

# The Causal Framework Wars in Economics

Lennart Ackermans



Munich Center for Mathematical Philosophy

October 5, 2025

- “Reflections on the 2021 Nobel Memorial Prize Awarded to David Card, Joshua Angrist, and Guido Imbens” (Ackermans 2023)
- Imbens (2020)
- Imbens’ response (via email and 2023 seminar)
- My response to Imbens: economists need causal graphs too

- Leamer (1983): **Econometrics is in a state of crisis.** “Hardly anyone takes data analyses seriously. Or perhaps more accurately, hardly anyone takes anyone else’s data analyses seriously.”

- Leamer (1983): **Econometrics is in a state of crisis.** “Hardly anyone takes data analyses seriously. Or perhaps more accurately, hardly anyone takes anyone else’s data analyses seriously.”
- 1990s: **the Causal diagnosis.** Imbens and Angrist believe the problem was a lack of causal methodology. Important studies and methodological papers: Angrist (1990), Angrist and Krueger (1991), Imbens and Angrist (1994), and Angrist, Imbens, and Rubin (1996).

- Leamer (1983): **Econometrics is in a state of crisis.** “Hardly anyone takes data analyses seriously. Or perhaps more accurately, hardly anyone takes anyone else’s data analyses seriously.”
- 1990s: **the Causal diagnosis.** Imbens and Angrist believe the problem was a lack of causal methodology. Important studies and methodological papers: Angrist (1990), Angrist and Krueger (1991), Imbens and Angrist (1994), and Angrist, Imbens, and Rubin (1996).
- The economists’ favourite causal framework: Potential Outcomes (Rubin 1974).

- Leamer (1983): **Econometrics is in a state of crisis.** “Hardly anyone takes data analyses seriously. Or perhaps more accurately, hardly anyone takes anyone else’s data analyses seriously.”
- 1990s: **the Causal diagnosis.** Imbens and Angrist believe the problem was a lack of causal methodology. Important studies and methodological papers: Angrist (1990), Angrist and Krueger (1991), Imbens and Angrist (1994), and Angrist, Imbens, and Rubin (1996).
- The economists’ favourite causal framework: Potential Outcomes (Rubin 1974).
- Pearl (2000) has a similar diagnosis, but offers a different framework as solution: the **graphical framework.**

- Econometricians were expert *pure statisticians*.

- Econometricians were expert *pure statisticians*.
- *Pure statistics*: estimating population parameters (means, partial correlation coefficients, etc.) from samples.

- Econometricians were expert *pure statisticians*.
- *Pure statistics*: estimating population parameters (means, partial correlation coefficients, etc.) from samples.
- For example, do linear regressions of the form:

$$Y_i = \alpha A_i + \beta B_i + \varepsilon_i, \quad (1)$$

where  $\alpha$  and  $\beta$  are understood to be population statistics and  $\varepsilon_i$  a residual.

- Econometricians were expert *pure statisticians*.
- *Pure statistics*: estimating population parameters (means, partial correlation coefficients, etc.) from samples.
- For example, do linear regressions of the form:

$$Y_i = \alpha A_i + \beta B_i + \varepsilon_i, \quad (1)$$

where  $\alpha$  and  $\beta$  are understood to be population statistics and  $\varepsilon_i$  a residual.

- However, economists engage(d) in semi-causal talk. Implying:  $\alpha$  and  $\beta$  are *not* population parameters but causal effect sizes.

- Econometricians were expert *pure statisticians*.
- *Pure statistics*: estimating population parameters (means, partial correlation coefficients, etc.) from samples.
- For example, do linear regressions of the form:

$$Y_i = \alpha A_i + \beta B_i + \varepsilon_i, \quad (1)$$

where  $\alpha$  and  $\beta$  are understood to be population statistics and  $\varepsilon_i$  a residual.

