Versions of Treatment

A causal inference debate that sociologists have ignored

Ian Lundberg
Princeton University
Sociology Statistics Reading Group
20 February 2019

Introductory note for those finding these slides online

These slides were prepared for the Sociology Statistics Reading Group at Princeton. Everyone read the following paper in advance:

> Hernán, Miguel A. 2016. Does water kill? A call for less casual causal inferences. *Annals of Epidemiology*, 26(10):674–680. [link]

At times, these slides intentionally emphasize alternative positions to those presented by Hernán (2016), such as the possibility that consistency is not an assumption but is rather a consequence of the assumptions embedded in a causal DAG (Pearl, 2010). I emphasize this alternative view not because my personal position is strongly one way or the other, but because it will promote better discussion among a group that read the former but not the latter. See references at the end for further reading.

Hernán (2016): Does Water Kill?

London cholera epidemic, 1854.
John Snow deduced that the water was the cause of death.

Source: Wikimedia Commons

Hernán (2016): Does Water Kill?

Does drinking water kill?

Hernán (2016): Does Water Kill?

Does drinking **fresh** water kill?

Hernán (2016): Does Water Kill?

Does drinking **a swig of** fresh water kill?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water **from the Broad Street pump** kill?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump **between August 31 and September 10** kill?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump between August 31 and September 10 kill

compared with drinking all your water from other pumps?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump between August 31 and September 10 **and not initiating a rehydration treatment if diarrhea starts** kill

compared with drinking all your water from other pumps?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump between August 31 and September 10 and not initiating a rehydration treatment if diarrhea starts kill

compared with drinking all your water from other pumps?

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump between August 31 and September 10 and not initiating a rehydration treatment if diarrhea starts kill

compared with drinking all your water from other pumps?

The **definition** of the causal effect is unclear without details.

Hernán (2016): Does Water Kill?

Does drinking a swig of fresh water from the Broad Street pump between August 31 and September 10 and not initiating a rehydration treatment if diarrhea starts kill

compared with drinking all your water from other pumps?

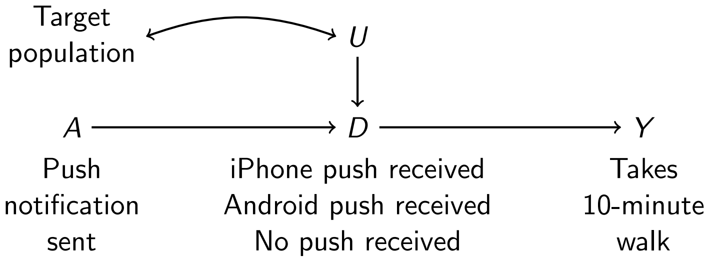
The **definition** of the causal effect is unclear without details.

Recommendation: Specify versions
"until no meaningful vagueness remains,"
(Hernán, 2016)

But some vagueness is

unavoidable

in both **experimental** and **observational** social science.


Target population
U
A
D
Y
Push notification sent
iPhone push received
Android push received
No push received
Takes 10-minute walk

Continuous treatment

Overwork

Collapsed by **researcher**

Researcher collapses continuous D

$$A = \mathcal{C}(D) = \mathbb{I}(D > 50)$$

$$D \longrightarrow Y$$

Employment hours → Hourly wage

Collapsed by **researcher**

Continuous treatment

Overwork

Researcher collapses continuous D

$$A = \mathcal{C}(D) = \mathbb{I}(D > 50)$$

$D \longrightarrow Y$

Employment hours → Hourly wage

Categorical treatment

Occupations

Researcher collapses categorical D

Class scheme: $A = \mathcal{C}(D)$

$D \longrightarrow Y$

Occupation → Status

	Continuous treatment	Categorical treatment
Collapsed by **researcher**	**Overwork** Researcher collapses continuous D $A = \mathcal{C}(D) = \mathbb{I}(D > 50)$ $D \longrightarrow Y$ Employment hours → Hourly wage	**Occupations** Researcher collapses categorical D Class scheme: $A = \mathcal{C}(D)$ $D \longrightarrow Y$ Occupation → Status
Collapsed by **respondent**	**Self-rated health** Respondent collapses continuous D 5-point scale: $A = \mathcal{C}(D)$ $D \longrightarrow Y$ Health → Lifespan	

	Continuous treatment	**Categorical** treatment
Collapsed by **researcher**	**Overwork** Researcher collapses continuous D $A = \mathcal{C}(D) = \mathbb{I}(D > 50)$ $D \longrightarrow Y$ Employment hours → Hourly wage	**Occupations** Researcher collapses categorical D Class scheme: $A = \mathcal{C}(D)$ $D \longrightarrow Y$ Occupation → Status
Collapsed by **respondent**	**Self-rated health** Respondent collapses continuous D 5-point scale: $A = \mathcal{C}(D)$ $D \longrightarrow Y$ Health → Lifespan	**Returns to college** Respondent collapses categorical D Completed college: $A = \mathcal{C}(D)$ $D \longrightarrow Y$ College degree (institution and major) → Earnings

