# 25. Future treatments as proxies for confounding

## Class guest: Felix Elwert
## University of Wisconsin, Madison

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

17 Nov 2022

# Learning goals for today

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

1. Reason about when future treatments can proxy for unmeasured confounding

Note that this class is based on:

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.

# A few things we've recently covered

# Hernán & Cole 2009

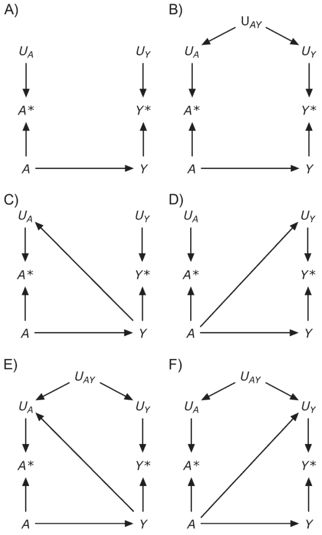



**Figure 2.** A structural classification of measurement error.

# Hernán & Cole 2009

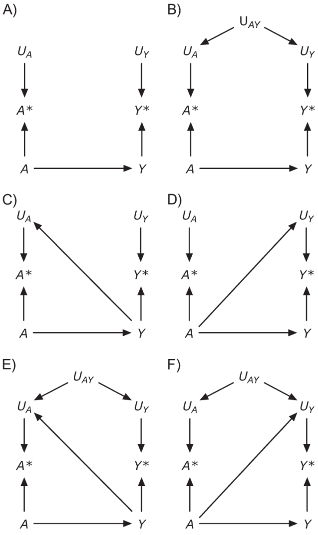



**Figure 2.** A structural classification of measurement error.

$$
\begin{tikzcd}
& \tilde{L} & \\
& L \arrow[u] \arrow[dl] \arrow[dr] & \\
A \arrow[rr] & & Y
\end{tikzcd}
$$

Hernán & Cole 2009

Elwert & Pfeffer 2022

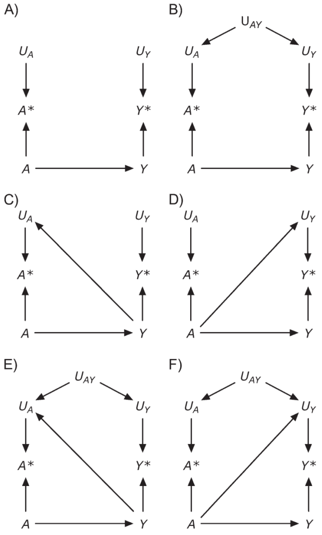
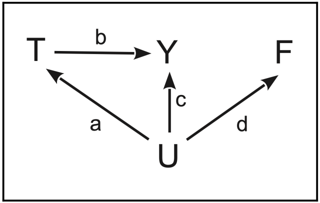


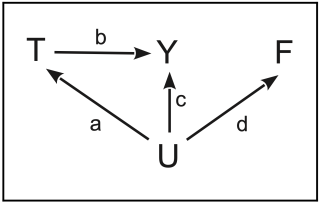

**Figure 2.** A structural classification of measurement error.



Elwert & Pfeffer 2022

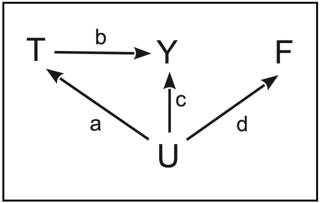
If you measure $U$

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



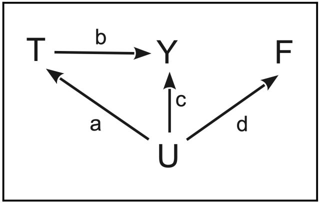
If you measure $U$

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




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

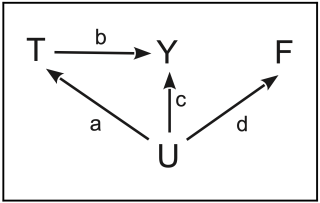
If you don't measure $U$



Elwert & Pfeffer 2022

If you don't measure $U$

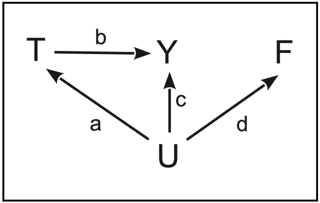
$$\mathsf{E}(Y \mid A, F) = \alpha + \beta A + \gamma F$$



