Setting the Target

Precise Estimands and the Gap Between Theory and Empirics

Ian Lundberg
Princeton Sociology
ilundberg@princeton.edu

Rebecca Johnson
Dartmouth Quantitative Social Science
rebecca.ann.johnson@dartmouth.edu

Brandon M. Stewart
Princeton Sociology
bms4@princeton.edu

4 August 2020. Annual Meeting of the American Sociological Association.
Paper on SocArxiv [link]. Replication code on GitHub [link]. Slides on GitHub [link].
Research reported in this publication was supported by The Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P2CHD047879.

State of the field: How methods are used in sociology

State of the field: How methods are used in sociology

"[Variables] empirically perform as theoretically predicted,

Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

State of the field: How methods are used in sociology

"[Variables] empirically perform as theoretically predicted,
by displaying statistically significant
effects net of other variables in the right direction"

Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

↓

The goal is only defined within the statistical model

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

↓

The goal is only defined within the statistical model

↓

It becomes impossible to reason about other estimation strategies

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

↓

The goal is only defined within the statistical model

↓

It becomes impossible to reason about other estimation strategies

Solution:

State the research goal
separately from the estimation strategy

The **target tautology**:

Research goals are defined by hypotheses about model coefficients

↓

The goal is only defined within the statistical model

↓

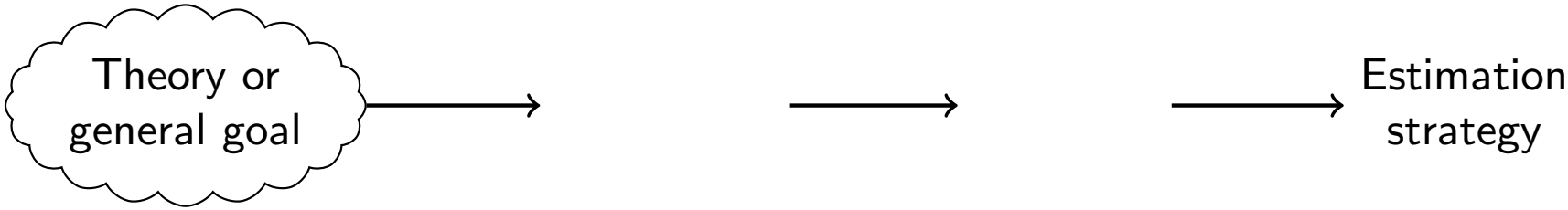
It becomes impossible to reason about other estimation strategies

Solution:

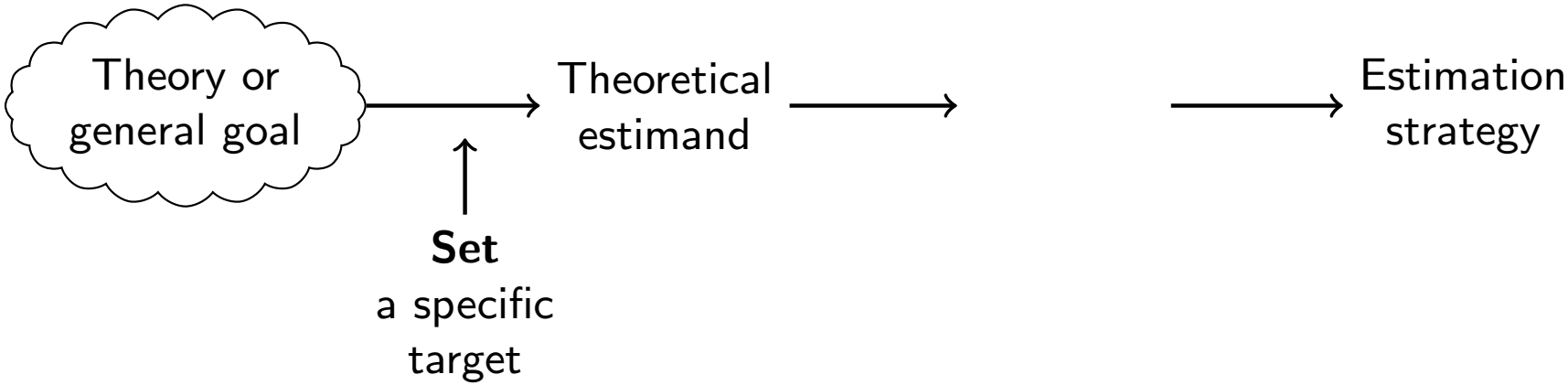
Our **diagnosis** for the source of many methodological problems

State the research goal separately from the estimation strategy

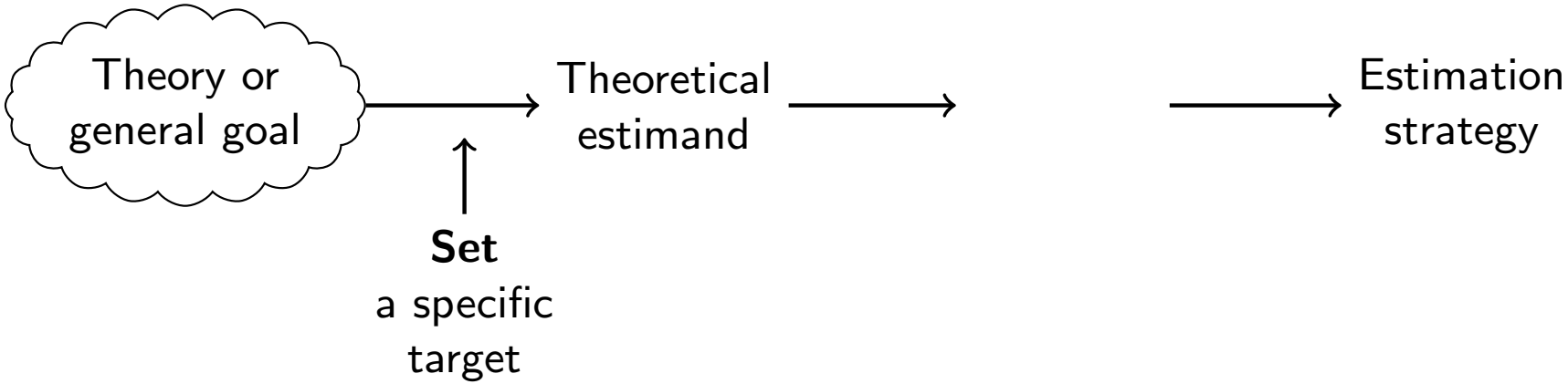
Our solution: A framework centered on **estimands**



Our solution: A framework centered on **estimands**



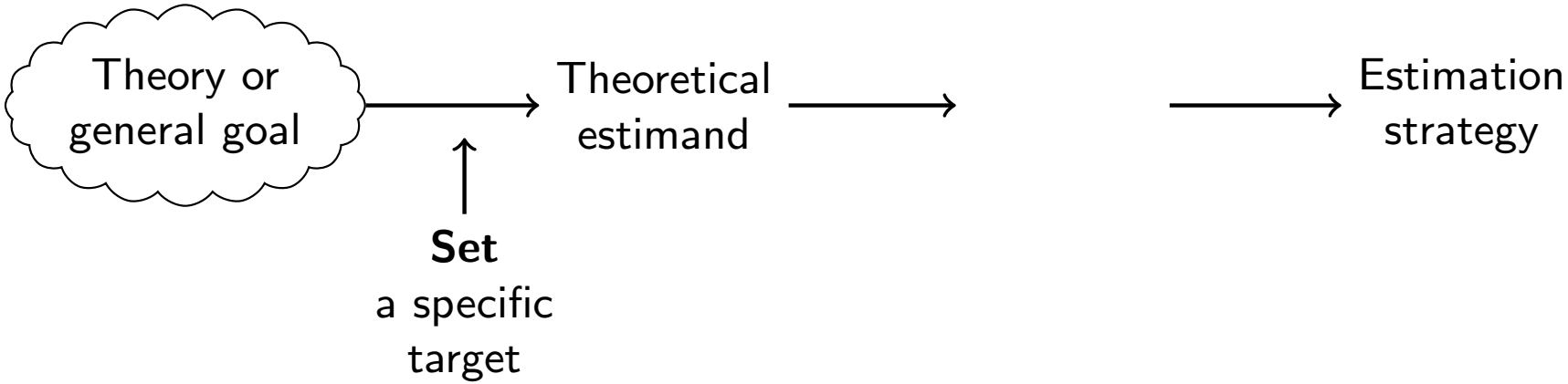
Our solution: A framework centered on **estimands**