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

→ 1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs

2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments

3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   → A) Potential outcomes
   B) Stochastic counterfactuals
   C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   A) Experimental studies: Effects may not generalize
   B) Observational studies:
      — "Effects" may be an unusual average
      — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

$$\text{Causal effect} = Y_i(a') - Y_i(a)$$

Potential outcomes

$$\underbrace{\{Y_i(a)\}}$$

Potential outcomes:
Deterministic consequence of a

$$\underbrace{Y_i^{\text{Obseved}} = Y_i(A_i)}$$

Observed outcome:
Random because A_i is random

Imbens and Rubin (2015, p. 10):

> ...for each unit, there are no different forms or versions of each treatment level which lead to different potential outcomes.

$$A \longrightarrow D \longrightarrow Y$$

A	D	Y
Push notification	iPhone push received	Walks
	Android push received	10 minutes
	No push received	

$Y_i(a)$ is deterministic under either:

- Deterministic detailed treatment assignment D given A

$$\mathbb{P}(D_i = d \mid A_i = a) = \begin{cases} 1 & \text{for one value of } d \\ 0 & \text{for all other values of } d \end{cases} \qquad \forall a$$

- Treatment variation irrelevance (adapted from VanderWeele 2009)

$$Y_i(d) = Y_i(d') \ \forall \ \{d, d'\} \text{ such that}$$

$$\mathbb{P}(D_i = d \mid A_i = a) > 0 \text{ and } \mathbb{P}(D_i = d' \mid A_i = a) > 0$$

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - → A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - → B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Stochastic counterfactuals[1] allow a more plausible assumption of treatment variation irrelevance.

Under **fixed** counterfactuals

Treatment-variation irrelevance:

$$Y_i(a, d_a) = Y_i(a, d_a') \quad \forall \quad \{d_a, d_a'\} \in \mathcal{D}_a$$

Thus can define $Y_i(a) \equiv Y_i(a, d_a)$ for any d_a.

Consistency:

$$\text{If } A_i = a, \quad \exists \; d_a \in \mathcal{D}_a \text{ such that } Y_i^{\text{Observed}} = Y_i(a, d_a)$$

---

[1] VanderWeele 2009, Imbens and Rubin 2015 p. 12

Stochastic counterfactuals[1] allow a more plausible assumption of treatment variation irrelevance.

Under **stochastic** counterfactuals

Treatment-variation irrelevance:

$$Y_i(a, d_a) \overset{D}{\sim} Y_i(a, d_a') \quad \forall \quad \{d_a, d_a'\} \in \mathcal{D}_a$$

Thus can define $Y_i(a) \sim Y_i(a, d_a)$ for any d_a.

Consistency:

$$\text{If } A_i = a, \quad \exists \ d_a \in \mathcal{D}_a \text{ such that } Y_i^{\text{Observed}} = Y_i(a, d_a)$$

[1] VanderWeele 2009, Imbens and Rubin 2015 p. 12

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - → B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   A) Potential outcomes
   B) Stochastic counterfactuals
   → C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   A) Experimental studies: Effects may not generalize
   B) Observational studies:
      — "Effects" may be an unusual average
      — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

In causal graphs, the absence of hidden versions is a **theorem** rather than an **assumption** (Pearl, 2010).

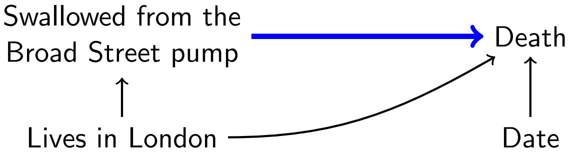
In causal graphs, the absence of hidden versions is a **theorem** rather than an **assumption** (Pearl, 2010).

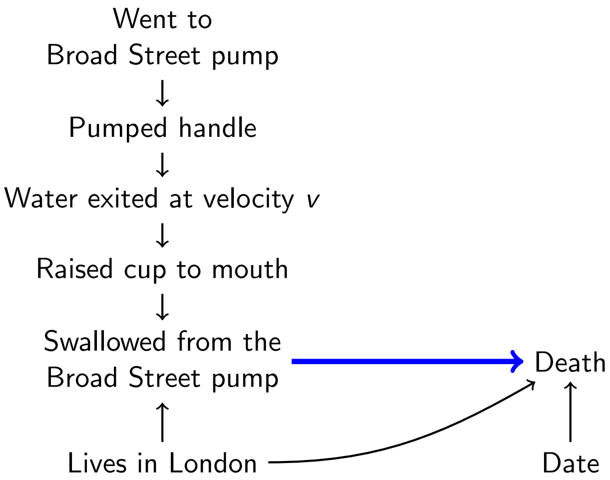
Treatment effects are defined by the DAG.

