What do we get?

When is Wright's rule helpful? Reasoning about biases

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Suppose we don't observe U. We regress Y on A.

$$E(Y \mid A) = \alpha + \beta A$$

What do we get?

$$\beta = \frac{\text{Cov}(A, Y)}{\text{V}(Y)} \qquad \text{from regression}$$

When is Wright's rule helpful? Reasoning about biases

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

Suppose we don't observe U. We regress Y on A.

$$\mathsf{E}(Y \mid A) = \alpha + \beta A$$

What do we get?

$$\begin{aligned} \beta &= \frac{\mathrm{Cov}(A, Y)}{\mathrm{V}(Y)} && \text{from regression} \\ &= \mathrm{Cov}(A, Y) && \text{since standardized} \end{aligned}$$

When is Wright's rule helpful? Reasoning about biases

$$
\begin{tikzcd}
A \arrow[r, "b"] & Y \\
U \arrow[u, "a"] \arrow[ur, "c"'] &
\end{tikzcd}
$$

Suppose we don't observe U. We regress Y on A.

$$\mathrm{E}(Y \mid A) = \alpha + \beta A$$

What do we get?

$$
\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y)}{\mathrm{V}(Y)} && \text{from regression} \\
&= \mathrm{Cov}(A, Y) && \text{since standardized} \\
&= \underbrace{b}_{\text{Estimand}} + \underbrace{ac}_{\text{Bias}} && \text{Wright's rule}
\end{aligned}
$$

When is Wright's rule helpful? Reasoning about biases

Discuss.
Think about that bias in a substantive example.

Conditional estimators meet Wright's rule

$$\mathrm{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"] & \end{tikzcd}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\beta = \frac{\mathsf{Cov}(A, Y) - \mathsf{Cov}(L, Y)\mathsf{Cov}(L, A)}{1 - [\mathsf{Cov}(L, A)]^2}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] &
\end{tikzcd}$$

Conditional estimators meet Wright's rule

$$\mathrm{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] &
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(L, Y)\mathrm{Cov}(L, A)}{1 - [\mathrm{Cov}(L, A)]^2} \\
&= \frac{(b + ac) - ((c + ab)a)}{1 - a^2}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] &
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(L, Y)\mathrm{Cov}(L, A)}{1 - [\mathrm{Cov}(L, A)]^2} \\
&= \frac{\big(b + ac\big) - \big((c + ab)a\big)}{1 - a^2} \\
&= \frac{b + ac - ac - a^2 b}{1 - a^2}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] &
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(L, Y)\mathrm{Cov}(L, A)}{1 - [\mathrm{Cov}(L, A)]^2} \\
&= \frac{(b + ac) - ((c + ab)a)}{1 - a^2} \\
&= \frac{b + ac - ac - a^2 b}{1 - a^2} \\
&= \frac{b(1 - a^2)}{1 - a^2}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"] & \end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(L, Y)\mathrm{Cov}(L, A)}{1 - [\mathrm{Cov}(L, A)]^2} \\
&= \frac{(b + ac) - ((c + ab)a)}{1 - a^2} \\
&= \frac{b + ac - ac - a^2 b}{1 - a^2} \\
&= \frac{b(1 - a^2)}{1 - a^2} \\
&= b
\end{aligned}$$

Conditional estimators meet Wright's rule

$$E(Y \mid A, L) = \alpha + \beta A + \gamma L$$

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"] & \end{tikzcd}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L}
\end{tikzcd}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\beta = \frac{\mathsf{Cov}(A, Y) - \mathsf{Cov}(\tilde{L}, Y)\mathsf{Cov}(\tilde{L}, A)}{1 - \left[\mathsf{Cov}(\tilde{L}, A)\right]^2}$$