Definition

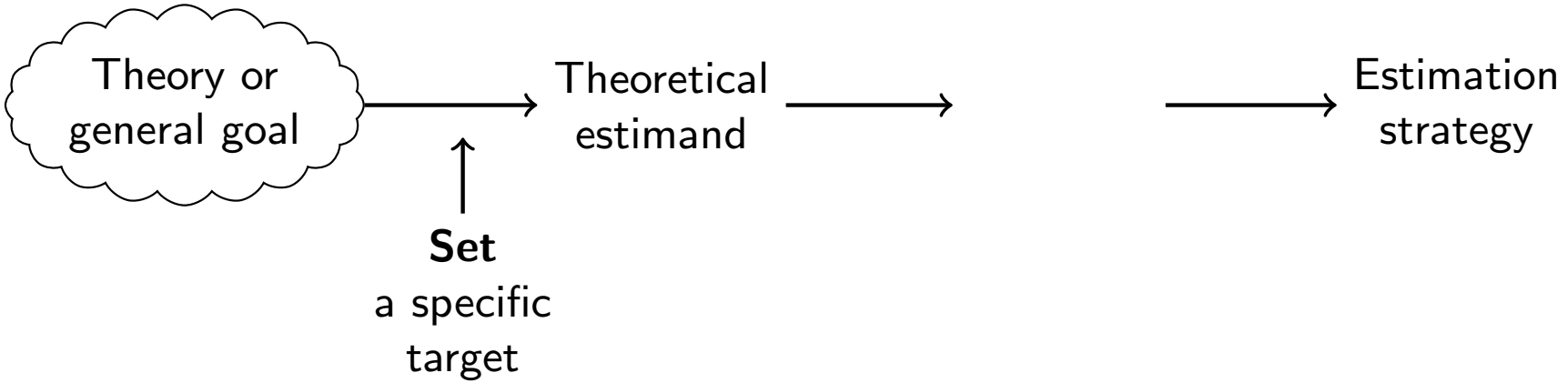
A **unit-specific quantity** aggregated over a **target population**

Our solution: A framework centered on **estimands**



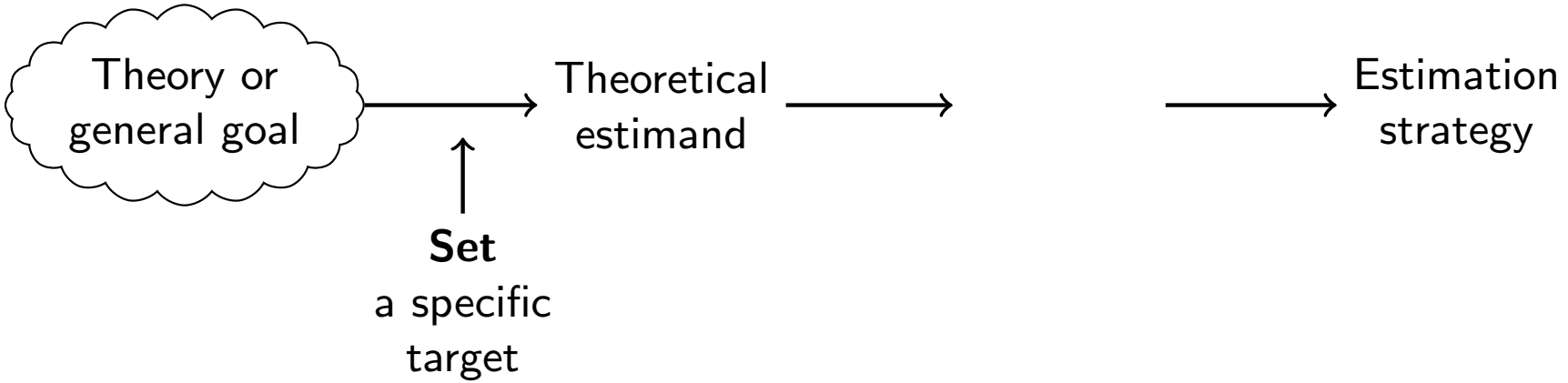
Definition	Example
A **unit-specific quantity** aggregated over a **target population**	$\frac{1}{\text{Size of U.S. adult population}} \sum_{i \text{ in U.S. adult population}} \left(\text{Employed}_i\right)$

Our solution: A framework centered on **estimands**



Definition	Example
A **unit-specific quantity** aggregated over a **target population**	$\frac{1}{\text{Size of U.S. adult population}} \sum_{i \text{ in U.S. adult population}} \Big(\underbrace{\text{Employed}_i(\text{Job training})}_{\text{Employment if received job training}} - \underbrace{\text{Employed}_i(\text{No job training})}_{\text{Employment if did not receive job training}} \Big)$

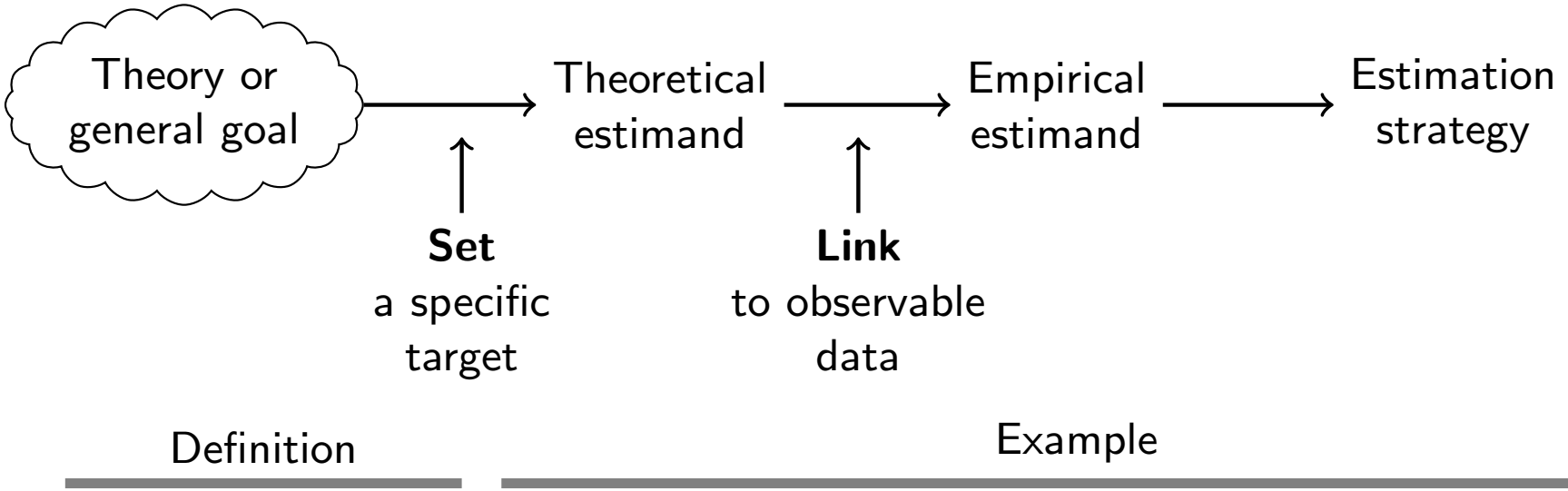
Our solution: A framework centered on **estimands**



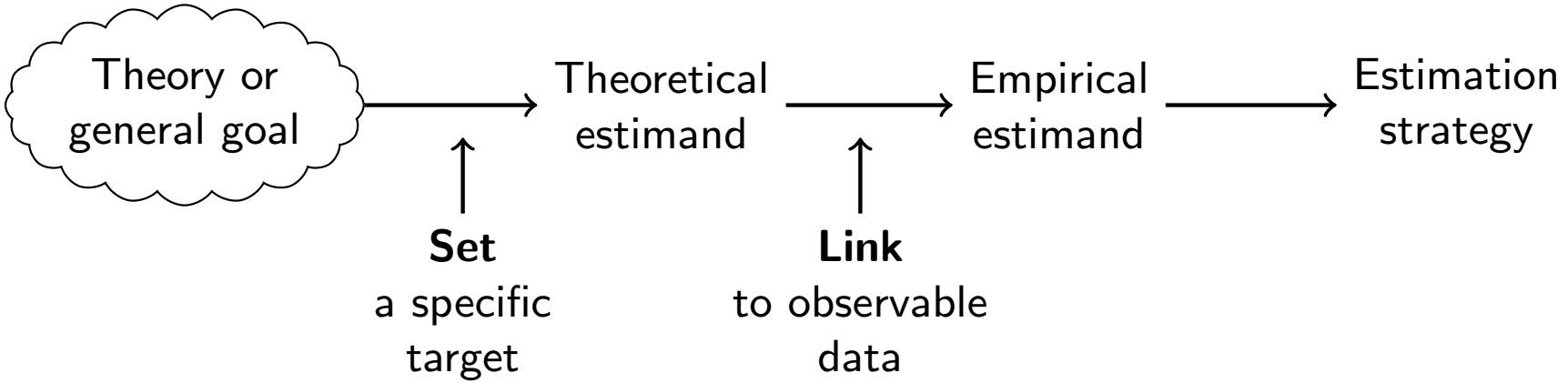
Definition	Example
A **unit-specific quantity** aggregated over a **target population**	$\frac{1}{\text{Size of U.S. adult population}} \sum_{i \text{ in U.S. adult population}} \Big(\underbrace{\text{Employed}_i(\text{Job training})}_{\text{Employment if received job training}} - \underbrace{\text{Employed}_i(\text{No job training})}_{\text{Employment if did not receive job training}} \Big)$

