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:


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.
Versions of Treatment

Formalizing the Problem

When It Matters

What To Do

Target population

\(U\)

\(A\)  \(\rightarrow\)  \(D\)  \(\rightarrow\)  \(Y\)

Push notification sent

iPhone push received

Android push received

No push received

Takes 10-minute walk
**Continuous treatment**

**Overwork**

Researcher collapses continuous $D$

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

$D \rightarrow Y$

Employment hours  Hourly wage

Collapsed by respondent
Versions of Treatment

Formalizing the Problem

When It Matters

What To Do

Continuous treatment

Overwork

Researcher collapses continuous $D$

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

$D \rightarrow Y$

Collapsed by researcher

Employment hours

Hourly wage

Categorical treatment

Occupations

Researcher collapses categorical $D$

Class scheme: $A = C(D)$

$D \rightarrow Y$

Occupation

Status

Respondent collapses continuous $D$

Respondent collapses categorical $D$

Self-rated health

Returns to college

Employment

Hourly wage

Occupation

Status

Earnings
**Continuous** treatment

**Overwork**

Researcher collapses continuous $D$

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

$D \rightarrow Y$

Employment hours    Hourly wage

**Categorical** treatment

**Occupations**

Researcher collapses categorical $D$

Class scheme: $A = C(D)$

$$D \rightarrow Y$$

Occupation    Status

**Self-rated health**

Respondent collapses continuous $D$

5-point scale: $A = C(D)$

$$D \rightarrow Y$$

Health    Lifespan
Versions of Treatment

Formalizing the Problem

When It Matters

What To Do

**Continuous treatment**

**Overwork**

- Researcher collapses continuous $D$
  
  - $A = C(D) = I(D > 50)$

- $D \rightarrow Y$
  
  - Employment hours
  
  - Hourly wage

**Collapsed by respondent**

**Self-rated health**

- Respondent collapses continuous $D$
  
  - 5-point scale: $A = C(D)$

- $D \rightarrow Y$
  
  - Health
  
  - Lifespan

**Categorical treatment**

**Occupations**

- Researcher collapses categorical $D$
  
  - Class scheme: $A = C(D)$

- $D \rightarrow Y$
  
  - Occupation
  
  - Status

**Collapsed by researcher**

**Returns to college**

- Respondent collapses categorical $D$

- Completed college: $A = C(D)$

- $D \rightarrow Y$
  
  - College degree
    
  - Earnings
    
    - (institution and major)
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
Causal effect = $Y_i(a') - Y_i(a)$

\begin{align*}
\{Y_i(a)\} & \quad \text{Potential outcomes:} \\
\text{Deterministic consequence of } a & \\
\end{align*}

\begin{align*}
Y_i^{\text{Observed}} & = Y_i(A_i) \\
\text{Observed outcome:} & \\
\text{Random because } A_i \text{ is random} & \\
\end{align*}

Imbens and Rubin (2015, p. 10):

\ldots for each unit, there are no different forms or versions of each treatment level which lead to different potential outcomes.
A $\rightarrow$ D $\rightarrow$ Y

- Push notification
- iPhone push received
- Android push received
- No push received

Y_i(a) is deterministic under either:

- Deterministic detailed treatment assignment D given A

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

- Treatment variation irrelevance (adapted from VanderWeele 2009)

$$Y_i(d) = Y_i(d') \quad \forall \{d, d'\} \text{ such that } \mathbb{P}(D_i = d \mid A_i = a) > 0 \text{ and } \mathbb{P}(D_i = d' \mid A_i = a) > 0$$
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:

\[
Y_i(a, d_a) = Y_i(a, d'_a) \quad \forall \quad \{d_a, d'_a\} \in D_a
\]

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

Consistency:

If \(A_i = a\), \(\exists\ d_a \in D_a\) such that \(Y_i^{\text{Observed}} = Y_i(a, d_a)\)