- However, economists engage(d) in semi-causal talk. Implying:  $\alpha$  and  $\beta$  are *not* population parameters but causal effect sizes.
- **The identification problem**: infinitely many causal mechanisms can generate identical populations.

$$Y_i = \alpha X_i + \varepsilon_i, \quad (2)$$

- Textbook “causal” econometrics:  $\varepsilon_i$  is the combined effect of omitted variables.  $\alpha$  can be estimated when  $\text{Cov}(X_i, \varepsilon_i) = 0$ . If not, there is “omitted variable bias”.

$$Y_i = \alpha X_i + \varepsilon_i, \quad (2)$$

- Textbook “causal” econometrics:  $\varepsilon_i$  is the combined effect of omitted variables.  $\alpha$  can be estimated when  $\text{Cov}(X_i, \varepsilon_i) = 0$ . If not, there is “omitted variable bias”.
- OK—but what about this model in which  $\varepsilon'_i = aX_i + b\varepsilon_i$ :

$$Y_i = \beta X_i + \varepsilon'_i, \quad (3)$$

$$Y_i = \alpha X_i + \varepsilon_i, \quad (2)$$

- Textbook “causal” econometrics:  $\varepsilon_i$  is the combined effect of omitted variables.  $\alpha$  can be estimated when  $\text{Cov}(X_i, \varepsilon_i) = 0$ . If not, there is “omitted variable bias”.
- OK—but what about this model in which  $\varepsilon'_i = aX_i + b\varepsilon_i$ :

$$Y_i = \beta X_i + \varepsilon'_i, \quad (3)$$

- But if  $\text{Cov}(X_i, \varepsilon_i) = 0$  then  $\text{Cov}(X_i, \varepsilon'_i) \neq 0$ .

$$Y_i = \alpha X_i + \varepsilon_i, \quad (2)$$

- Textbook “causal” econometrics:  $\varepsilon_i$  is the combined effect of omitted variables.  $\alpha$  can be estimated when  $\text{Cov}(X_i, \varepsilon_i) = 0$ . If not, there is “omitted variable bias”.
- OK—but what about this model in which  $\varepsilon'_i = aX_i + b\varepsilon_i$ :

$$Y_i = \beta X_i + \varepsilon'_i, \quad (3)$$

- But if  $\text{Cov}(X_i, \varepsilon_i) = 0$  then  $\text{Cov}(X_i, \varepsilon'_i) \neq 0$ .
- To determine the **correct** model, and to know whether  $\text{Cov}(X_i, \varepsilon_i) = 0$ , one needs a causal theory that answers at least:

$$Y_i = \alpha X_i + \varepsilon_i, \quad (2)$$

- Textbook “causal” econometrics:  $\varepsilon_i$  is the combined effect of omitted variables.  $\alpha$  can be estimated when  $\text{Cov}(X_i, \varepsilon_i) = 0$ . If not, there is “omitted variable bias”.
- OK—but what about this model in which  $\varepsilon'_i = aX_i + b\varepsilon_i$ :

$$Y_i = \beta X_i + \varepsilon'_i, \quad (3)$$

- But if  $\text{Cov}(X_i, \varepsilon_i) = 0$  then  $\text{Cov}(X_i, \varepsilon'_i) \neq 0$ .
- To determine the **correct** model, and to know whether  $\text{Cov}(X_i, \varepsilon_i) = 0$ , one needs a causal theory that answers at least:
  - What is a causal variable? What is an omitted causal variable? What is a causal effect size?

- Imagine a population in which individuals are assigned a treatment  $T_i = 1$  or no treatment  $T_i = 0$ . The *potential outcomes* if they were treated and not treated are  $Y_i(1)$  and  $Y_i(0)$ . The **individual treatment effect** is

$$Y_i(1) - Y_i(0).$$

- Imagine a population in which individuals are assigned a treatment  $T_i = 1$  or no treatment  $T_i = 0$ . The *potential outcomes* if they were treated and not treated are  $Y_i(1)$  and  $Y_i(0)$ . The **individual treatment effect** is

$$Y_i(1) - Y_i(0).$$

- We can often estimate the **average treatment effect**

$$E[Y_i(1) - Y_i(0)].$$

- Imagine a population in which individuals are assigned a treatment  $T_i = 1$  or no treatment  $T_i = 0$ . The *potential outcomes* if they were treated and not treated are  $Y_i(1)$  and  $Y_i(0)$ . The **individual treatment effect** is

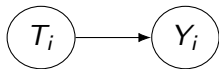
$$Y_i(1) - Y_i(0).$$

- We can often estimate the **average treatment effect**

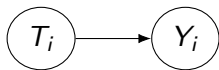
$$E[Y_i(1) - Y_i(0)].$$

- Such methods (RCT, instrumental variables) rely on the assumption of **unconfoundedness**:  $T_i$  is independent of the potential population outcomes  $(Y_i(0), Y_i(1))$ .