Lieberson 1987, Abbott 1988, Freedman 1991, Xie 2013, Hernán 2018

Our solution: A framework centered on **estimands**

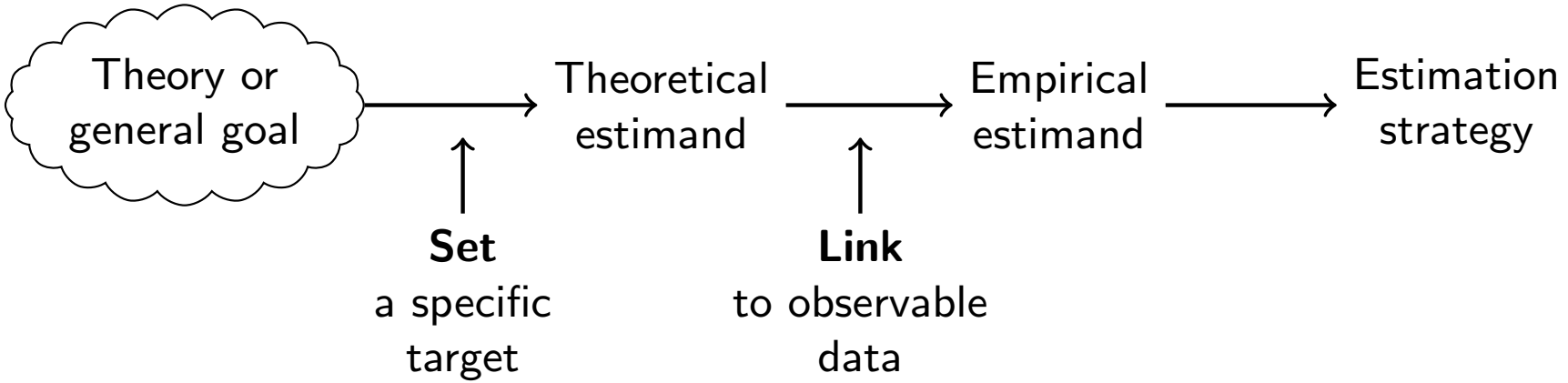


Our solution: A framework centered on **estimands**

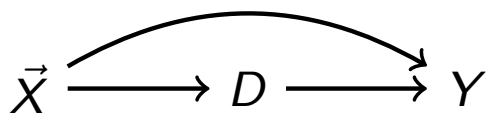


Definition	Example
A quantity involving **observable data**	

Our solution: A framework centered on **estimands**

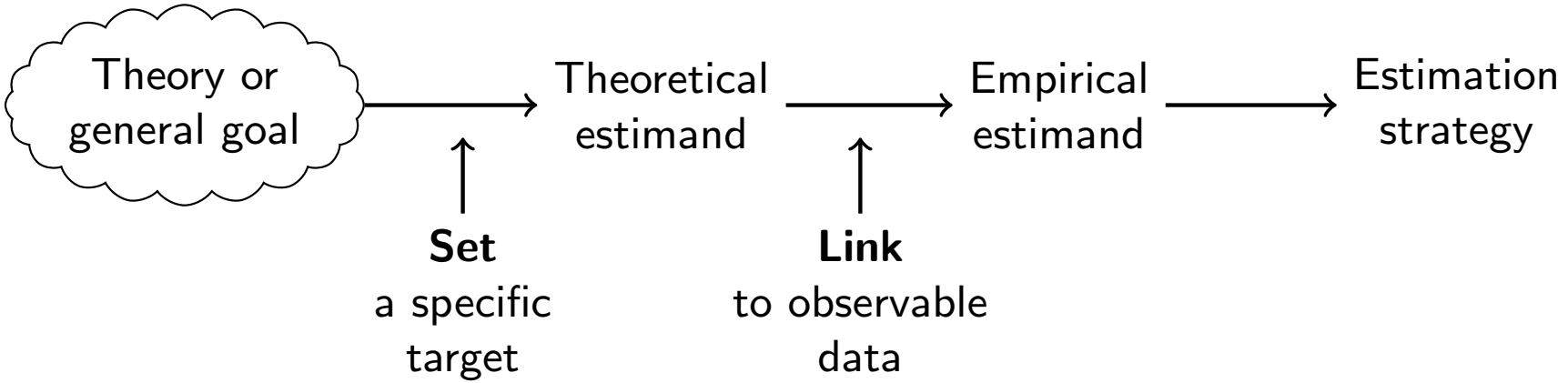


Definition	Example
A quantity involving **observable data**	$\theta = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\text{Sample average}} \left(\underbrace{\mathrm{E}(Y \mid \vec{X} = \vec{x}_i, D = 1)}_{\text{Expected outcome among those treated}} - \underbrace{\mathrm{E}(Y \mid \vec{X} = \vec{x}_i, D = 0)}_{\text{Expected outcome among those untreated}} \right)$

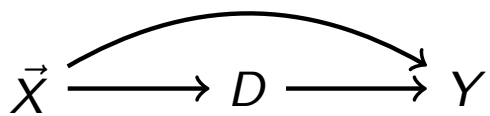


Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

Our solution: A framework centered on **estimands**

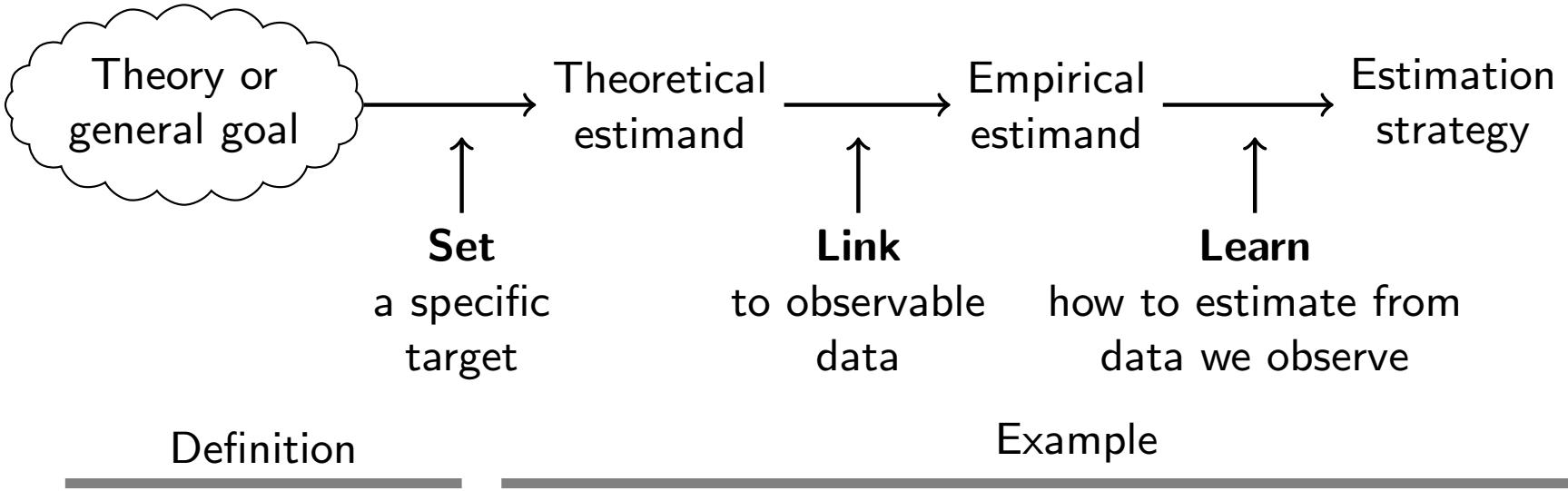


Definition	Example
A quantity involving **observable data**	$\theta = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\text{Sample average}} \left(\underbrace{\mathrm{E}(Y \mid \vec{X} = \vec{x}_i, D = 1)}_{\text{Expected outcome among those treated}} - \underbrace{\mathrm{E}(Y \mid \vec{X} = \vec{x}_i, D = 0)}_{\text{Expected outcome among those untreated}} \right)$