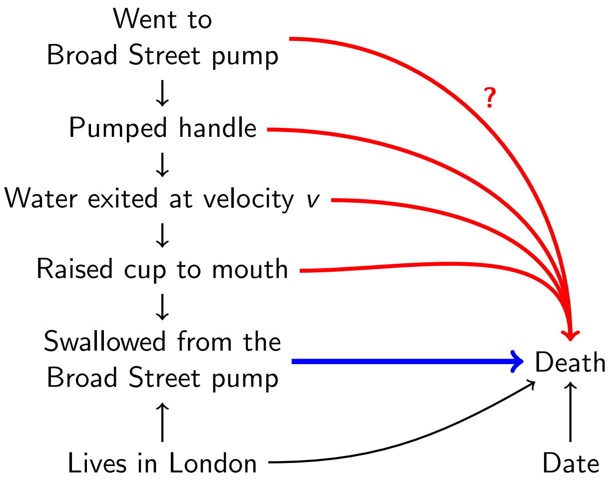
In causal graphs, the absence of hidden versions is a **theorem** rather than an **assumption** (Pearl, 2010).

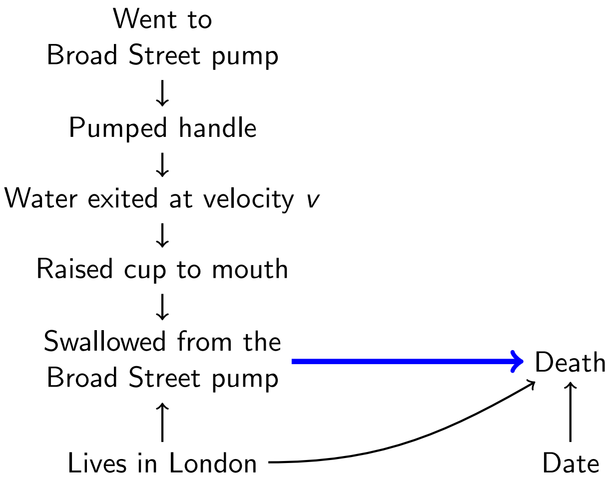
Treatment effects are defined by the DAG.

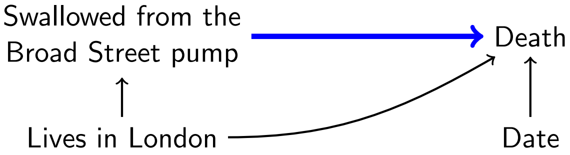
A correct DAG **implies** a well-defined effect.

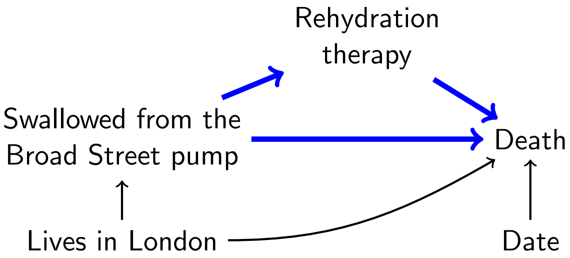

Swallowed from the Broad Street pump
Death
Lives in London
Date

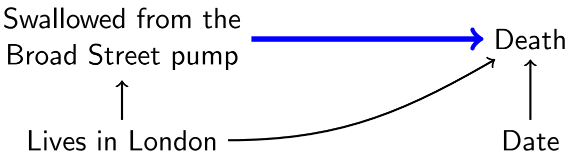

Went to
Broad Street pump
Pumped handle
Water exited at velocity v
Raised cup to mouth
Swallowed from the
Broad Street pump
Death
Lives in London
Date


Went to
Broad Street pump
Pumped handle
Water exited at velocity v
Raised cup to mouth
Swallowed from the
Broad Street pump
Lives in London
?
Death
Date


Went to
Broad Street pump
Pumped handle
Water exited at velocity v
Raised cup to mouth
Swallowed from the
Broad Street pump
Death
Lives in London
Date

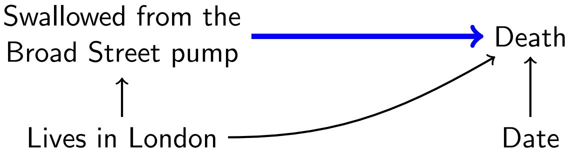

Swallowed from the Broad Street pump
Death
Lives in London
Date


Rehydration therapy
Swallowed from the Broad Street pump
Death
Lives in London
Date

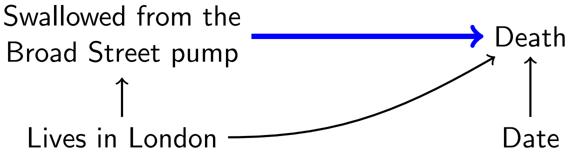

Swallowed from the Broad Street pump
Death
Lives in London
Date

$$\mathbb{E}\left(\text{Death} \mid \begin{array}{c}\text{do(Swallowed from Broad Street pump),}\\ \text{Living in London on August 31 – September 10}\end{array}\right)$$

$$-\,\mathbb{E}\left(\text{Death} \mid \begin{array}{c}\text{do(Did not swallow from Broad Street pump),}\\ \text{Living in London on August 31 – September 10}\end{array}\right)$$



Things are vague only if the **graph** is insufficiently precise (and thus wrong).