Elwert & Pfeffer 2022

If you don't measure $U$

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




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

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

$$= \frac{b + ac - acd^2 - a^2 bd^2}{1 - a^2 d^2}$$

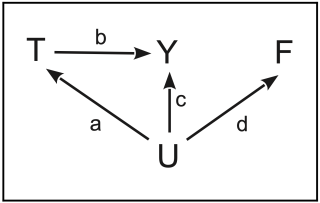
$$= \frac{b(1 - a^2 d^2) + ac(1 - d^2)}{1 - a^2 d^2}$$

$$= b + \underbrace{ac}_{\substack{\text{Bias} \\ \text{without} \\ \text{control}}} \underbrace{\frac{1 - d^2}{1 - a^2 d^2}}_{\substack{\text{Bias} \\ \text{Multiplier} \\ |M|<1}}$$

Elwert & Pfeffer 2022

If you don't measure $U$

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



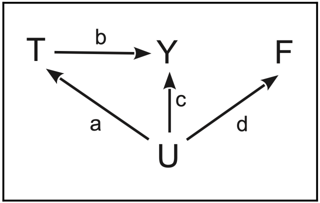

Control Estimator

$$\beta = b + \underbrace{ac}_{\substack{\text{Bias}\\ \text{without}\\ \text{control}}} \underbrace{\frac{1 - d^2}{1 - a^2 d^2}}_{\substack{\text{Bias}\\ \text{Multiplier}\\ |M|<1}}$$

Elwert & Pfeffer 2022

If you don't measure $U$

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




Control Estimator

$$\beta = b + \underbrace{ac}_{\substack{\text{Bias}\\\text{without}\\\text{control}}} \underbrace{\frac{1-d^2}{1-a^2d^2}}_{\substack{\text{Bias}\\\text{Multiplier}\\|M|<1}}$$

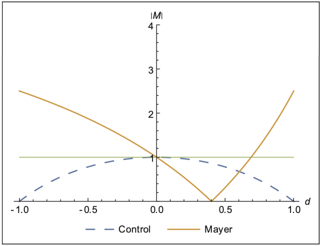
Difference (Mayer) Estimator

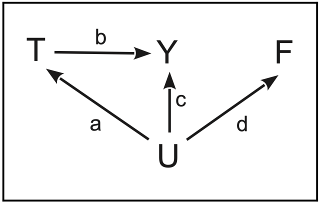
$$\beta - \gamma = b + \underbrace{ac}_{\substack{\text{Bias}\\\text{without}\\\text{control}}} \underbrace{\frac{a-d}{a-a^2d}}_{\substack{\text{Bias}\\\text{Multiplier}\\|M|}}$$

Elwert & Pfeffer 2022

If you don't measure $U$

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




Control Estimator

$$\beta = b + \underbrace{ac}_{\substack{\text{Bias}\\\text{without}\\\text{control}}} \underbrace{\frac{1-d^2}{1-a^2d^2}}_{\substack{\text{Bias}\\\text{Multiplier}\\|M|<1}}$$

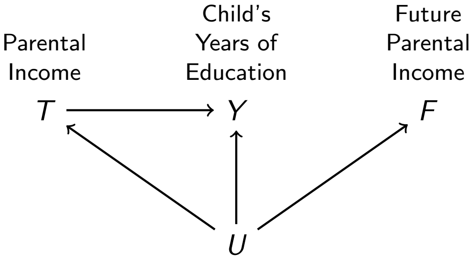
Difference (Mayer) Estimator

$$\beta - \gamma = b + \underbrace{ac}_{\substack{\text{Bias}\\\text{without}\\\text{control}}} \underbrace{\frac{a-d}{a-a^2d}}_{\substack{\text{Bias}\\\text{Multiplier}\\|M|}}$$

Parts of the paper we have not yet covered

# Empirical example

What is the effect of log parental income on years of education?