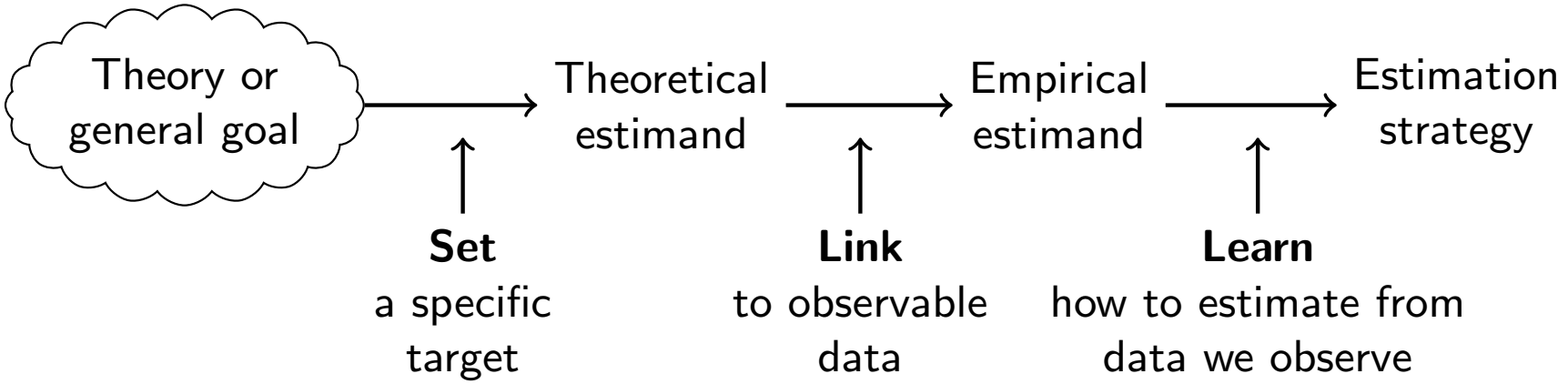


Pearl 2009, Imbens and Rubin 2015,
Morgan and Winship 2015, Elwert and Winship 2014

Our solution: A framework centered on **estimands**

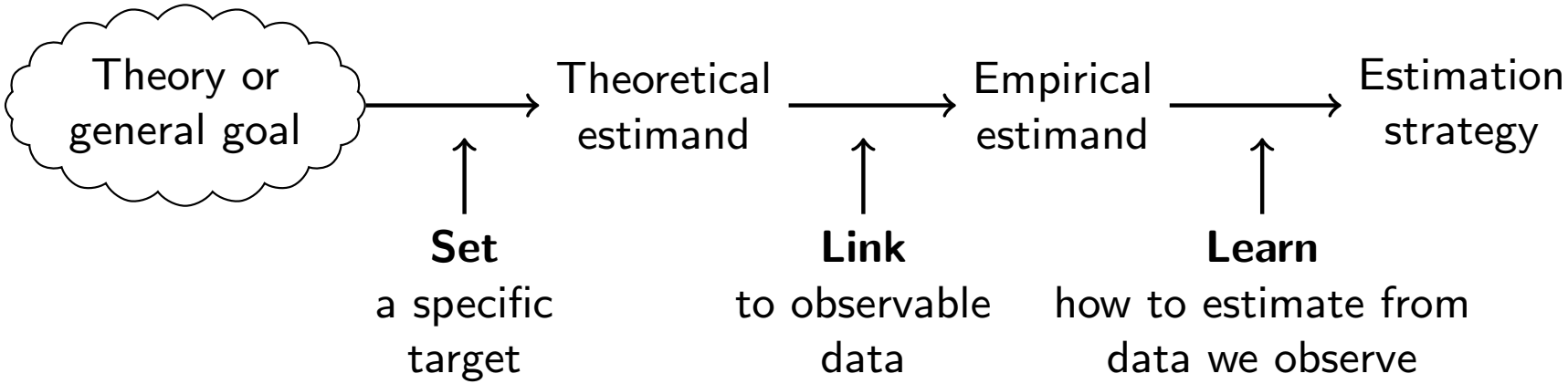


Our solution: A framework centered on **estimands**



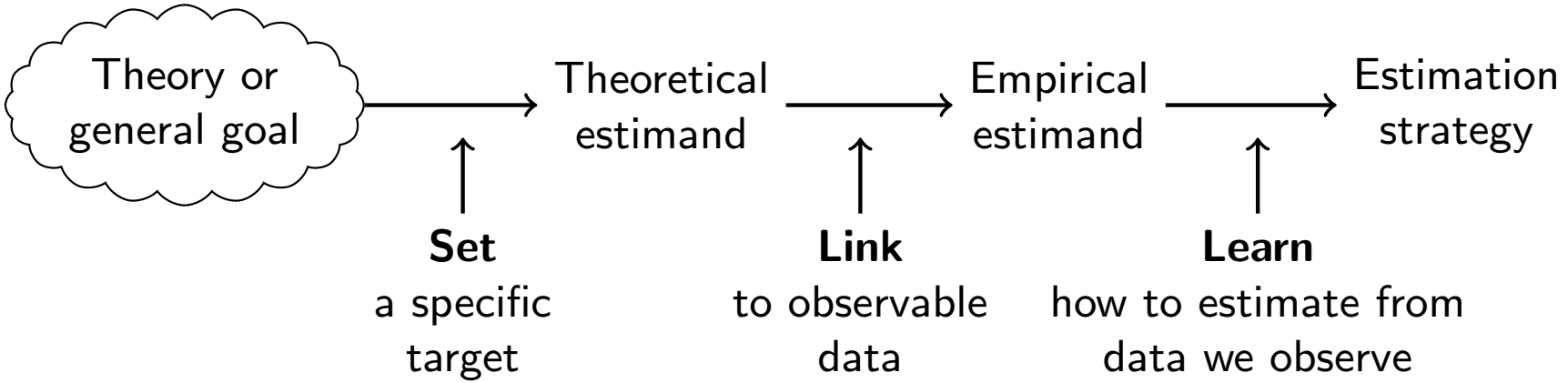
Definition	Example
An algorithm applied to data	

Our solution: A framework centered on **estimands**



Definition	Example
An algorithm applied to data	$\hat{\theta} = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\text{Sample average}} \left(\underbrace{\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 1)}_{\text{Regression prediction if treated}} - \underbrace{\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 0)}_{\text{Regression prediction if untreated}} \right)$

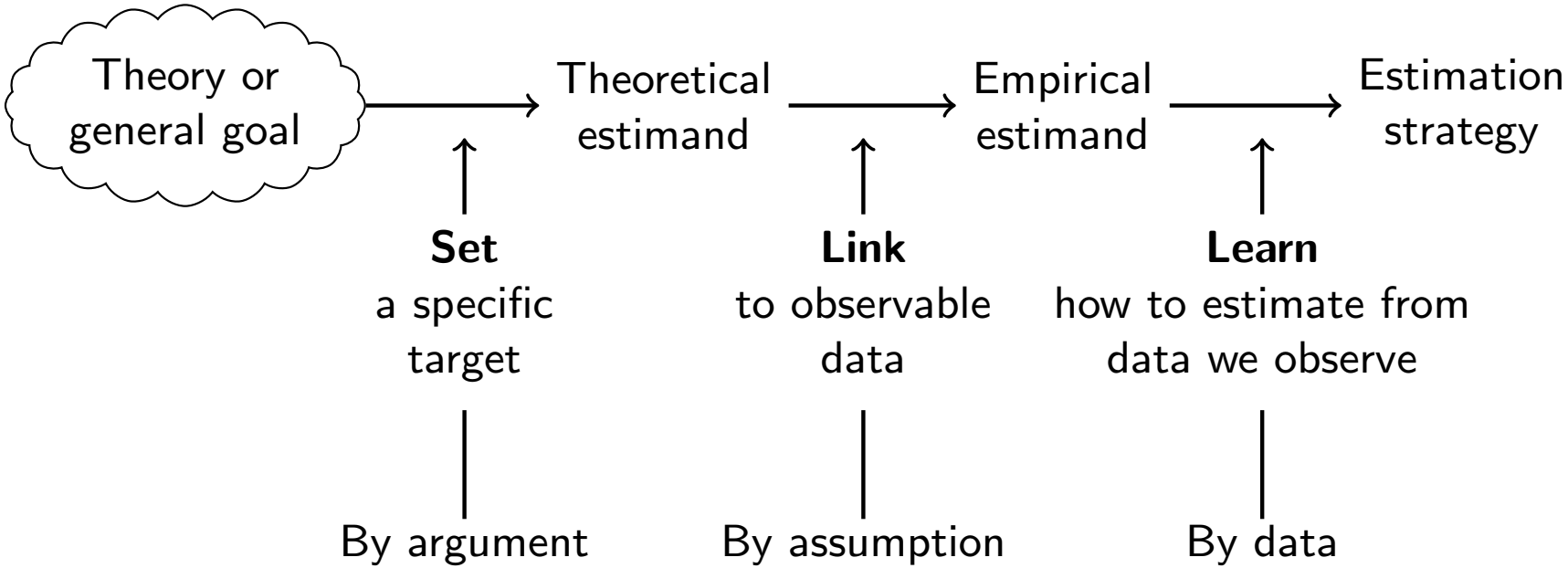
Our solution: A framework centered on **estimands**