Consistency: $Y_i^{\text{Observed}} = Y_i(a_i)$. Is this an assumption?

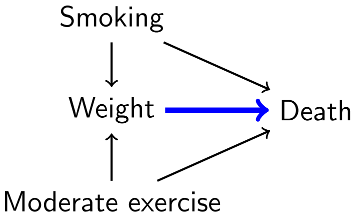
Consistency: $Y_i^{\text{Observed}} = Y_i(a_i)$. Is this an assumption?

Hernán (2016): Potential death Y under weight a depends on whether weight is set by smoking or by moderate exercise.

Consistency: $Y_i^{\text{Observed}} = Y_i(a_i)$. Is this an assumption?

Hernán (2016): Potential death Y under weight a depends on whether weight is set by smoking or by moderate exercise.

But we could just label these as confounding variables.
Not clear that consistency is an assumption.



Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - → C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   A) Potential outcomes
   B) Stochastic counterfactuals
   C) Causal graphs

→ 2. When it matters: **Consequences** of collapsed versions
   A) Experimental studies: Effects may not generalize
   B) Observational studies:
      — "Effects" may be an unusual average
      — Heterogeneous treatment effects may really be the effects of heterogeneous treatments

3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - → A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Experiments identify causal effects with minimal assumptions,

but they often seek to generalize to a **target population**.

Experiments identify causal effects with minimal assumptions,

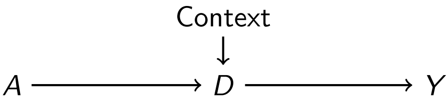
but they often seek to generalize to a **target population**.

Versions of treatment make generalization difficult.
(Hernán and VanderWeele, 2011)

If there is **contextual variation** in $A \rightarrow Y$, then the effect may not generalize to new contexts.

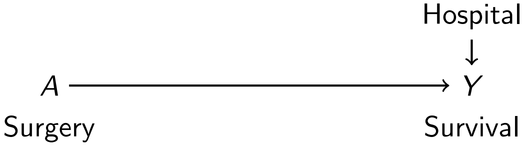
If there is **contextual variation** in $A \rightarrow Y$, then the effect may not generalize to new contexts.

If this arises from variation in $A \rightarrow D$, then measuring D may promote transportability of inference.



If there is **contextual variation** in $A \to Y$, then the effect may not generalize to new contexts.

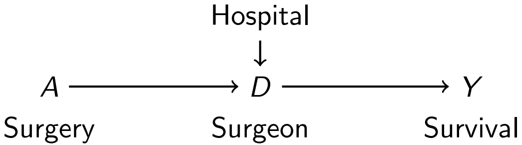
If this arises from variation in $A \to D$, then measuring D may promote transportability of inference.



$A \to Y$ may differ across hospitals

If there is **contextual variation** in $A \to Y$, then the effect may not generalize to new contexts.

If this arises from variation in $A \to D$, then measuring D may promote transportability of inference.



$A \to Y$ may differ across hospitals

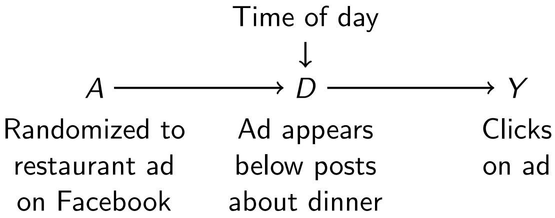
If there is **contextual variation** in $A \to Y$, then the effect may not generalize to new contexts.

If this arises from variation in $A \to D$, then measuring D may promote transportability of inference.

$A \to Y$ may differ by time of day

If there is **contextual variation** in $A \to Y$, then the effect may not generalize to new contexts.

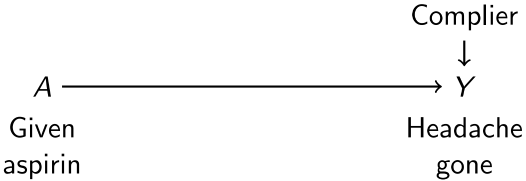
If this arises from variation in $A \to D$, then measuring D may promote transportability of inference.