Panel Study of Income Dynamics. Born 1956–1968. ($n = 1,513$)

- $T$ log family income averaged at child age 13–17
- $Y$ years of education by age 24
- $F$ log family income at child age 25–29

**Table 3.** Estimating the Causal Effect of Parental Income on Children's Years of Education With and Without Future Treatments.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Coefficients | | | | |
| *T*: Parental income | .448 | .319 | .185 | .118 |
| | (.039)*** | (.049)*** | (.039)*** | (.041)** |
| *F*: Future parental income | | .274 | | .202 |
| | | (.088)** | | (.076)** |
| *X*: Controls | | | Yes | Yes |
| Difference in coefficients | | | | |
| $T - F$ | | .045 | | −.084 |
| | | (.126) | | (.098) |
| Test of equality of coefficients on *T*: *p* values | | | | |
| Model (1) versus (2): .0006 | | | | |
| Model (1) versus (3): .0000 | | | | |
| Model (3) versus (4): .0006 | | | | |
| Model (1) versus (4): .0000 | | | | |
| *N* | 1,513 | 1,513 | 1,513 | 1,513 |

*Note*. Standardized OLS regression coefficients (standard errors in parentheses); weighted. Significance tests for the difference between coefficients across models are using seemingly unrelated regression.

Statistical significance at $^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$ (two-tailed test).

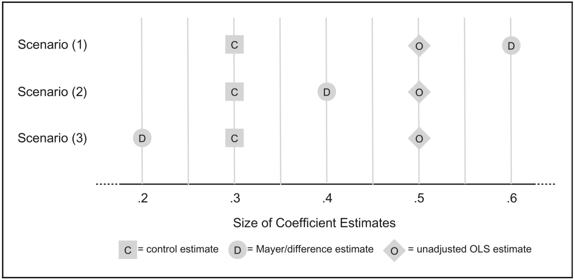
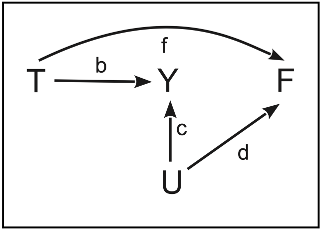
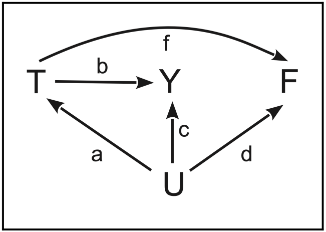



**Figure 4.** Illustration of the heuristic for choosing between estimates. The relative position of the control (*C*), difference (*D*), and unadjusted OLS (*O*) estimates can help the analyst decide between alternative estimates. In data generated by Figure 2, the location of the control estimate indicates the direction of unadjusted OLS bias (in this example, upward bias). In scenarios (1) and (2), the control estimate is preferred. In scenario (3), additional assumptions are needed to decide between the control and difference estimates.

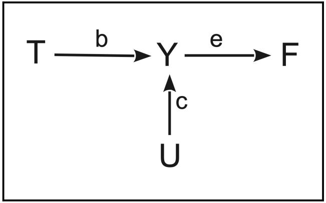
# Challenge 1: True State Dependence



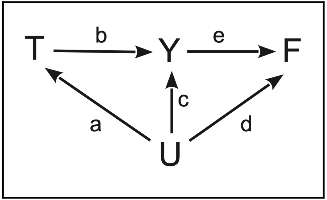
# Challenge 2: Confounded True State Dependence



# Challenge 3: Unconfounded Study with Selection



# Challenge 4: Confounded Study with Selection



# Challenge 4: Confounded Study with Selection

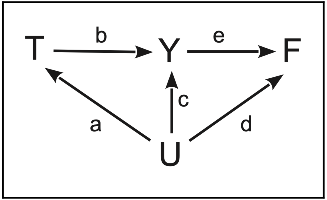



**Table 2.** Performance of the Control Estimator and the Mayer/Difference Estimator in the Presence of Selection and Weak to Moderate Path Parameters, $|p| < .5$.

| | Bias With | |
|---|---|---|
| Selection (e) | Control Estimator | Mayer/Difference Estimator |
| Selection $(e \neq 0)$ | Negligibly amplified or weakly reduced (see below) | Mostly amplified |
| Mild selection $(\|e\| \lesssim 0.3)$ | Weakly reduced | Mostly amplified |

# Nonparametric results

Previous results relied on a linear path model.
Next results rely only on a DAG (nonparametric).