---

**Stochastic counterfactuals**\(^1\) allow a more plausible assumption of treatment variation irrelevance.

Under **stochastic** counterfactuals

Treatment-variation irrelevance:

\[
Y_i(a, d_a) \overset{D}{\sim} Y_i(a, d'_a) \quad \forall \quad \{d_a, d'_a\} \in D_a
\]

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

Consistency:

If \(A_i = a\), \(\exists \ d_a \in D_a\) such that \(Y_i^{\text{Observed}} = Y_i(a, d_a)\)

---

**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 \textbf{theorem} rather than an \textbf{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

$\text{Went to Broad Street pump}$

$\text{Pumped handle}$

$\text{Water exited at velocity } v$

$\text{Raised cup to mouth}$

$\text{Rehydration}$

$\text{E}(\text{Death} \mid \text{do(Swallowed from Broad Street pump)}), \text{Living in London on August 31 – September 10})$

$\text{−} \text{E}(\text{Death} \mid \text{do(} \text{Did not swallow from Broad Street pump}) , \text{Living in London on August 31 – September 10}))$

$\text{Things are vague only if the graph is insufficiently precise (and thus wrong).}$
Went to Broad Street pump
\[\Downarrow\]
Pumped handle
\[\Downarrow\]
Water exited at velocity \(v\)
\[\Downarrow\]
Raised cup to mouth
\[\Downarrow\]
Swallowed from the Broad Street pump
\[\Uparrow\]
Lives in London
\[\rightarrow\]
Death
\[\Uparrow\]
Date

\[\text{Date} \rightarrow \text{Went to Broad Street pump} \rightarrow \text{Pumped handle} \rightarrow \text{Water exited at velocity } v \rightarrow \text{Raised cup to mouth} \rightarrow \text{Swallowed from the Broad Street pump} \rightarrow \text{Death}\]
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

Things are vague only if the graph is insufficiently precise (and thus wrong).
Went to
Broad Street pump
\[ \Downarrow \]
Pumped handle
\[ \Downarrow \]
Water exited at velocity \( v \)
\[ \Downarrow \]
Raised cup to mouth
\[ \Downarrow \]
Swallowed from the
Broad Street pump
\[ \Uparrow \]
Lives in London
\[ \Rightarrow \]
Death
\[ \Uparrow \]
Date
Swallowed from the Broad Street pump → Death
Lives in London → Date

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

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

(Rehydration therapy)

Death (Swallowed from Broad Street pump, Lives in London on August 31 – September 10) - Death (Did not swallow from Broad Street pump, Lives in London on August 31 – September 10)

Things are vague only if the graph is insufficiently precise (and thus wrong).
\begin{align*}
&\mathbb{E}\left( \text{Death} \mid \text{do(Swallowed from Broad Street pump), Living in London on August 31 – September 10} \right) \\
&\quad - \mathbb{E}\left( \text{Death} \mid \text{do(Did not swallow from Broad Street pump), Living in London on August 31 – September 10} \right)
\end{align*}
Things are vague only if the graph is insufficiently precise (and thus wrong).

- Swallowed from the Broad Street pump
- Lives in London
- Date
- Death
**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.

\[
\text{Context} \quad \downarrow \\
A \quad \rightarrow \quad Y
\]
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 ↠ D ↠ Y
  ↓
Context
```
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 across hospitals.
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.

$$
\text{Hospital} \\
\downarrow \\
A \rightarrow D \rightarrow Y
$$

$A \rightarrow Y$ may differ across hospitals
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 time of day

- Randomized to restaurant ad on Facebook
- Clicks on ad
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.