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L} \end{tikzcd}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L} \end{tikzcd}$$

$$\beta = \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(\tilde{L}, Y)\mathrm{Cov}(\tilde{L}, A)}{1 - \left[\mathrm{Cov}(\tilde{L}, A)\right]^2}$$

$$= \frac{(b + ac) - ((cd + abd)ad)}{1 - a^2 d^2}$$

Conditional estimators meet Wright's rule

$$\mathrm{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L}
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(\tilde{L}, Y)\mathrm{Cov}(\tilde{L}, A)}{1 - \left[\mathrm{Cov}(\tilde{L}, A)\right]^2} \\
&= \frac{\left(b + ac\right) - \left((cd + abd)ad\right)}{1 - a^2 d^2} \\
&= \frac{b + ac - acd^2 - a^2 bd^2}{1 - a^2 d^2}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L}
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(\tilde{L}, Y)\mathrm{Cov}(\tilde{L}, A)}{1 - \left[\mathrm{Cov}(\tilde{L}, A)\right]^2} \\
&= \frac{(b + ac) - ((cd + abd)ad)}{1 - a^2 d^2} \\
&= \frac{b + ac - acd^2 - a^2 bd^2}{1 - a^2 d^2} \\
&= \frac{b(1 - a^2 d^2) + ac(1 - d^2)}{1 - a^2 d^2}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathsf{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L}
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(\tilde{L}, Y)\mathrm{Cov}(\tilde{L}, A)}{1 - \left[\mathrm{Cov}(\tilde{L}, A)\right]^2} \\
&= \frac{\left(b + ac\right) - \left((cd + abd)ad\right)}{1 - a^2 d^2} \\
&= \frac{b + ac - acd^2 - a^2 bd^2}{1 - a^2 d^2} \\
&= \frac{b(1 - a^2 d^2) + ac(1 - d^2)}{1 - a^2 d^2} \\
&= b + \underbrace{ac}_{\substack{\text{Bias} \\ \text{without} \\ \text{control}}} \underbrace{\frac{1 - d^2}{1 - a^2 d^2}}_{\substack{\text{Bias} \\ \text{Multiplier} \\ <1}}
\end{aligned}$$

Conditional estimators meet Wright's rule

$$\mathrm{E}(Y \mid A, \tilde{L}) = \alpha + \beta A + \gamma \tilde{L}$$

$$\begin{tikzcd}
A \arrow[r, "b"] & Y \\
L \arrow[u, "a"] \arrow[ur, "c"] \arrow[r, "d"'] & \tilde{L}
\end{tikzcd}$$

$$\begin{aligned}
\beta &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(\tilde{L}, Y)\mathrm{Cov}(\tilde{L}, A)}{1 - \left[\mathrm{Cov}(\tilde{L}, A)\right]^2} \\
&= \frac{(b + ac) - ((cd + abd)ad)}{1 - a^2d^2} \\
&= \frac{b + ac - acd^2 - a^2bd^2}{1 - a^2d^2} \\
&= \frac{b(1 - a^2d^2) + ac(1 - d^2)}{1 - a^2d^2} \\
&= b + \underbrace{ac}_{\substack{\text{Bias}\\\text{without}\\\text{control}}} \underbrace{\frac{1 - d^2}{1 - a^2d^2}}_{\substack{\text{Bias}\\\text{Multiplier}\\<1}}
\end{aligned}$$

Result.
If you don't measure L, then controlling for $\tilde{L}$ is strictly better than controlling for nothing

Elwert & Pfeffer 2022

Article

The Future Strikes Back: Using Future Treatments to Detect and Reduce Hidden Bias

Sociological Methods & Research
1-38

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0049124119875958
journals.sagepub.com/home/smr
SAGE

Felix Elwert[1] and Fabian T. Pfeffer[2]