$A \to Y$ may differ by time of day

If there is **contextual variation** in $A \rightarrow Y$, then the effect may not generalize to new contexts.

If this arises from variation in $A \rightarrow D$, then measuring D may promote transportability of inference.



$A \rightarrow Y$ may differ by rates of compliance

If there is **contextual variation** in $A \rightarrow Y$, then the effect may not generalize to new contexts.

If this arises from variation in $A \rightarrow D$, then measuring D may promote transportability of inference.

Complier

$\downarrow$

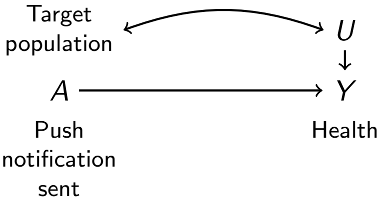
$A \longrightarrow D \longrightarrow Y$

Given aspirin | Takes Aspirin | Headache gone

$A \rightarrow Y$ may differ by rates of compliance

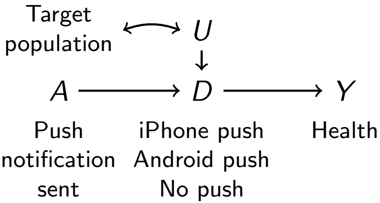
Generalizing experimental evidence

Generalization **impossible**

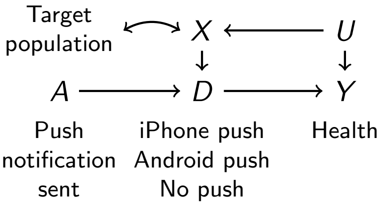


Generalization **possible**

by randomizing D

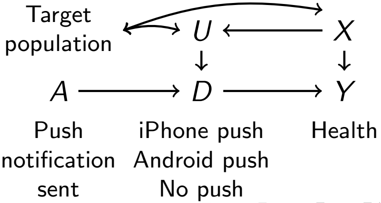


Generalization **impossible**



Generalization **possible**

by randomizing D and observing X



Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - → A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - → B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Continuous treatment | Categorical treatment

Collapsed by **researcher**

Overwork

Researcher collapses continuous D

$$A = \mathcal{C}(D) = \mathbb{I}(D > 50)$$

$$D \longrightarrow Y$$

Employment hours → Hourly wage

Occupations

Researcher collapses categorical D

Class scheme: $A = \mathcal{C}(D)$

$$D \longrightarrow Y$$

Occupation → Status

Collapsed by **respondent**

Self-rated health

Respondent collapses continuous D

5-point scale: $A = \mathcal{C}(D)$

$$D \longrightarrow Y$$

Health → Lifespan

Returns to college

Respondent collapses categorical D

Completed college: $A = \mathcal{C}(D)$

$$D \longrightarrow Y$$

College degree (institution and major) → Earnings

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - → — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined.

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$

Observed detailed treaments d mapping to collapsed treatment c'

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$

Observed detailed treaments d mapping to collapsed treatment c'	Observed detailed treaments d mapping to collapsed treatment c

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$

Observed detailed treaments d mapping to collapsed treatment c'

↓

Random draw

Observed detailed treaments d mapping to collapsed treatment c

$$D \longrightarrow Y$$

Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$

Observed detailed treaments d mapping to collapsed treatment c'	Observed detailed treaments d mapping to collapsed treatment c
↓ Random draw	↓ Random draw

$$D \longrightarrow Y$$

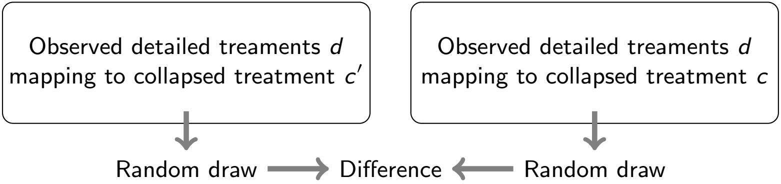
Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$



$$D \longrightarrow Y$$

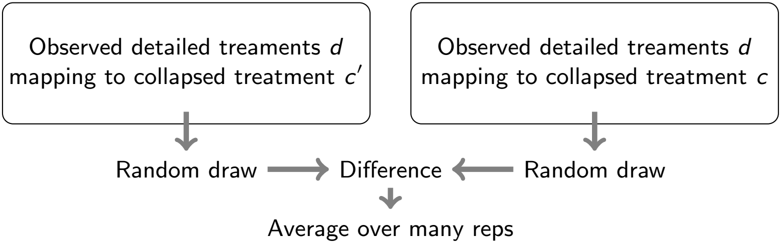
Because the collapsed $\mathcal{C}(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

(VanderWeele and Hernàn (2013, Prop. 8), though notation differs.)