- Causal structure is encoded in **directed acyclic graphs** (DAGs).



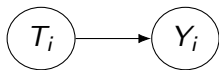
- Causal structure is encoded in **directed acyclic graphs** (DAGs).



- Much like potential outcomes, we imagine a counterfactual model in which values of  $T_i$  are intervened on.



- Causal structure is encoded in **directed acyclic graphs** (DAGs).

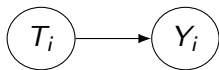


- Much like potential outcomes, we imagine a counterfactual model in which values of  $T_i$  are intervened on.



- The average treatment effect is  $E[Y_i | \hat{T}_i = 1] - E[Y_i | \hat{T}_i = 0]$ .

- Causal structure is encoded in **directed acyclic graphs** (DAGs).



- Much like potential outcomes, we imagine a counterfactual model in which values of  $T_i$  are intervened on.



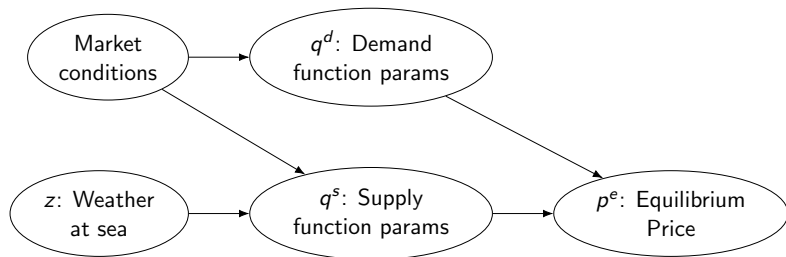
- The average treatment effect is  $E[Y_i | \hat{T}_i = 1] - E[Y_i | \hat{T}_i = 0]$ .
- Assumptions like unconfoundedness can be assessed based on the graph's structure, assuming it is accurate.

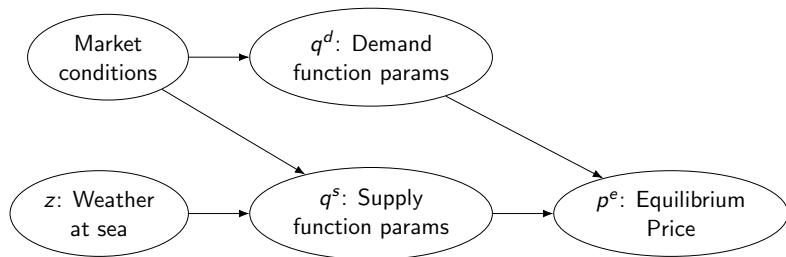
- PO more easily allows one to use causal methods that rely on **non-graphical** assumptions (e.g., shape restrictions). (Imbens 2020; Ackermans 2023)

- PO more easily allows one to use causal methods that rely on **non-graphical** assumptions (e.g., shape restrictions). (Imbens 2020; Ackermans 2023)
- Causal graphs are essential for encoding the causal assumptions on which many methods rely. Our causal background information comes in terms of **causal structure**, not counterfactual population statistics. (Ackermans 2023)