# Nonparametric results

Previous results relied on a linear path model.
Next results rely only on a DAG (nonparametric).

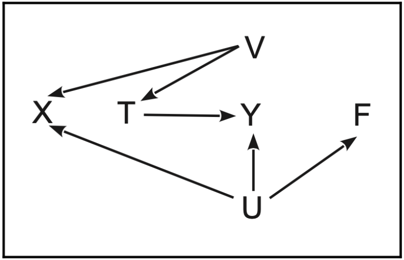
Motivating (but incomplete) intuition:

> *A future treatment $F$ cannot affect an outcome $Y$.*
>
> *If the outcome $Y$ is related to $F$,*
> *then there must be unobserved confounding.*

# Nonparametric Example 1.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$?
Is $T \to Y$ confounded given $X$?

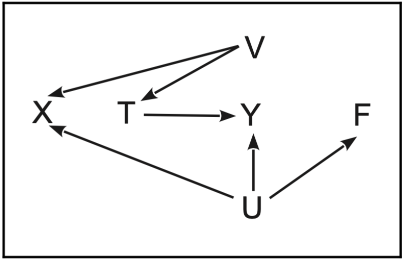


# Nonparametric Example 1.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow U \rightarrow F$

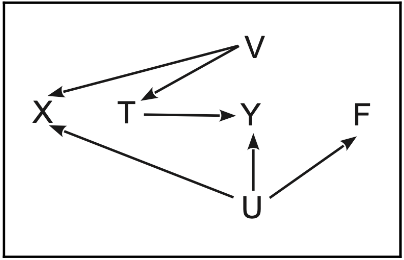
Is $T \rightarrow Y$ confounded given $X$?



# Nonparametric Example 1.

Does a relationship between $Y$ and $F$ imply confounding?

| | |
|---|---|
| Is $F$ related to $Y$? | Yes. $Y \leftarrow U \rightarrow F$ |
| Is $T \rightarrow Y$ confounded given $X$? | Yes. $T \leftarrow V \rightarrow \boxed{X} \leftarrow U \rightarrow Y$ |

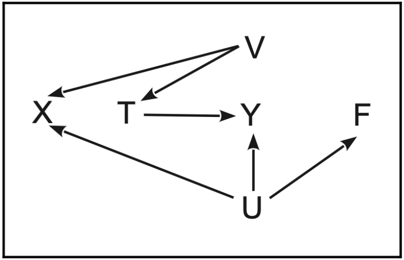


# Nonparametric Example 1.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow U \rightarrow F$

Is $T \rightarrow Y$ confounded given $X$? Yes. $T \leftarrow V \rightarrow \boxed{X} \leftarrow U \rightarrow Y$



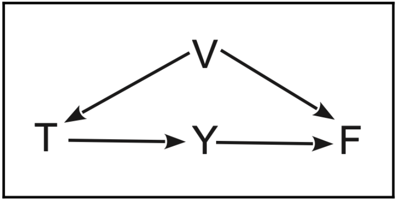

A conditional relationship between $Y$ and $F$ implies confounding.

# Nonparametric results: Example 2.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$?

Is $T \rightarrow Y$ confounded?

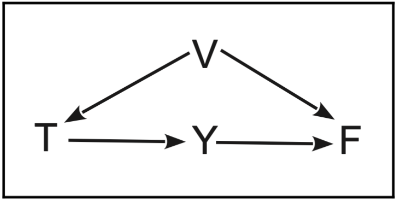


# Nonparametric results: Example 2.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow V \rightarrow F$

Is $T \rightarrow Y$ confounded?

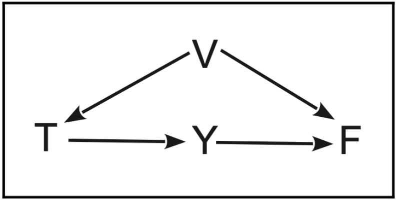


# Nonparametric results: Example 2.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow V \rightarrow F$

Is $T \rightarrow Y$ confounded? No.

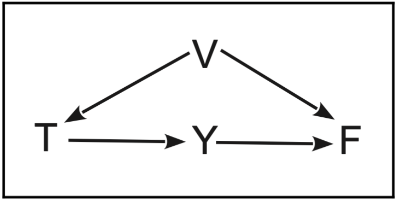


# Nonparametric results: Example 2.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow V \rightarrow F$

Is $T \rightarrow Y$ confounded? No.