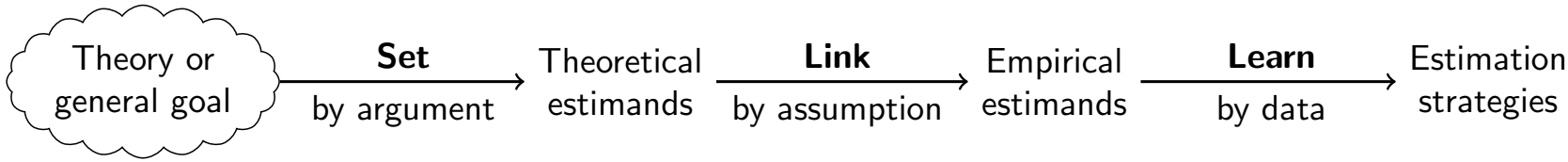


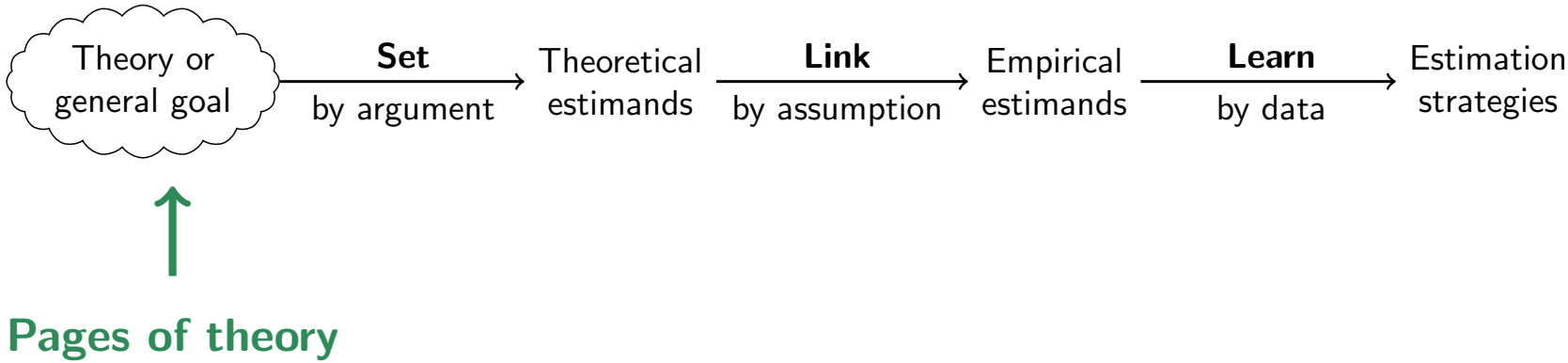
Definition	Example
An algorithm applied to data	$\hat{\theta} = \underbrace{\frac{1}{n}\sum_{i=1}^{n}}_{\text{Sample average}} \left(\underbrace{\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 1)}_{\text{Regression prediction if treated}} - \underbrace{\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, D = 0)}_{\text{Regression prediction if untreated}} \right)$

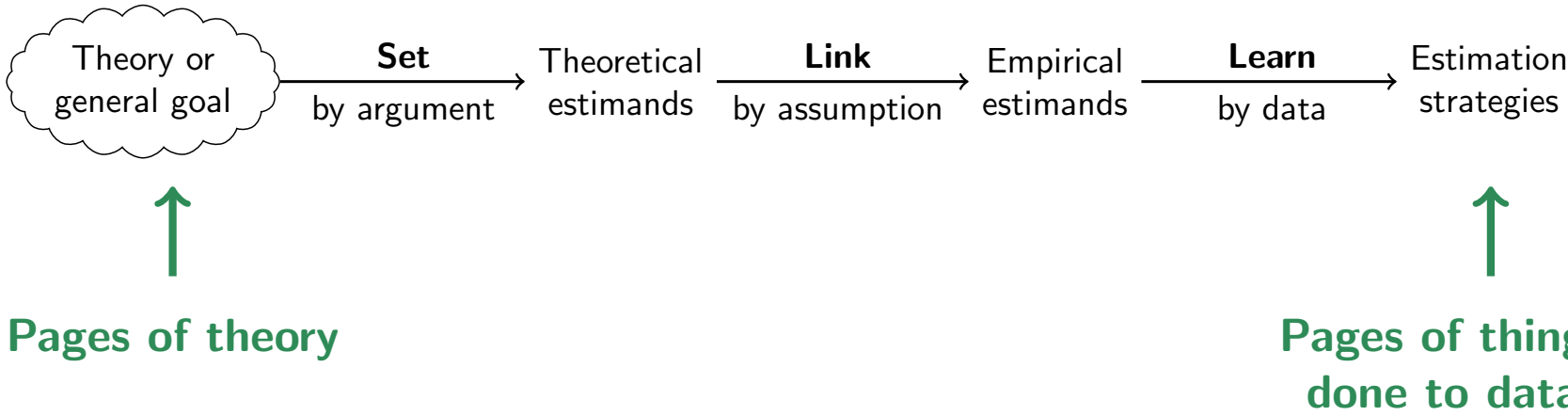
Young 2009, Watts 2014, Berk et al. 2019, Molina and Garip 2019

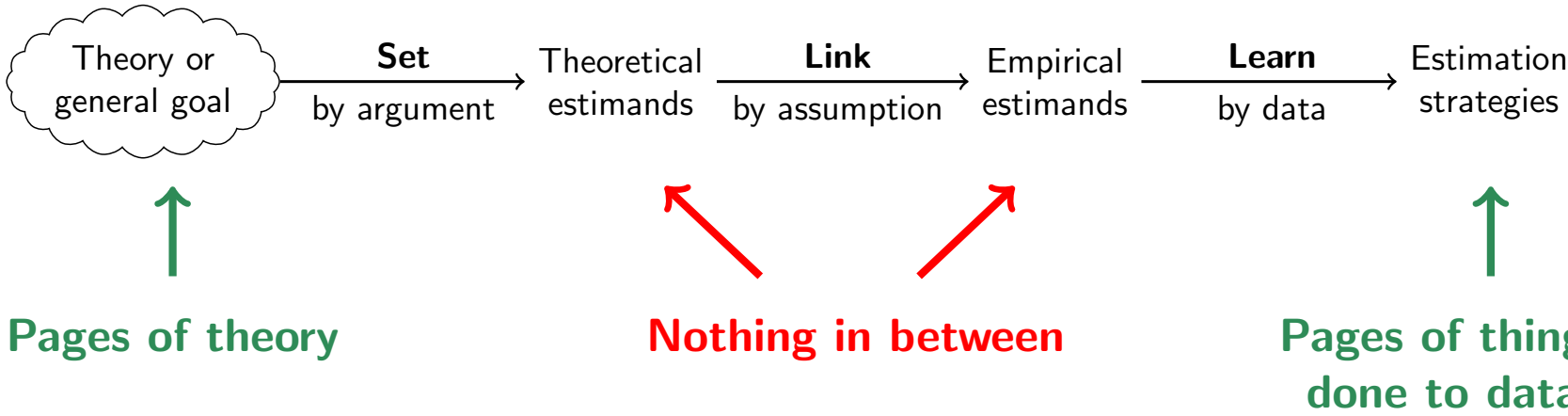
Our solution: A framework centered on **estimands**




Theory or general goal
Set
by argument
Theoretical estimands
Link
by assumption
Empirical estimands
Learn
by data
Estimation strategies


Theory or general goal
Set
by argument
Theoretical estimands
Link
by assumption
Empirical estimands
Learn
by data
Estimation strategies
Pages of theory


Theory or general goal
Set
by argument
Theoretical estimands
Link
by assumption
Empirical estimands
Learn
by data
Estimation strategies
Pages of theory
Pages of things done to data



Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

Systematic Review

All 32 articles in ASR 2018 using quantitative data

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

Clarity about the **target population**

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles | 56% of *ASR* → Causal target: Observational data

7 articles | 22% of *ASR* → Causal target: Experiments

7 articles | 22% of *ASR* → Descriptive target

Clarity about the **target population**

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles | 56% of *ASR*

Causal target: Observational data

0 out of 18 articles was clear

7 articles | 22% of *ASR*

Causal target: Experiments

7 articles | 22% of *ASR*

Descriptive target

Clarity about the **target population**

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles | 56% of *ASR*

Causal target: Observational data

0 out of 18 articles was clear

7 articles | 22% of *ASR*

Causal target: Experiments

4 out of 7 articles were clear

7 articles | 22% of *ASR*

Descriptive target

Clarity about the **target population**

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

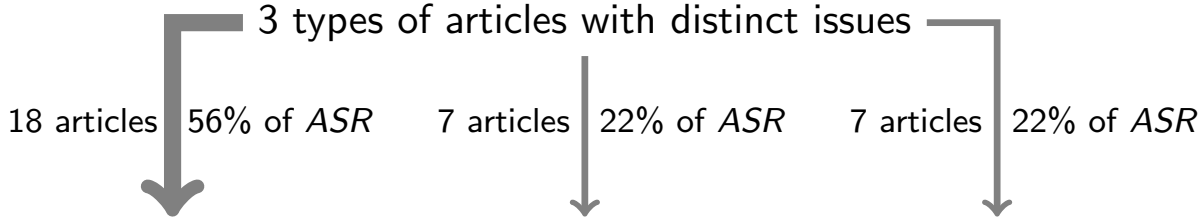
3 types of articles with distinct issues

18 articles (56% of *ASR*)	7 articles (22% of *ASR*)	7 articles (22% of *ASR*)
Causal target: Observational data	Causal target: Experiments	Descriptive target
0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Clarity about the **target population**