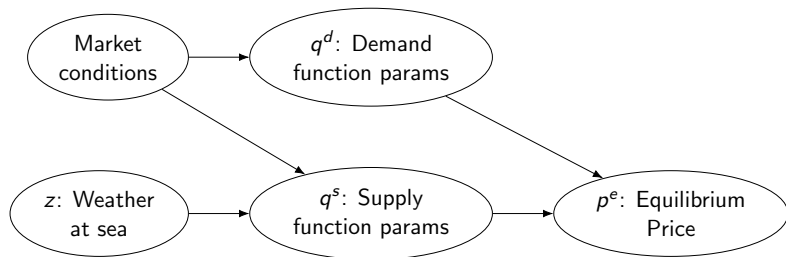
- PO more easily allows one to use causal methods that rely on **non-graphical** assumptions (e.g., shape restrictions). (Imbens 2020; Ackermans 2023)
- Causal graphs are essential for encoding the causal assumptions on which many methods rely. Our causal background information comes in terms of **causal structure**, not counterfactual population statistics. (Ackermans 2023)
- My conclusion: use PO when useful but always combine with graphs for verifying assumptions.

- PO more easily allows one to use causal methods that rely on **non-graphical** assumptions (e.g., shape restrictions). (Imbens 2020; Ackermans 2023)
- Causal graphs are essential for encoding the causal assumptions on which many methods rely. Our causal background information comes in terms of **causal structure**, not counterfactual population statistics. (Ackermans 2023)
- My conclusion: use PO when useful but always combine with graphs for verifying assumptions.
- Imbens' reply: (1) economists are familiar enough with assumptions like unconfoundedness that they don't need graphs (Imbens 2020). (2) PO can be used for estimating parameters of equilibrium models, while graphical models can't: see Angrist, Graddy, and Imbens (2000).

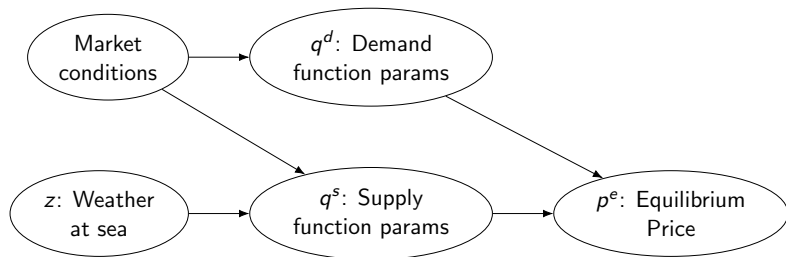




- **Remarkable:** The authors use  $z$  as an instrumental variable to get an estimate of the *demand* elasticity!



- **Remarkable:** The authors use  $z$  as an instrumental variable to get an estimate of the *demand* elasticity!
- This is possible because the study utilises additional theoretical knowledge (shape restrictions)



- **Remarkable:** The authors use  $z$  as an instrumental variable to get an estimate of the *demand* elasticity!
- This is possible because the study utilises additional theoretical knowledge (shape restrictions)
- A central assumption is **unconfoundedness**:  $z \perp\!\!\!\perp (q_{z^*, p^*}^d, q_{z^*, p^*}^s)$  for all counterfactual values  $z^*$  and  $p^*$ .

- **Unconfoundedness:** Take a large hypothetical sample of days with

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),
  - the counterfactual supply functions given potential values of the weather  $z^*$

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),
  - the counterfactual supply functions given potential values of the weather  $z^*$
  - the counterfactual demand functions given each  $z^*$ .

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),
  - the counterfactual supply functions given potential values of the weather  $z^*$
  - the counterfactual demand functions given each  $z^*$ .

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),
  - the counterfactual supply functions given potential values of the weather  $z^*$
  - the counterfactual demand functions given each  $z^*$ .

Observed weather must be independent of the counterfactual supply and demand function for every  $z^*$ .

- **Unconfoundedness:** Take a large hypothetical sample of days with
  - observed weather at sea ( $z$ ),
  - the counterfactual supply functions given potential values of the weather  $z^*$
  - the counterfactual demand functions given each  $z^*$ .

Observed weather must be independent of the counterfactual supply and demand function for every  $z^*$ .

- How do we know this is true? Unconfoundedness refers to unobservable (counterfactual) quantities.

- Scientific observations and causal/counterfactual quantities are **fundamentally different things**. To connect the two we need empirical postulates.