Abstract

Conventional advice discourages controlling for postoutcome variables in regression analysis. By contrast, we show that controlling for commonly available postoutcome (i.e., future) values of the treatment variable can help detect, reduce, and even remove omitted variable bias (unobserved confounding). The premise is that the same unobserved confounder that affects treatment also affects the future value of the treatment. Future treatments thus proxy for the unmeasured confounder, and researchers can exploit these proxy measures productively. We establish several new results: Regarding a commonly assumed data-generating process involving future treatments, we (1) introduce a simple new approach and show that it strictly reduces bias, (2) elaborate on existing approaches and show that they can increase bias, (3) assess the relative merits of alternative approaches, and (4) analyze true state dependence and selection as key challenges. (5) Importantly, we also introduce a new nonparametric test that uses future treatments to detect hidden bias even when future-treatment estimation fails to reduce bias. We illustrate these results

[1] University of Wisconsin–Madison, WI, USA
[2] University of Michigan, Ann Arbor, MI, USA

Corresponding Author:
Felix Elwert, University of Wisconsin-Madison, Department of Sociology,1180 Observatory Drive, Madison, WI 53706, USA.
Email: elwert@wisc.edu

Elwert & Pfeffer 2022

Check for updates

Article

Sociological Methods & Research
1-38

DOI: 10.1177/0049124119875958
journals.sagepub.com/home/smr
SAGE

The Future Strikes Back: Using Future Treatments to Detect and Reduce Hidden Bias

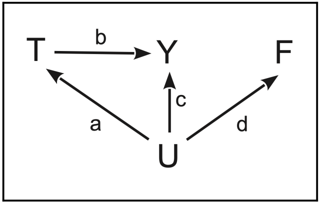
Felix Elwert[1] and Fabian T. Pfeffer[2]

Abstract

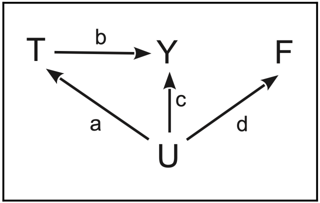
Conventional advice discourages controlling for postoutcome variables in regression analysis. By contrast, we show that controlling for commonly available postoutcome (i.e., future) values of the treatment variable can help detect, reduce, and even remove omitted variable bias (unobserved confounding). The premise is that the same unobserved confounder that affects treatment also affects the future value of the treatment. Future treatments thus proxy for the unmeasured confounder, and researchers can exploit these proxy measures productively. We establish several new results: Regarding a commonly assumed data-generating process involving future treatments, we (1) introduce a simple new approach and show that it strictly reduces bias, (2) elaborate on existing approaches and show that they can increase bias, (3) assess the relative merits of alternative approaches, and (4) analyze true state dependence and selection as key challenges. (5) Importantly, we also introduce a new nonparametric test that uses future treatments to detect hidden bias even when future-treatment estimation fails to reduce bias. We illustrate these results

[1] University of Wisconsin–Madison, WI, USA
[2] University of Michigan, Ann Arbor, MI, USA

Corresponding Author:
Felix Elwert, University of Wisconsin-Madison, Department of Sociology,1180 Observatory Drive, Madison, WI 53706, USA.
Email: elwert@wisc.edu



Elwert & Pfeffer 2022

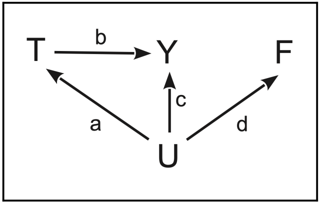


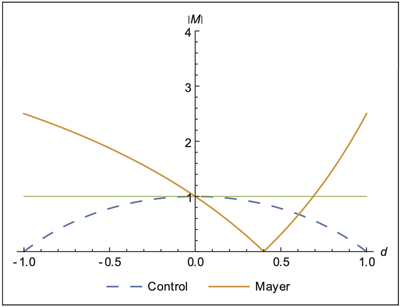
$$E(Y \mid T, F) = \alpha + \beta T + \gamma F$$

Two estimators:

1. Control estimator: β
2. Difference estimator: $\beta - \gamma$

Elwert & Pfeffer 2022



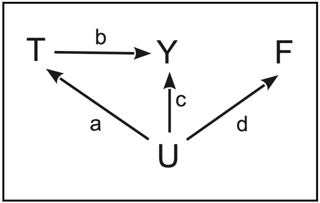


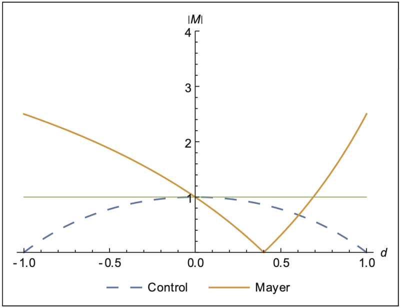
$$E(Y \mid T, F) = \alpha + \beta T + \gamma F$$

Two estimators:

1. Control estimator: β
2. Difference estimator: $\beta - \gamma$

Elwert & Pfeffer 2022





Reading before Thursday

Elwert, F., & Pfeffer, F. T. (2022). The future strikes back: using future treatments to detect and reduce hidden bias. Sociological Methods & Research, 51(3), 1014-1051.

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\beta_{\text{Unadjusted}} = \text{Cov}(A, Y)$$

$$\beta_{\text{Adjusted}}$$

More practice with path diagrams

$$\begin{tikzcd} Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\ & U \arrow[u, "c"] \arrow[ur, "d"'] & \end{tikzcd}$$

$\beta_{\text{Unadjusted}} = \text{Cov}(A, Y)$

β_{Adjusted}

What is $\text{Cov}(A, Y)$?

What is $\text{Cov}(Z, A)$?

What is $\text{Cov}(Z, Y)$?

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$\beta_{\text{Unadjusted}} = \text{Cov}(A, Y)$

β_{Adjusted}

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?

What is $\text{Cov}(Z, Y)$?

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$\beta_{\text{Unadjusted}} = \text{Cov}(A, Y)$

β_{Adjusted}

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?

More practice with path diagrams

$$\beta_{\text{Unadjusted}} = \text{Cov}(A, Y)$$

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\beta_{\text{Adjusted}}$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
 & U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}$$

$$\beta_{\text{Adjusted}}$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}$$

β_{Adjusted}

$$= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2}$$

What is $\text{Cov}(A, Y)$?

$b + cd$

What is $\text{Cov}(Z, A)$?

a

What is $\text{Cov}(Z, Y)$?

ab

More practice with path diagrams

$$\begin{tikzcd} Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\ & U \arrow[u, "c"] \arrow[ur, "d"'] & \end{tikzcd}$$

What is $\mathrm{Cov}(A, Y)$?
$b + cd$

What is $\mathrm{Cov}(Z, A)$?
a

What is $\mathrm{Cov}(Z, Y)$?
ab

$$\begin{aligned} \beta_{\text{Unadjusted}} &= \mathrm{Cov}(A, Y) \\ &= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}} \end{aligned}$$

$$\begin{aligned} &\beta_{\text{Adjusted}} \\ &= \frac{\mathrm{Cov}(A, Y) - \mathrm{Cov}(Z, Y)\mathrm{Cov}(Z, A)}{1 - [\mathrm{Cov}(Z, A)]^2} \\ &= \frac{(b + cd) - (ab)(a)}{1 - a^2} \end{aligned}$$

More practice with path diagrams

$$
\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}
$$

What is $\text{Cov}(A, Y)$?