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

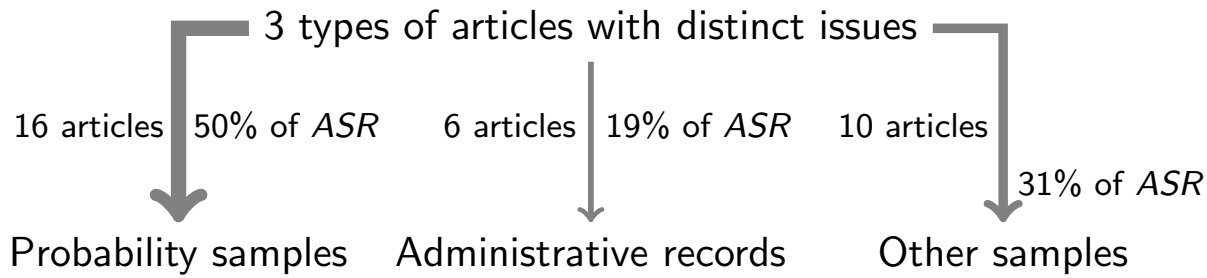


0 out of 18 articles was clear

4 out of 7 articles were clear

1 out of 7 articles was clear

Clarity about the **target population**



Systematic Review

All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles	7 articles	7 articles
56% of *ASR*	22% of *ASR*	22% of *ASR*
Causal target: Observational data	Causal target: Experiments	Descriptive target
0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Clarity about the **target population**

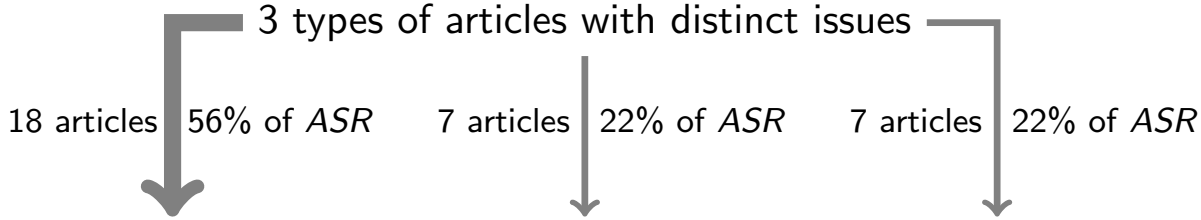
3 types of articles with distinct issues

16 articles	6 articles	10 articles
50% of *ASR*	19% of *ASR*	31% of *ASR*
Probability samples	Administrative records	Other samples
6 out of 16 articles were clear		

Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics

Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

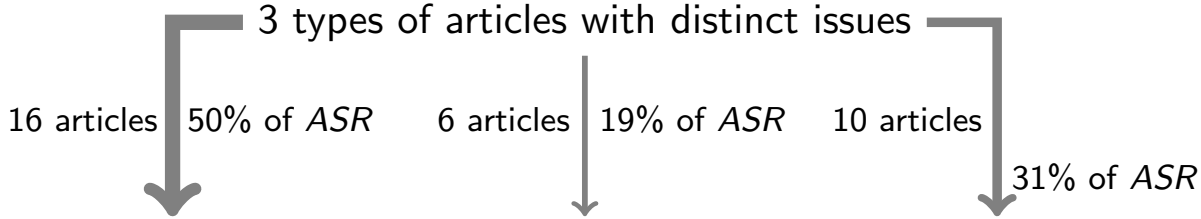


0 out of 18 articles was clear

4 out of 7 articles were clear

1 out of 7 articles was clear

Clarity about the **target population**

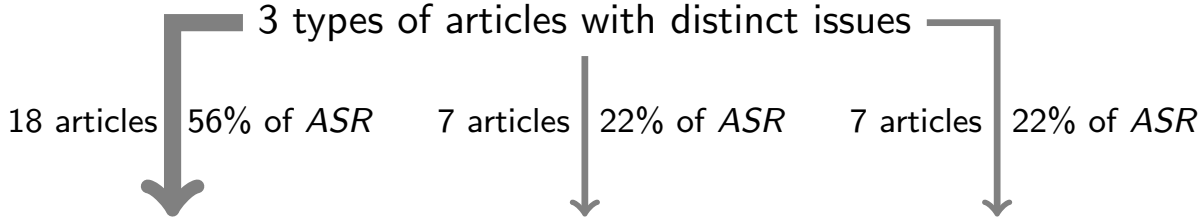


6 out of 16 articles were clear

1 out of 6 articles was clear

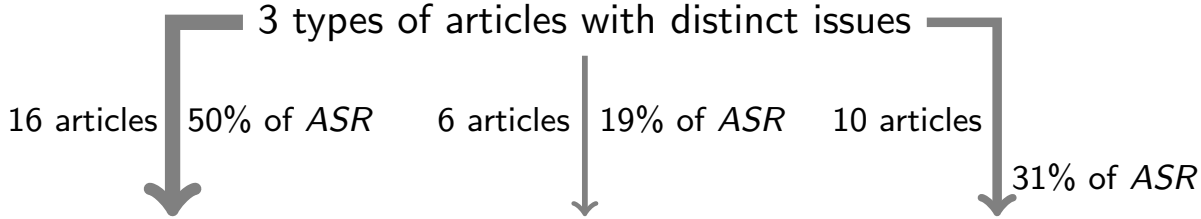
Systematic Review
All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)



0 out of 18	**4 out of 7**	**1 out of 7**
articles was clear	articles were clear	articles was clear

Clarity about the **target population**



6 out of 16	**1 out of 6**	**2 out of 10**
articles were clear	articles was clear	articles were clear

Systematic Review

All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles	7 articles	7 articles
56% of *ASR*	22% of *ASR*	22% of *ASR*
Causal target: Observational data	Causal target: Experiments	Descriptive target

Consequences

0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Clarity about the **target population**

3 types of articles with distinct issues

16 articles	6 articles	10 articles
50% of *ASR*	19% of *ASR*	31% of *ASR*
Probability samples	Administrative records	Other samples

6 out of 16 articles were clear	**1 out of 6** articles was clear	**2 out of 10** articles were clear

Systematic Review

All 32 articles in ASR 2018 using quantitative data

Consequences

Confusion about theory

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles	7 articles	7 articles
56% of *ASR*	22% of *ASR*	22% of *ASR*
Causal target: Observational data	Causal target: Experiments	Descriptive target
0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Clarity about the **target population**

3 types of articles with distinct issues

16 articles	6 articles	10 articles
50% of *ASR*	19% of *ASR*	31% of *ASR*
Probability samples	Administrative records	Other samples
6 out of 16 articles were clear	**1 out of 6** articles was clear	**2 out of 10** articles were clear

Systematic Review

All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles	7 articles	7 articles
56% of *ASR*	22% of *ASR*	22% of *ASR*
Causal target: Observational data	Causal target: Experiments	Descriptive target
0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Consequences

Confusion about theory

Confusion about methods

Clarity about the **target population**

3 types of articles with distinct issues

16 articles	6 articles	10 articles
50% of *ASR*	19% of *ASR*	31% of *ASR*
Probability samples	Administrative records	Other samples
6 out of 16 articles were clear	**1 out of 6** articles was clear	**2 out of 10** articles were clear

Systematic Review

All 32 articles in ASR 2018 using quantitative data

Clarity about **causal claims** (if any)

3 types of articles with distinct issues

18 articles	7 articles	7 articles
56% of *ASR*	22% of *ASR*	22% of *ASR*
Causal target: Observational data	Causal target: Experiments	Descriptive target
0 out of 18 articles was clear	**4 out of 7** articles were clear	**1 out of 7** articles was clear

Consequences

Confusion about theory

Clarity about the **target population**

3 types of articles with distinct issues

16 articles	6 articles	10 articles
50% of *ASR*	19% of *ASR*	31% of *ASR*
Probability samples	Administrative records	Other samples
6 out of 16 articles were clear	**1 out of 6** articles was clear	**2 out of 10** articles were clear