- Scientific observations and causal/counterfactual quantities are **fundamentally different things**. To connect the two we need empirical postulates.
  - Reichenbach's common cause principle
  - The Causal Markov Principle
  - Faithfulness

- Scientific observations and causal/counterfactual quantities are **fundamentally different things**. To connect the two we need empirical postulates.
  - Reichenbach's common cause principle
  - The Causal Markov Principle
  - Faithfulness

## **The Causal Markov Principle:**

Let  $G$  be a DAG describing some population. If

## **The Causal Markov Principle:**

Let  $G$  be a DAG describing some population. If

- $G$  accurately describes the causal structure at work,

## **The Causal Markov Principle:**

Let  $G$  be a DAG describing some population. If

- $G$  accurately describes the causal structure at work,
- Every common cause of two variables in  $G$  is either (a) itself in  $G$  or (b) has the same value for all units in the population,

## The Causal Markov Principle:

Let  $G$  be a DAG describing some population. If

- $G$  accurately describes the causal structure at work,
- Every common cause of two variables in  $G$  is either (a) itself in  $G$  or (b) has the same value for all units in the population,
- $G$  is not otherwise deficient (see later),

## **The Causal Markov Principle:**

Let  $G$  be a DAG describing some population. If

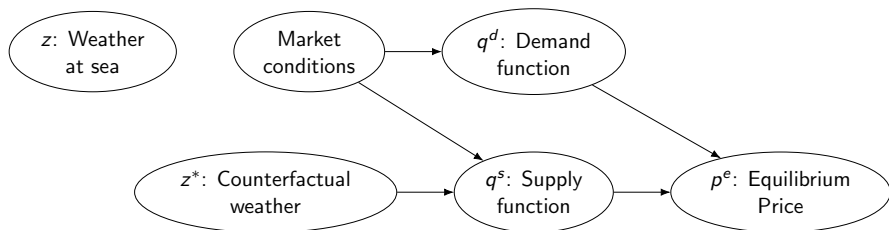
- $G$  accurately describes the causal structure at work,
- Every common cause of two variables in  $G$  is either (a) itself in  $G$  or (b) has the same value for all units in the population,
- $G$  is not otherwise deficient (see later),

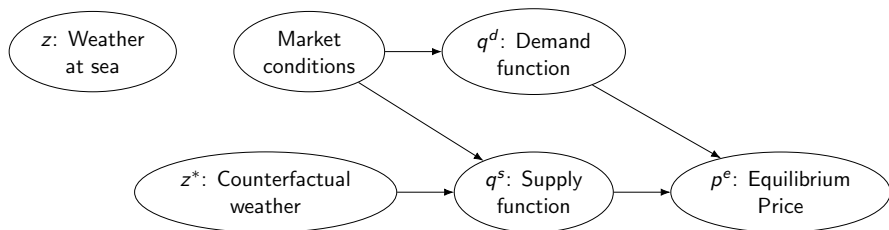
## The Causal Markov Principle:

Let  $G$  be a DAG describing some population. If

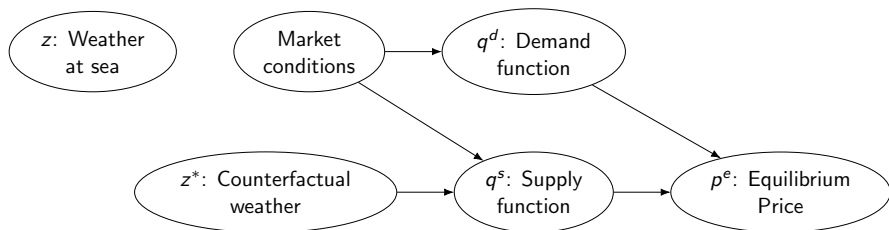
- $G$  accurately describes the causal structure at work,
- Every common cause of two variables in  $G$  is either (a) itself in  $G$  or (b) has the same value for all units in the population,
- $G$  is not otherwise deficient (see later),

then each variable in  $G$  is statistically independent of its nondescendants conditional on its parents. (Spirtes, Glymour, and Scheines 2000)