$b + cd$

What is $\text{Cov}(Z, A)$?

a

What is $\text{Cov}(Z, Y)$?

ab

$$
\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}
$$

$$
\begin{aligned}
&\beta_{\text{Adjusted}} \\
&= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2} \\
&= \frac{(b + cd) - (ab)(a)}{1 - a^2} \\
&= \frac{b(1 - a^2) + cd}{1 - a^2}
\end{aligned}
$$

More practice with path diagrams

$$
\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}
$$

$$
\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}
$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

$$
\begin{aligned}
&\beta_{\text{Adjusted}} \\
&= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2} \\
&= \frac{(b + cd) - (ab)(a)}{1 - a^2} \\
&= \frac{b(1 - a^2) + cd}{1 - a^2} \\
&= \underbrace{b} + \underbrace{cd}\,\underbrace{\frac{1}{1 - a^2}}
\end{aligned}
$$

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

$$\begin{aligned}\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\ &= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}\end{aligned}$$

$$\begin{aligned}&\beta_{\text{Adjusted}} \\ &= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2} \\ &= \frac{(b + cd) - (ab)(a)}{1 - a^2} \\ &= \frac{b(1 - a^2) + cd}{1 - a^2} \\ &= \underbrace{b}_{\text{Effect}} + \underbrace{cd}\,\underbrace{\frac{1}{1 - a^2}}\end{aligned}$$

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

$$\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}$$

$$\begin{aligned}
&\beta_{\text{Adjusted}} \\
&= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2} \\
&= \frac{(b + cd) - (ab)(a)}{1 - a^2} \\
&= \frac{b(1 - a^2) + cd}{1 - a^2} \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}} \underbrace{\frac{1}{1 - a^2}}
\end{aligned}$$

More practice with path diagrams

$$\begin{tikzcd}
Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\
& U \arrow[u, "c"] \arrow[ur, "d"'] &
\end{tikzcd}$$

$$\begin{aligned}
\beta_{\text{Unadjusted}} &= \text{Cov}(A, Y) \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}}
\end{aligned}$$

What is $\text{Cov}(A, Y)$?
$b + cd$

What is $\text{Cov}(Z, A)$?
a

What is $\text{Cov}(Z, Y)$?
ab

$$\begin{aligned}
&\beta_{\text{Adjusted}} \\
&= \frac{\text{Cov}(A, Y) - \text{Cov}(Z, Y)\text{Cov}(Z, A)}{1 - [\text{Cov}(Z, A)]^2} \\
&= \frac{(b + cd) - (ab)(a)}{1 - a^2} \\
&= \frac{b(1 - a^2) + cd}{1 - a^2} \\
&= \underbrace{b}_{\text{Effect}} + \underbrace{cd}_{\text{Bias}} \underbrace{\frac{1}{1 - a^2}}_{\substack{\text{Bias} \\ \text{Amplifier} \\ \geq 1}}
\end{aligned}$$

What we have learned

Bias

We can quantify the bias
from unmeasured variables

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] & \end{tikzcd}$$

$$ac$$

We can reduce the bias
by adjusting for a proxy

$$\begin{tikzcd} A \arrow[r, "b"] & Y \\ U \arrow[u, "a"] \arrow[ur, "c"'] \arrow[r, "d"'] & \tilde{L} \end{tikzcd}$$

$$ac\frac{1-d^2}{1-a^2d^2}$$

We can do many things!
Example:
Bad controls magnify bias

$$\begin{tikzcd} Z \arrow[r, "a"] & A \arrow[r, "b"] & Y \\ & U \arrow[u, "c"] \arrow[ur, "d"'] & \end{tikzcd}$$

$$cd\frac{1}{1-a^2}$$

Cost of all of these things: Linear path model

- Homogeneous causal effects
- Linear causal effects

$$\mathsf{E}(Y \mid A, L) = \alpha + \beta_1 A + \beta_2 L$$

Learning goals for today

At the end of class, you will be able to:

1. Recognize differences between DAGs and linear path models
2. Determine the covariance between standardized variables using a linear path model
3. Quantify the biases of coefficient estimators using path models
4. Be equipped to read Elwert & Pfeffer (2022)

Let me know what you are thinking

tinyurl.com/CausalQuestions

Office hours TTh 11am-12pm and at
calendly.com/ianlundberg/office-hours
Come say hi!