```
Time of day

\[ A \longrightarrow D \longrightarrow Y \]

Randomized to restaurant ad
Ad appears below posts about dinner

clicks
```

$A \rightarrow Y$ 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.

\[
\begin{array}{c}
A \\
\text{Given aspirin}
\end{array} \quad \Downarrow \\
\text{Complier} \quad \rightarrow \\
\begin{array}{c}
Y \\
\text{Headache gone}
\end{array}
\]

$A \rightarrow Y$ 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.

$$\begin{align*}
\text{Complier} \\
\downarrow \\
A \rightarrow D \rightarrow Y \\
\text{Given aspirin} \quad \text{Takes Aspirin} \quad \text{Headache gone}
\end{align*}$$

$A \rightarrow Y$ may differ by rates of compliance
Generalizing experimental evidence

Generalization **impossible**

- Target population
- $A \rightarrow U$ (Push notification sent)
- $U \rightarrow Y$ (Health)

Generalization **possible** by randomizing $D$

- Target population
- $A \rightarrow D$ (iPhone push, Android push, No push)
- $D \rightarrow Y$ (Health)

Generalization **impossible**

- Target population
- $A \rightarrow X$ (Push notification sent)
- $X \leftrightarrow U$ (Health)

Generalization **possible** by randomizing $D$ and observing $X$

- Target population
- $A \rightarrow D$ (iPhone push, Android push, No push)
- $D \rightarrow Y$ (Health)

Note: The diagram illustrates the possible and impossible generalizations of experimental evidence in different scenarios involving target populations and interventions.
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

Formalizing the Problem

When It Matters

What To Do

**Continuous treatment**

**Overwork**

Researcher collapses continuous $D$

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

Collapsed by researcher

| Employment hours | Hourly wage |

**Categorical treatment**

**Occupations**

Researcher collapses categorical $D$

Class scheme: $A = C(D)$

Collapsed by researcher

| Occupation | Status |

**Self-rated health**

Respondent collapses continuous $D$

5-point scale: $A = C(D)$

Collapsed by respondent

| Health | Lifespan |

**Returns to college**

Respondent collapses categorical $D$

Completed college: $A = C(D)$

| 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 $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined.
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: (VanderWeele and Hernán (2013, Prop. 8), though notation differs.)

$$
\mathbb{E}(Y \mid C(D) = c) = \sum_{d \in C^{-1}(c)} \mathbb{E}(Y \mid \text{do}(D = d)) \mathbb{P}(D = d \mid D \in C^{-1}(c))
$$
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: (VanderWeele and Hernán (2013, Prop. 8), though notation differs.)

$$
E(Y | C(D) = c) = \sum_{d \in C^{-1}(c)} E(Y | do(D = d)) P(D = d | D \in C^{-1}(c))
$$

**One causal contrast**

$$
E(Y | C(D) = c') - E(Y | C(D) = c)
$$
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: (VanderWeele and Hernán (2013, Prop. 8), though notation differs.)

$$E(Y | C(D) = c) = \sum_{d \in C^{-1}(c)} E(Y | do(D = d)) P(D = d | D \in C^{-1}(c))$$

**One causal contrast**

$$E(Y | C(D) = c') - E(Y | C(D) = c)$$

Observed detailed treatments $d$ mapping to collapsed treatment $c'$
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

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

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

**One causal contrast**

\[
\mathbb{E}(Y \mid C(D) = c') - \mathbb{E}(Y \mid C(D) = c)
\]
Because the collapsed \( C(D) \) is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine:

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

**One causal contrast**

\[
\mathbb{E}(Y \mid C(D) = c') - \mathbb{E}(Y \mid C(D) = c)
\]
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: (VanderWeele and Hernán (2013, Prop. 8), though notation differs.)

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

**One causal contrast**

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

- **Observed detailed treatments $d$ mapping to collapsed treatment $c'$**
  - Random draw