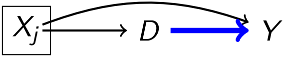
$$\mathbb{E}(Y \mid \mathcal{C}(D) = c) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d)\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c)\big)$$

One causal contrast

$$\mathbb{E}(Y \mid \mathcal{C}(D) = c') - \mathbb{E}(Y \mid \mathcal{C}(D) = c)$$



With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)

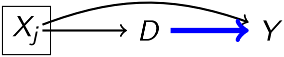


$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



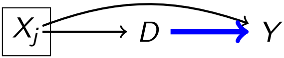
$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\textstyle\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\Big)$$

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

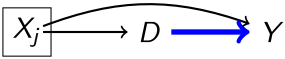
$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\left(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\right)$$

among $\vec{X} = \vec{x}$

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

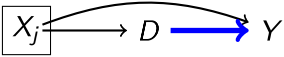
One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\Big)$$

Observed detailed treaments d mapping to collapsed treatment c'

among $\vec{X} = \vec{x}$

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

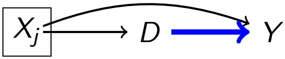
$$\textstyle\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\Big)$$

Observed detailed treaments d mapping to collapsed treatment c'

Observed detailed treaments d mapping to collapsed treatment c

among $\vec{X} = \vec{x}$

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\Big)$$

Observed detailed treaments d mapping to collapsed treatment c'

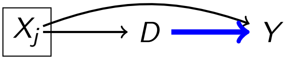
↓

Random draw

Observed detailed treaments d mapping to collapsed treatment c

among $\vec{X} = \vec{x}$

With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)

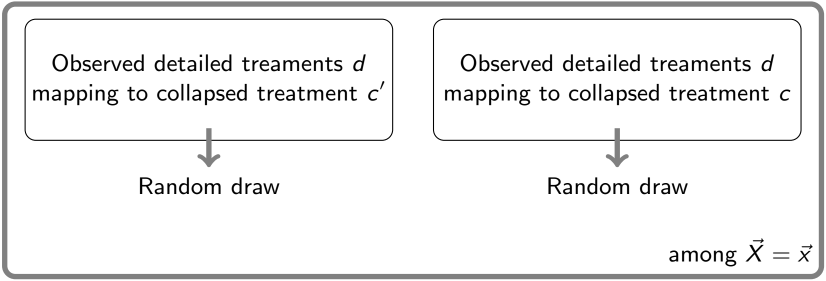


$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

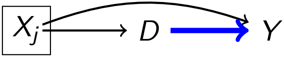
$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\Big)$$



With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)

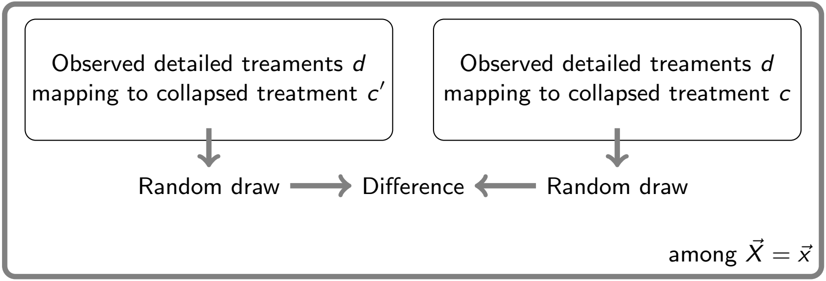


$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

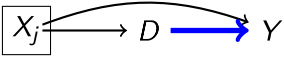
$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big) \mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\textstyle\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x}) \Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) \Big)$$



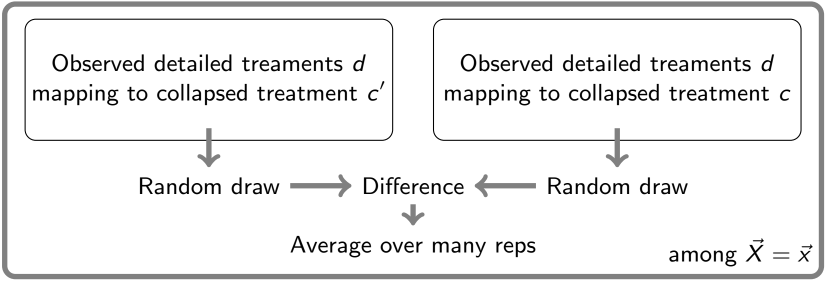
With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)



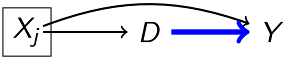
$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) = \sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}(Y \mid \text{do}(D = d), \vec{X} = \vec{x}) \mathbb{P}(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x})$$

One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x}) \Big(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) \Big)$$



With covariates. (equivalent to VanderWeele and Hernàn 2013, Prop. 8)