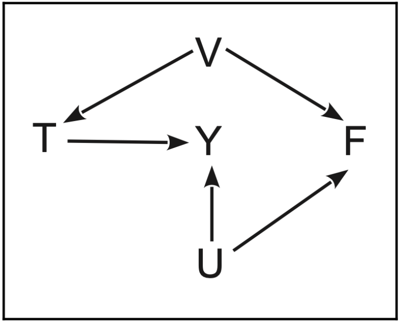



A conditional relationship between $Y$ and $F$ does not imply confounding.

# Nonparametric results: Example 3.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$?
Is $T \to Y$ confounded?

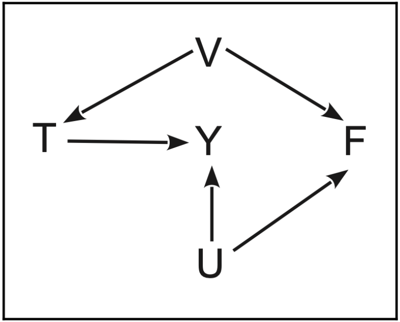


# Nonparametric results: Example 3.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow V \rightarrow F$

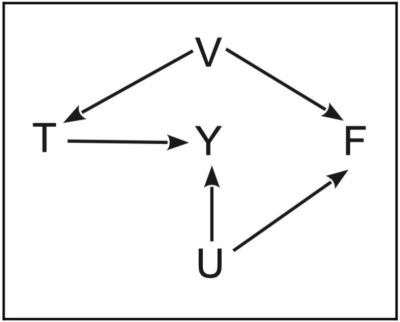
Is $T \rightarrow Y$ confounded?



# Nonparametric results: Example 3.

Does a relationship between $Y$ and $F$ imply confounding?

| | |
|---|---|
| Is $F$ related to $Y$? | Yes. $Y \leftarrow V \rightarrow F$ |
| Is $T \rightarrow Y$ confounded? | No. |

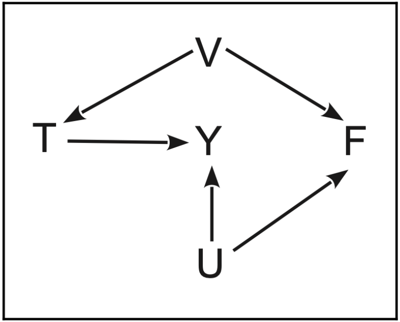


# Nonparametric results: Example 3.

Does a relationship between $Y$ and $F$ imply confounding?

Is $F$ related to $Y$? Yes. $Y \leftarrow V \rightarrow F$

Is $T \rightarrow Y$ confounded? No.




A conditional relationship between $Y$ and $F$ does not imply confounding.

# Two nonparametric results: A formal answer

What does the relationship between $F$ and $Y$ tell us about confounding?

# Two nonparametric results: A formal answer

What does the relationship between $F$ and $Y$ tell us about confounding?

$$\begin{tikzcd}
X \arrow[r] \arrow[rr, bend left] \arrow[rrr, bend left] \arrow[drr, bend right] & T \arrow[r] & Y & F \\
 & & V \arrow[ul] \arrow[u] \arrow[ur] &
\end{tikzcd}$$

Result 11. (Requires Assumption 1)

$$F \perp\!\!\!\perp Y \mid \{T, X\} \qquad \rightarrow \qquad \{Y^0, Y^1\} \perp\!\!\!\perp T \mid X$$

Result 12. (Requires Assumptions 1–3)

$$F \not\perp\!\!\!\perp Y \mid \{T, X\} \qquad \rightarrow \qquad \{Y^0, Y^1\} \not\perp\!\!\!\perp T \mid X$$

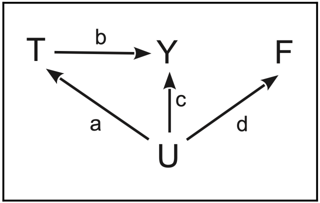
Assumption 1. There exists some unobserved $V$ such that $V \rightarrow T$ and $V \not\perp\!\!\!\perp F \mid \{T, X\}$

Assumption 2. All unobserved causes of $F$ also cause $T$

Assumption 3. $Y$ does not directly or indirectly cause $F$

# Discussion: Applied examples

The usefulness of future treatments relies heavily on this DAG.