- **Observed detailed treatments $d$ mapping to collapsed treatment $c$**
  - Random draw
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: (VanderWeele and Hernán (2013, Prop. 8), though notation differs.)

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

**One causal contrast**

$$E(Y \mid C(D) = c') - E(Y \mid C(D) = c)$$
Because the collapsed $C(D)$ is not in the causal graph, the effect of a collapsed treatment is undefined. We might examine: 

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

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

**One causal contrast**

$$E(Y \mid C(D) = c') - E(Y \mid C(D) = c)$$

- Observed detailed treatments $d$ mapping to collapsed treatment $c'$
- Observed detailed treatments $d$ mapping to collapsed treatment $c$

Random draw $\rightarrow$ Difference $\rightarrow$ Random draw $\leftrightarrow$ Random draw

Average over many reps
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

One causal contrast
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

One causal contrast

\[
\sum_{\tilde{x}} \mathbb{P}(\tilde{X} = \tilde{x}) \left( \mathbb{E}(Y \mid C(D) = c', \tilde{X} = \tilde{x}) - \mathbb{E}(Y \mid C(D) = c, \tilde{X} = \tilde{x}) \right)
\]
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

**One causal contrast**

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

among \( \vec{X} = \vec{x} \)
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

**One causal contrast**

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

Observed detailed treatments \(d\) mapping to collapsed treatment \(c'\)

among \(\tilde{X} = \tilde{x}\)
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

\[
E(Y \mid C(D) = c, \tilde{X} = \tilde{x}) = \sum_{d \in C^{-1}(c)} E(Y \mid \text{do}(D = d), \tilde{X} = \tilde{x}) P(D = d \mid D \in C^{-1}(c), \tilde{X} = \tilde{x})
\]

**One causal contrast**

\[
\sum_{\tilde{x}} P(\tilde{X} = \tilde{x}) \left( E(Y \mid C(D) = c', \tilde{X} = \tilde{x}) - E(Y \mid C(D) = c, \tilde{X} = \tilde{x}) \right)
\]

Observed detailed treatments \(d\) mapping to collapsed treatment \(c'\)

Observed detailed treatments \(d\) mapping to collapsed treatment \(c\)

among \(\tilde{X} = \tilde{x}\)
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

\[
E(Y | C(D) = c, \vec{X} = \vec{x}) = \sum_{d \in C^{-1}(c)} E(Y | \text{do}(D = d), \vec{X} = \vec{x}) P(D = d | D \in C^{-1}(c), \vec{X} = \vec{x})
\]

**One causal contrast**

\[
\sum_{\vec{x}} P(\vec{X} = \vec{x}) \left( E(Y | C(D) = c', \vec{X} = \vec{x}) - E(Y | C(D) = c, \vec{X} = \vec{x}) \right)
\]

- Observed detailed treatments \(d\) mapping to collapsed treatment \(c'\)
- Observed detailed treatments \(d\) mapping to collapsed treatment \(c\)
- Random draw
- Among \(\vec{X} = \vec{x}\)
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

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

One causal contrast

$$\sum{\vec{x}} P(\vec{X} = \vec{x}) \left( E(Y | C(D) = c', \vec{X} = \vec{x}) - E(Y | C(D) = c, \vec{X} = \vec{x}) \right)$$

- Observed detailed treatments $d$ mapping to collapsed treatment $c'$
- Observed detailed treatments $d$ mapping to collapsed treatment $c$

Random draw among $\vec{X} = \vec{x}$
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

One causal contrast

\[ \sum_{\tilde{x}} \mathbb{P}(\tilde{X} = \tilde{x}) \left( \mathbb{E}(Y \mid C(D) = c', \tilde{X} = \tilde{x}) - \mathbb{E}(Y \mid C(D) = c, \tilde{X} = \tilde{x}) \right) \]
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

One causal contrast

\[ \sum_{\vec{x}} \mathbb{P}(\vec{X} = \vec{x}) \left( \mathbb{E}(Y \mid C(D) = c', \vec{X} = \vec{x}) - \mathbb{E}(Y \mid C(D) = c, \vec{X} = \vec{x}) \right) \]
With covariates. (equivalent to VanderWeele and Hernán 2013, Prop. 8)

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

One causal contrast

\[
\sum_{\tilde{x}} \mathbb{P}(\tilde{X} = \tilde{x}) \left( \mathbb{E}(Y \mid C(D) = c', \tilde{X} = \tilde{x}) - \mathbb{E}(Y \mid C(D) = c, \tilde{X} = \tilde{x}) \right)
\]