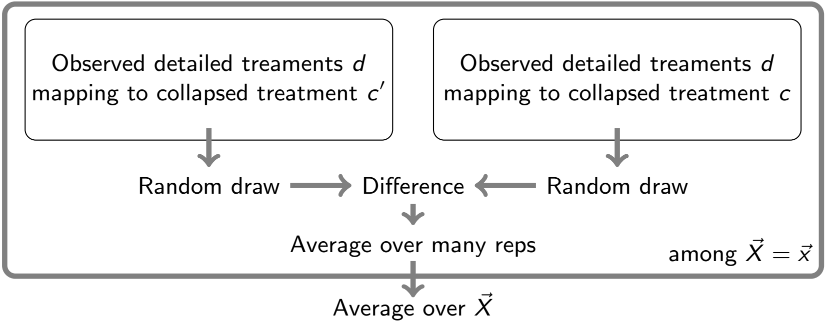


$$\mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x}) =$$

$$\sum_{d \in \mathcal{C}^{-1}(c)} \mathbb{E}\big(Y \mid \text{do}(D = d), \vec{X} = \vec{x}\big)\mathbb{P}\big(D = d \mid D \in \mathcal{C}^{-1}(c), \vec{X} = \vec{x}\big)$$

One causal contrast

$$\sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x})\left(\mathbb{E}(Y \mid \mathcal{C}(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid \mathcal{C}(D) = c, \vec{X} = \vec{x})\right)$$



Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   A) Potential outcomes
   B) Stochastic counterfactuals
   C) Causal graphs

2. When it matters: **Consequences** of collapsed versions
   A) Experimental studies: Effects may not generalize
   B) Observational studies:
      → — "Effects" may be an unusual average
      — Heterogeneous treatment effects may really be the effects of heterogeneous treatments

3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - → — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

What is the effect of **overwork** ($>$ 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

What is the effect of **overwork** ($>$ 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

What is the effect of **overwork** ($>$ 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

$$\mathbb{E}\big(\text{Wage} \mid \text{do(Overwork)}, \text{Man}\big) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\big(\text{Wage} | \text{do(Hours}=h), \text{Man}\big) \\ \times \mathbb{P}\big(\text{Hours}=h | \text{Hours}>50, \text{Man}\big) \end{pmatrix}$$

What is the effect of **overwork** (> 50 hours) on the **hourly wage** of those working at least 40 hours per week?
Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

$$\mathbb{E}\big(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Man}\big) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\big(\text{Wage}|\text{do}(\text{Hours}=h),\text{Man}\big) \\ \times\mathbb{P}\big(\text{Hours}=h|\text{Hours}>50,\text{Man}\big) \end{pmatrix}$$

$$\mathbb{E}\big(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Woman}\big) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\big(\text{Wage}|\text{do}(\text{Hours}=h),\text{Woman}\big) \\ \times\mathbb{P}\big(\text{Hours}=h|\text{Hours}>50,\text{Woman}\big) \end{pmatrix}$$

What is the effect of **overwork** (> 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

$$\mathbb{E}\big(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Man}\big) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\big(\text{Wage}|\text{do}(\text{Hours}=h),\text{Man}\big) \\ \times\mathbb{P}\big(\text{Hours}=h|\text{Hours}>50,\text{Man}\big) \end{pmatrix}$$

$$\mathbb{E}\big(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Woman}\big) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\big(\text{Wage}|\text{do}(\text{Hours}=h),\text{Woman}\big) \\ \times\mathbb{P}\big(\text{Hours}=h|\text{Hours}>50,\text{Woman}\big) \end{pmatrix}$$

Heterogeneous treatment effects

What is the effect of **overwork** (> 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

$$\mathbb{E}\left(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Man}\right) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\left(\text{Wage}|\text{do}(\text{Hours}=h),\text{Man}\right) \\ \times\mathbb{P}\left(\text{Hours}=h|\text{Hours}>50,\text{Man}\right) \end{pmatrix}$$

$$\mathbb{E}\left(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Woman}\right) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\left(\text{Wage}|\text{do}(\text{Hours}=h),\text{Woman}\right) \\ \times\mathbb{P}\left(\text{Hours}=h|\text{Hours}>50,\text{Woman}\right) \end{pmatrix}$$

Effects of heterogeneous treatments

What is the effect of **overwork** ($>$ 50 hours) on the **hourly wage** of those working at least 40 hours per week? Suppose **gender** is the only source of confounding.

$$\text{Overwork} = \mathcal{C}(\text{Employment hours}) = \mathbb{I}(\text{Hours} > 50)$$

$$\mathbb{E}\left(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Man}\right) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\left(\text{Wage}|\text{do}(\text{Hours}=h),\text{Man}\right) \\ \times\mathbb{P}\left(\text{Hours}=h|\text{Hours}>50,\text{Man}\right) \end{pmatrix}$$

$$\mathbb{E}\left(\text{Wage} \mid \text{do}(\text{Overwork}), \text{Woman}\right) = \sum_{h>50} \begin{pmatrix} \mathbb{E}\left(\text{Wage}|\text{do}(\text{Hours}=h),\text{Woman}\right) \\ \times\mathbb{P}\left(\text{Hours}=h|\text{Hours}>50,\text{Woman}\right) \end{pmatrix}$$