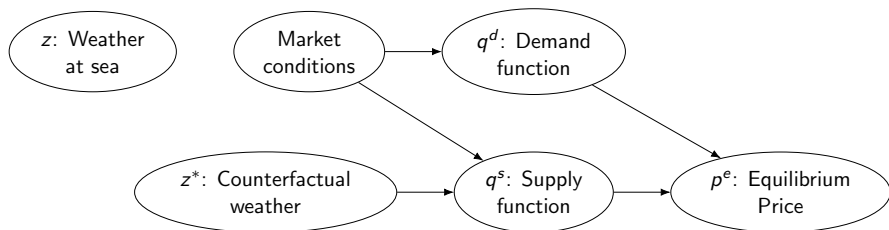




- **Unconfoundedness:**  $z$  is independent of counterfactual supply and demand functions  $q^s$  and  $q^d$  given  $z^*$ .



- **Unconfoundedness:**  $z$  is independent of counterfactual supply and demand functions  $q^s$  and  $q^d$  given  $z^*$ .
- **Causal Markov Principle:** “variables in  $G$  are independent of their nondescendants conditional on their parents”



- **Unconfoundedness:**  $z$  is independent of counterfactual supply and demand functions  $q^s$  and  $q^d$  given  $z^*$ .
- **Causal Markov Principle:** “variables in  $G$  are independent of their nondescendants conditional on their parents”
- Causal Markov Principle  $\rightarrow$  Unconfoundedness

- Steps of causal inference:

- Steps of causal inference:
  1. Gain knowledge about causal structure

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method
    - *Causal Markov Principle* or similar
    - Knowledge from (1)

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method
    - *Causal Markov Principle* or similar
    - Knowledge from (1)
  3. Execute a method to estimate a causal effect size

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method
    - *Causal Markov Principle* or similar
    - Knowledge from (1)
  3. Execute a method to estimate a causal effect size
- The PO framework is very versatile with respect to (3). But has less to say about (1) and (2).

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method
    - *Causal Markov Principle* or similar
    - Knowledge from (1)
  3. Execute a method to estimate a causal effect size
- The PO framework is very versatile with respect to (3). But has less to say about (1) and (2).
- Economists still do steps 1 and 2, but in a more informal and implicit way.

- Steps of causal inference:
  1. Gain knowledge about causal structure
    - *Faithfulness*
    - Experimentation
    - Theory
  2. Assess causal assumptions needed for a method
    - *Causal Markov Principle* or similar
    - Knowledge from (1)
  3. Execute a method to estimate a causal effect size
- The PO framework is very versatile with respect to (3). But has less to say about (1) and (2).
- Economists still do steps 1 and 2, but in a more informal and implicit way.
  - E.g.: “The weather has no common cause with demand, nor causes demand, so they are independent”

- Economists still do steps 1 and 2, but in a more informal and implicit way.

- Economists still do steps 1 and 2, but in a more informal and implicit way.
- This is problematic:

- Economists still do steps 1 and 2, but in a more informal and implicit way.
- This is problematic:
  1. Informal reasoning is bound to go wrong

- Economists still do steps 1 and 2, but in a more informal and implicit way.
- This is problematic:
  1. Informal reasoning is bound to go wrong
  2. There are various pitfalls of the *Causal Markov Principle* and similar postulates

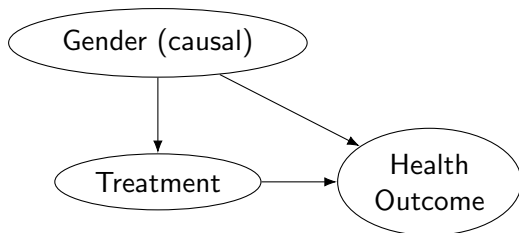
- Economists still do steps 1 and 2, but in a more informal and implicit way.
- This is problematic:
  1. Informal reasoning is bound to go wrong
  2. There are various pitfalls of the *Causal Markov Principle* and similar postulates
  3. Economists miss out on a large literature about empirical postulates and their application