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.

Gender → Employment hours → Hourly wage

Overwork = C(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.

\[
\begin{align*}
\text{Overwork} &= C(\text{Employment hours}) = \mathbb{1}(\text{Hours} > 50) \\
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Man}) &= \sum_{h>50} \left( \mathbb{E}(\text{Wage} \mid \text{do(Hours=h)}, \text{Man}) \times \mathbb{P}(\text{Hours=h} \mid \text{Hours>50}, \text{Man}) \right) 
\end{align*}
\]
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.

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

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Man}) = \sum_{h>50} \left( \mathbb{E}(\text{Wage} \mid \text{do(Hours=}\, h\), \text{Man}) \times \mathbb{P}(\text{Hours=}\, h \mid \text{Hours}>50, \text{Man}) \right)
\]

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Woman}) = \sum_{h>50} \left( \mathbb{E}(\text{Wage} \mid \text{do(Hours=}\, h\), \text{Woman}) \times \mathbb{P}(\text{Hours=}\, h \mid \text{Hours}>50, \text{Woman}) \right)
\]
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.

![Diagram](attachment:image.png)

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

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Man}) = \sum_{h>50} \left( \mathbb{E}(\text{Wage} \mid \text{do(Hours}=h\text{)}, \text{Man}) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Man}) \right)
\]

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Woman}) = \sum_{h>50} \left( \mathbb{E}(\text{Wage} \mid \text{do(Hours}=h\text{)}, \text{Woman}) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Woman}) \right)
\]

**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.

\[
\text{Gender} \quad \rightarrow \quad \text{Employment hours} \quad \rightarrow \quad \text{Hourly wage}
\]

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

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Man}) = \sum_{h>50} \left( \mathbb{E}\left( \text{Wage} \mid \text{do(Hours}=h), \text{Man} \right) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Man}) \right)
\]

\[
\mathbb{E}(\text{Wage} \mid \text{do(Overwork)}, \text{Woman}) = \sum_{h>50} \left( \mathbb{E}\left( \text{Wage} \mid \text{do(Hours}=h), \text{Woman} \right) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Woman}) \right)
\]

**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.

![Diagram](https://via.placeholder.com/150)

\[
\begin{align*}
\text{Gender} & \rightarrow \text{Employment hours} \rightarrow \text{Hourly wage} \\
\text{Overwork} = C(\text{Employment hours}) &= \mathbb{I}(\text{Hours} > 50) \\
\mathbb{E}(Wage \mid \text{do(Overwork)}, \text{Man}) &= \sum_{h>50} \left( \mathbb{E}(Wage \mid \text{do(Hours}=h), \text{Man}) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Man}) \right) \\
\mathbb{E}(Wage \mid \text{do(Overwork)}, \text{Woman}) &= \sum_{h>50} \left( \mathbb{E}(Wage \mid \text{do(Hours}=h), \text{Woman}) \times \mathbb{P}(\text{Hours}=h \mid \text{Hours}>50, \text{Woman}) \right)
\end{align*}
\]
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 $P(\text{Treated})$ 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 Hernán, 2013).

**Versions of treatment**

$$A \rightarrow D \rightarrow Y$$

$Y_i(a, d_a)$ 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**

$$A \rightarrow M \rightarrow Y$$

$Y_i(a, m)$ 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 $D$.

\[
D \rightarrow A \rightarrow Y
\]

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 $D \rightarrow A$ 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 $A \rightarrow D$ is confounded?

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

\[ U \rightarrow \quad \downarrow \quad \rightarrow \quad A \rightarrow D \rightarrow Y \]

The reasons not to do this are

1. The edge $U \rightarrow A$ 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 $U \rightarrow D$ alone. Omitting $U \rightarrow A$ helps to highlight these problems in the scenario when $A$ is randomized.
References