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - → — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
   - A) Potential outcomes
   - B) Stochastic counterfactuals
   - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
   - A) Experimental studies: Effects may not generalize
   - B) Observational studies:
     - — "Effects" may be an unusual average
     - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments

→ 3. **Recommendations:** What to do

What To Do

In randomized **experiments** aiming to generalize:

- Randomize a detailed treatment
- Theorize context-specific versions likely to remain

In **observational** studies:

- Estimate at the finest level of detail measured
  - Promotes a simple definition of the effect
  - Promotes transportability
  - Promotes clear policy implications
- If treatment remains vague, state the implied intervention.

Versions of Treatment

A Causal Inference Debate Sociologists Have Ignored

1. Formalizing the **problem**
    - A) Potential outcomes
    - B) Stochastic counterfactuals
    - C) Causal graphs
2. When it matters: **Consequences** of collapsed versions
    - A) Experimental studies: Effects may not generalize
    - B) Observational studies:
        - — "Effects" may be an unusual average
        - — Heterogeneous treatment effects may really be the effects of heterogeneous treatments
3. **Recommendations:** What to do

Appendix

Some define the **assumptions** for causal inference as:

- Ignorable treatment assignment
  — Violated if the treated would do better even without treatment
- Positivity
  — Violated if $\mathbb{P}$(Treated) is 0 or 1 for some units
- Stable Unit Treatment Value Assumption
  — Violated if there is interference
  — Violated if there are hidden versions of treatment

In this setup, **hidden versions** are the second part of SUTVA. Social scientists often focus on the first assumptions and give less thought to this part of SUTVA.

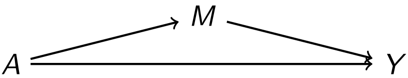
Why not use the $Y(a, d_a)$ notation?

One could state potential outcomes as a function of both treatment and treatment version (VanderWeele, 2009; Hernán and VanderWeele, 2011; VanderWeele and Hernàn, 2013).

Versions of treatment

$Y_i(a, d_a)$ is unnecessary notation, though. Because only one a exists for any given d, $Y_i(d)$ carries the same information.
In contrast, this is useful in mediation.

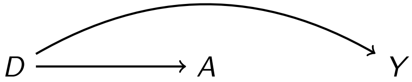
Mediation



$Y_i(a, m)$ is valuable. Because A is not fixed given M, there exist multiple $\{a, a'\}$ with $Y_i(a, m) \neq Y_i(a', m)$.

Why not put A on the DAG as a consequence of D?

When the researchers take a detailed treatment D and coarsen it into an aggregate treatment A, Hernán and VanderWeele (2011) put it in the DAG as a consequence of D.

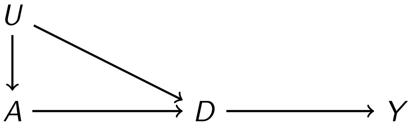


The reasons not to do this are

1. In a DAG, it is useful to be able to conceive of an intervention to any given node. Because $D \to A$ is deterministic, it is hard to imagine an intervention to A which has no consequence for D. By the DAG, this intervention would have no consequence for D. This seems hard to swallow.
2. Perhaps A is not deterministic: it is *reported* D. But this seems like a whole different set of issues, and it is clear even without the DAG that intervening to change a report would have no consequence for Y.

What about when $A \to D$ is confounded?

When treatment precedes version, Hernán and VanderWeele (2011) also include cases like below:



The reasons not to do this are

1. The edge $U \to A$ implies this is an observational study rather than an experiment. In observational studies, I usually do not believe the story that A is assigned first, followed by D. I think in observational studies D is typically the only variable involved.
2. We already have transportability issues from $U \to D$ alone. Omitting $U \to A$ helps to highlight these problems in the scenario when A is randomized.

References

Hernán, M. A. 2016. Does water kill? A call for less casual causal inferences. *Annals of epidemiology*, 26(10):674–680.

Hernán, M. A. and T. J. VanderWeele 2011. Compound treatments and transportability of causal inference. *Epidemiology*, 22(3):368.

Imbens, G. W. and D. B. Rubin 2015. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Pearl, J. 2010. Brief report: On the consistency rule in causal inference: Axiom, definition, assumption, or theorem? *Epidemiology*, Pp. 872–875.

VanderWeele, T. J. 2009. Concerning the consistency assumption in causal inference. *Epidemiology*, 20(6):880–883.

VanderWeele, T. J. and M. A. Hernàn 2013. Causal inference under multiple versions of treatment. *Journal of Causal Inference*, 1(1):1–20.