Confusion about methods

Confusion about conflicting studies

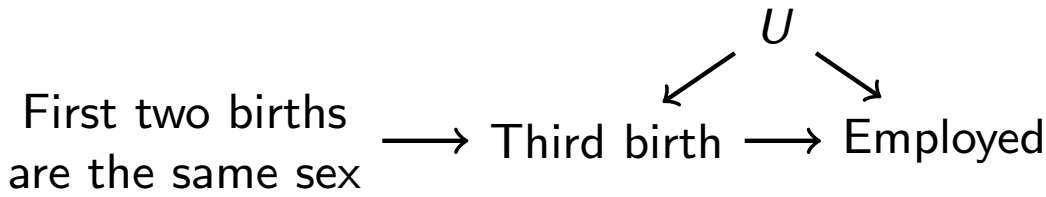
Angrist and Evans 1998

First two births
are the same sex

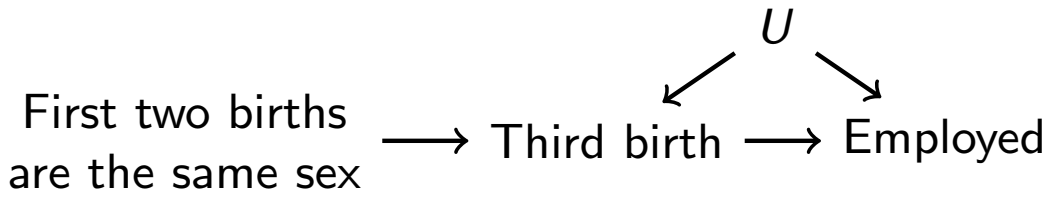
Angrist and Evans 1998

First two births are the same sex $\longrightarrow$ Third birth

Angrist and Evans 1998

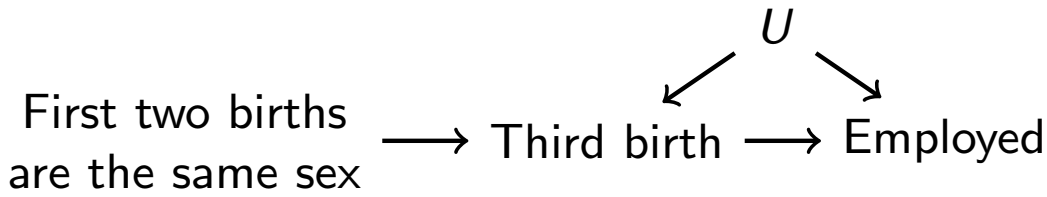


Angrist and Evans 1998



Target population is limited:

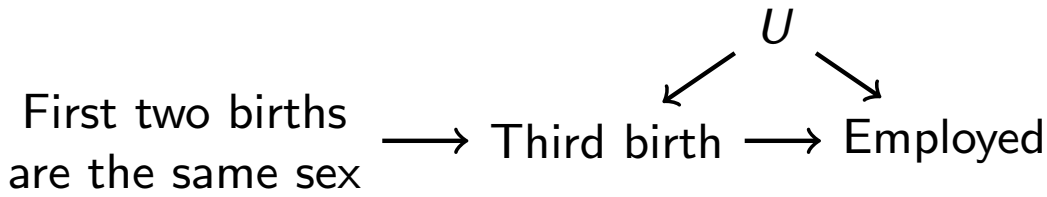
Angrist and Evans 1998



Target population is limited:

— At most 53% of mothers have 2+ children

Angrist and Evans 1998

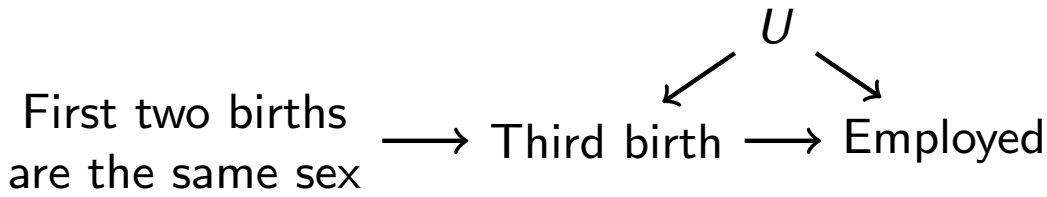


Target population is limited:

— At most 53% of mothers have 2+ children

— The complier population is at most 7%

Angrist and Evans 1998



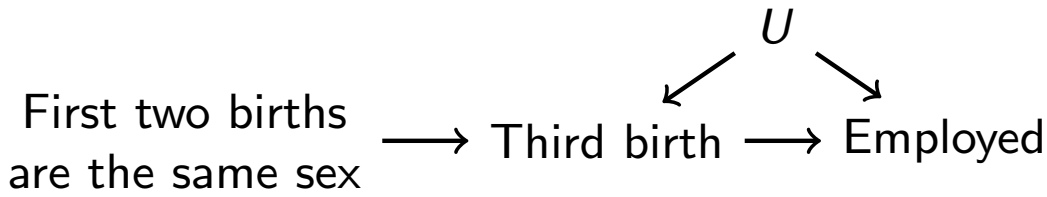
Target population is limited:

— At most 53% of mothers have 2+ children

— The complier population is at most 7%

— Target population: at most 53% $\times$ 7% $=$ 4% of mothers

Angrist and Evans 1998



Target population is limited:

— At most 53% of mothers have 2+ children

— The complier population is at most 7%

— Target population: at most 53% $\times$ 7% $=$ 4% of mothers

Causal contrast is limited:

— Having 3 vs. 2 children

Angrist and Evans 1998

Vague statement

Effect of motherhood on employment

Vague statement

Effect of motherhood on employment

Precise statement

Vague statement

Effect of motherhood on employment

Precise statement

Effect of having **3 vs. 2 children**

causal contrast

Vague statement

Effect of motherhood on employment

Precise statement

Effect of having **3 vs. 2 children**

among those with at least two children who would have a third birth if and only if the first two were of the same sex

causal contrast

target population

Vague statement

Effect of motherhood on employment

Precise statement

Effect of having **3 vs. 2 children**

among those with at least two children who would have a third birth if and only if the first two were of the same sex

causal contrast

target population

Not just problems with instruments.

Vague statement

Effect of motherhood on employment

Precise statement

causal contrast

Effect of having **3 vs. 2 children**

among those with at least two children who would have a third birth if and only if the first two were of the same sex

target population

Not just problems with instruments.

— Causal contrast is vague when the treatment is continuous

Vague statement

Effect of motherhood on employment

Precise statement

Effect of having **3 vs. 2 children**

among those with at least two children who would have a third birth if and only if the first two were of the same sex

causal contrast

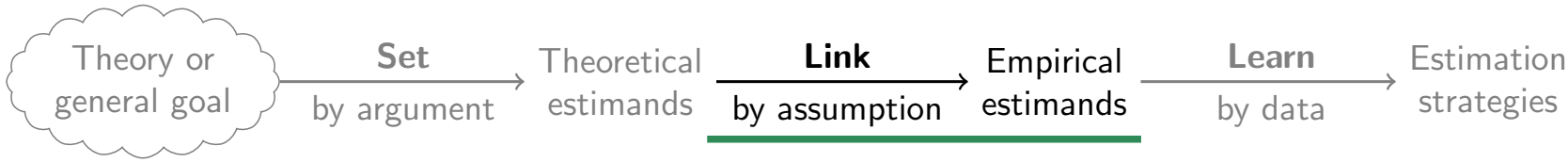
target population

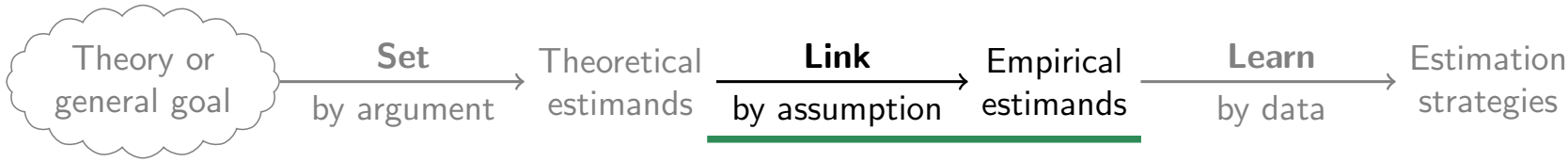
Not just problems with instruments.

— Causal contrast is vague when the treatment is continuous

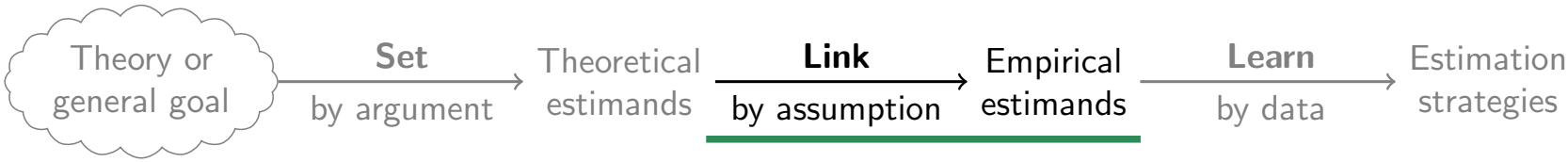
— Target population is vague if there are issues of common support


Theory or general goal
Set
by argument
Theoretical estimands
Link
by assumption
Empirical estimands
Learn
by data
Estimation strategies


Theory or general goal
Set
by argument
Theoretical estimands
Link
by assumption
Empirical estimands
Learn
by data
Estimation strategies