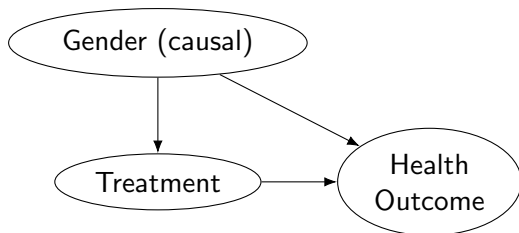
- Economists still do steps 1 and 2, but in a more informal and implicit way.
- This is problematic:
  1. Informal reasoning is bound to go wrong
  2. There are various pitfalls of the *Causal Markov Principle* and similar postulates
  3. Economists miss out on a large literature about empirical postulates and their application
- Combining PO with the graphical framework would solve this

- **Imbens:** PO has a useful distinction between *causal variables* and *non-causal attributes* which the graphical framework lacks

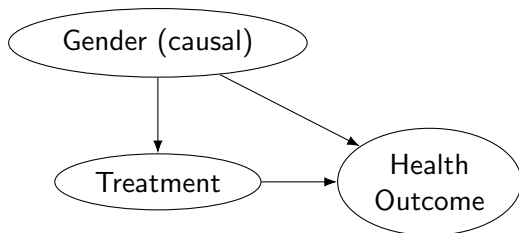
- **Imbens:** PO has a useful distinction between *causal variables* and *non-causal attributes* which the graphical framework lacks
- A **causal variable** has clearly defined manipulations, a **non-causal attribute** does not

- **Imbens:** PO has a useful distinction between *causal variables* and *non-causal attributes* which the graphical framework lacks
- A **causal variable** has clearly defined manipulations, a **non-causal attribute** does not
- **Markov pitfall:** To apply the Causal Markov Principle, manipulations on all variables need to be defined and non-ambiguous (Spirtes and Scheines 2004)

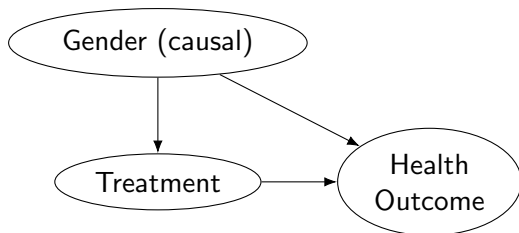




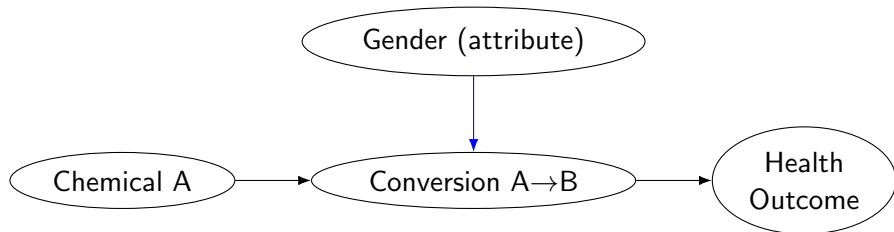
- One should condition on *Gender* to adjust for confounding

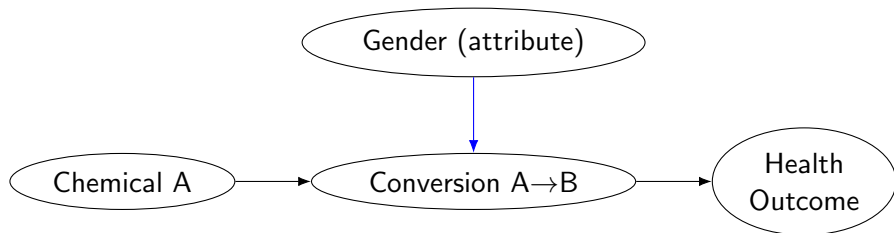


- One should condition on *Gender* to adjust for confounding
- **Causal Markov**: we know this works *only* if *Gender* is a causal variable!