**Exercise.** Discuss the plausibility of this DAG in applied cases.

tinyurl.com/FutureTreatmentsExamples

# Group 1

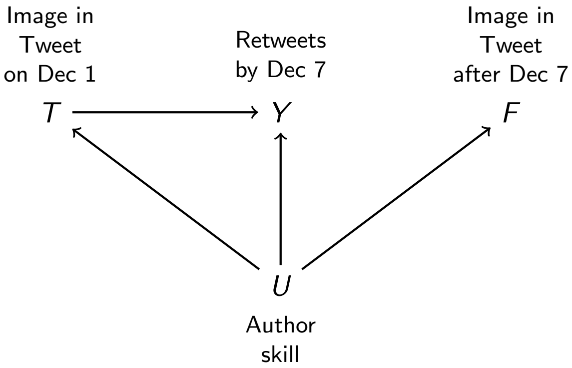
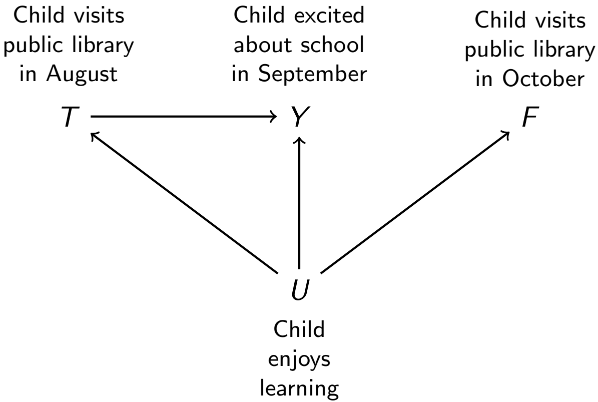
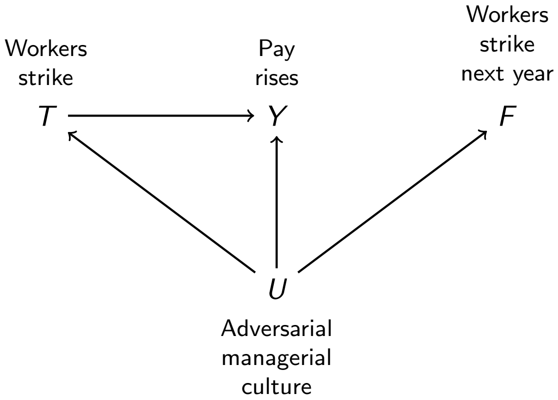


Image in
Tweet
on Dec 1
Retweets
by Dec 7
Image in
Tweet
after Dec 7
T
Y
F
U
Author
skill

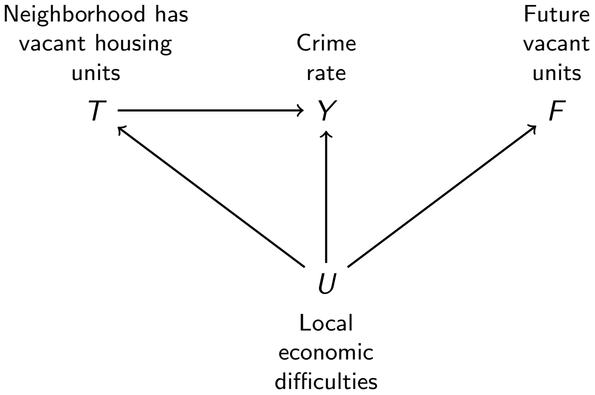
# Group 2


Child visits
public library
in August
T
Child excited
about school
in September
Y
Child visits
public library
in October
F
U
Child
enjoys
learning

# Group 3


Workers
strike
T
Pay
rises
Y
Workers
strike
next year
F
U
Adversarial
managerial
culture

# Group 4


Neighborhood has vacant housing units
T
Crime rate
Y
Future vacant units
F
U
Local economic difficulties

# Learning goals for today

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

1. Reason about when future treatments can proxy for unmeasured confounding

Note that this class is based on:

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.

# 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!