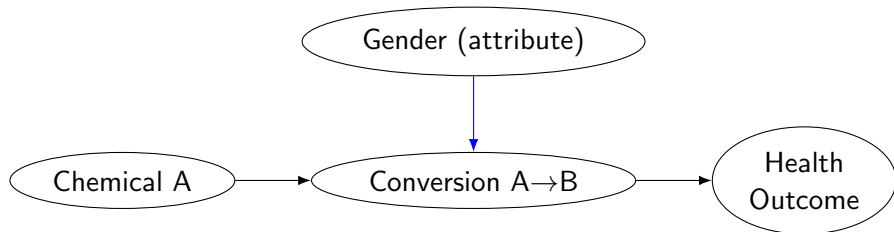


- One should condition on *Gender* to adjust for confounding
- **Causal Markov**: we know this works *only* if *Gender* is a causal variable!
- This might be overlooked in PO studies in which *Gender* is an attribute

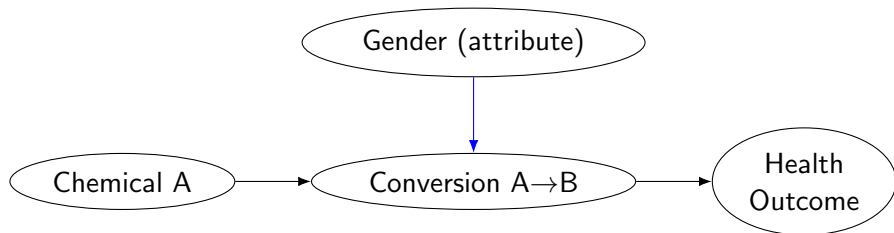




- Since there is no confounding,  
 $E(\text{Outcome}|A) - E(\text{Outcome}|\text{no } A)$   
is the average causal effect of *chemical A* on *Outcome*.



- Since there is no confounding,  
 $E(\text{Outcome}|A) - E(\text{Outcome}|\text{no } A)$   
is the average causal effect of *chemical A* on *Outcome*.
- There might still be non-causal bias: The average causal effect is uninformative if the sign of the causal effect is different for men and women



- Since there is no confounding,  
 $E(\text{Outcome}|A) - E(\text{Outcome}|\text{no } A)$   
is the average causal effect of *chemical A* on *Outcome*.
- There might still be non-causal bias: The average causal effect is uninformative if the sign of the causal effect is different for men and women
- The graphical framework stops at causality; PO is concerned with additional problems.

- Summarising,

- Summarising,
  1. Causal methods work only if all variables are causal
  2. *Non-causal* biases may require conditioning on non-causal attributes

- Summarising,
  1. Causal methods work only if all variables are causal
  2. *Non-causal* biases may require conditioning on non-causal attributes
- The graphical framework focuses only on (1); the PO framework on (1) and (2) combined

- Summarising,
  1. Causal methods work only if all variables are causal
  2. *Non-causal* biases may require conditioning on non-causal attributes
- The graphical framework focuses only on (1); the PO framework on (1) and (2) combined
- Use of PO exclusively runs the risk of mistakenly conditioning on non-causal attributes to solve causal problems.

- Econometrics has improved greatly due to better causal methodology and use of the PO framework.
- To better assess assumptions of causal methods, economists should incorporate causal graphs.
- Imbens' arguments against graphical models fail.

- Ackermans, Lennart B. 2023. “Reflections on the 2021 Nobel Memorial Prize Awarded to David Card, Joshua Angrist, and Guido Imbens.” *Erasmus Journal for Philosophy and Economics* 16, no. 1 (July): 77–96.
- Angrist, Joshua D. 1990. “Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records.” *The American Economic Review* 80 (3): 313–336.
- Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens. 2000. “The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish.” *The Review of Economic Studies* 67, no. 3 (July): 499–527.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables.” *Journal of the American Statistical Association* 91 (434): 444–455.

- Angrist, Joshua D., and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106, no. 4 (November): 979–1014.
- Imbens, Guido W. 2020. "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics." *Journal of Economic Literature* 58, no. 4 (December): 1129–1179.
- Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–475.
- Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." *The American Economic Review* 73 (1): 31–43.

- Pearl, Judea. 2000. *Causality*. 1st ed. Cambridge: Cambridge University Press.
- Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology* 66 (5): 688–701.
- Spirtes, Peter, Clark Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press.
- Spirtes, Peter, and Richard Scheines. 2004. "Causal inference of ambiguous manipulations." *Philosophy of Science* 71 (5): 833–845.