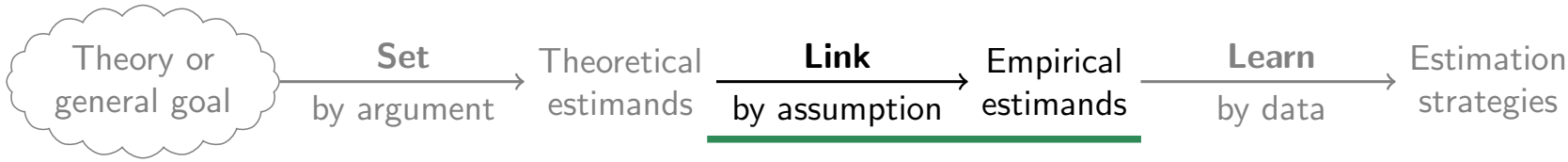


Identification assumptions

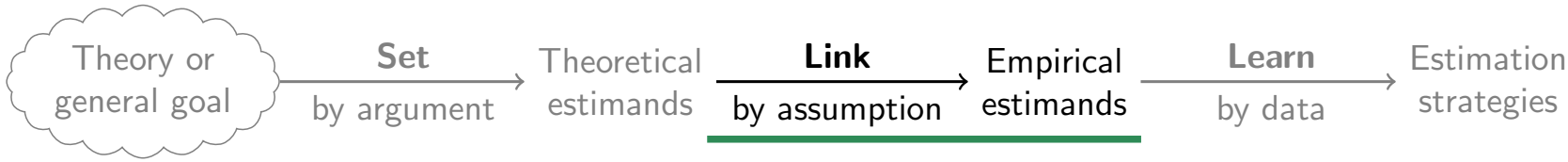


Identification assumptions

are essential even for claims that appear to be **non-causal**

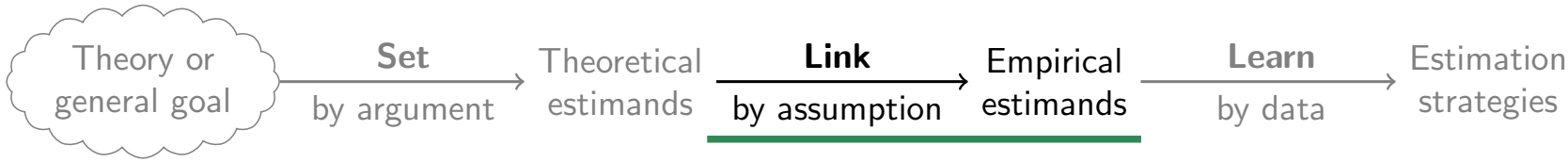


Fryer 2019. See a fuller critique by Knox et al. 2020.



Evidence: Police use lethal force at the same rate against black and white civilians who are stopped.

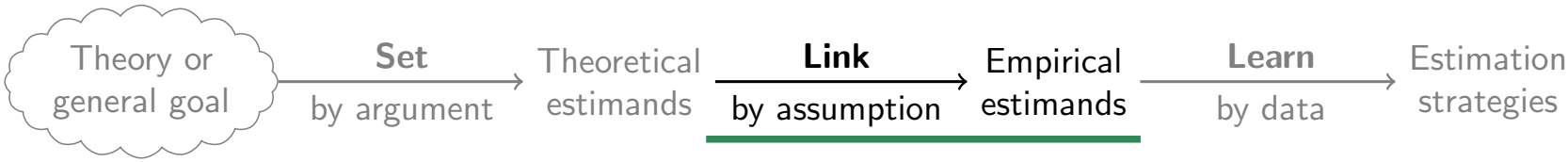
Fryer 2019. See a fuller critique by Knox et al. 2020.



Evidence: Police use lethal force at the same rate against black and white civilians who are stopped.

Claim: Police are unbiased

Fryer 2019. See a fuller critique by Knox et al. 2020.



Evidence: Police use lethal force at the same rate against black and white civilians who are stopped.

Claim: Police are unbiased

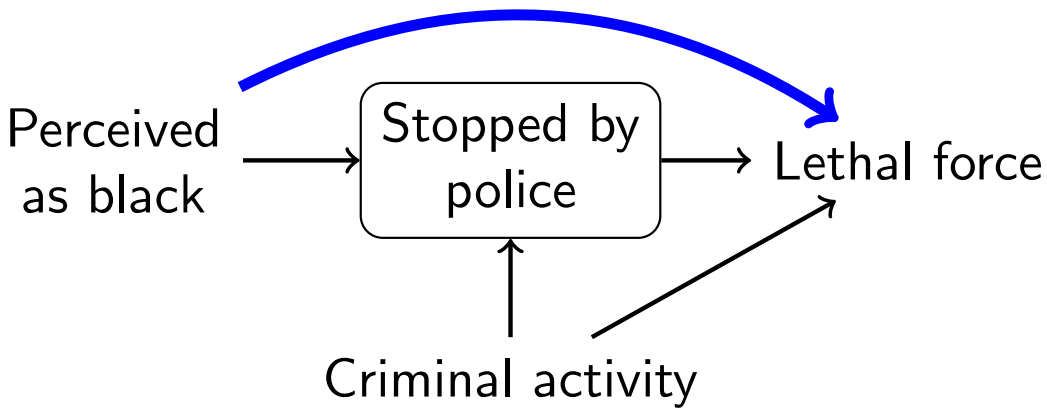
Why wrong: Only the most dangerous whites are stopped. We should expect **more** force used against whites.

Fryer 2019. See a fuller critique by Knox et al. 2020.

Evidence: Police use lethal force at the same rate against black and white civilians who are stopped.

Claim: Police are unbiased

Why wrong: Only the most dangerous whites are stopped. We should expect **more** force used against whites.



Fryer 2019. See a fuller critique by Knox et al. 2020.

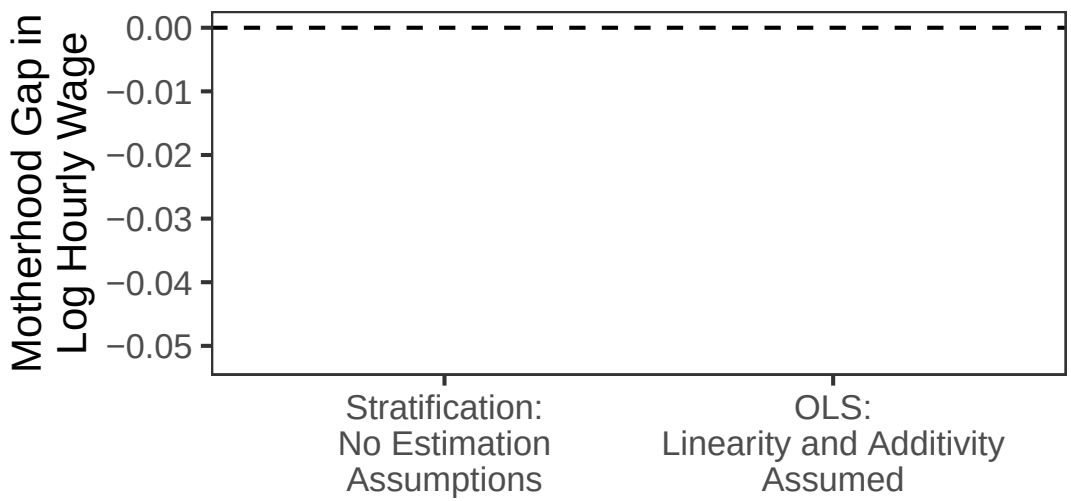
A **statistical model** enters here (and only here)

Pal and Waldfogel 2017

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \Big(\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Mother}) - \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Non-mother}) \Big)$$

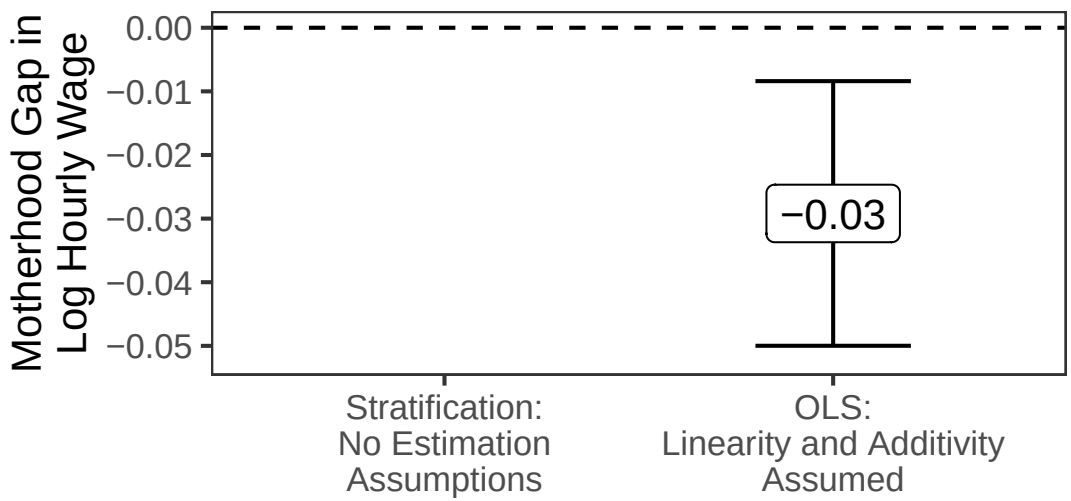
Pal and Waldfogel 2017

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\Big(\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Mother}) - \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Non-mother})\Big)$$



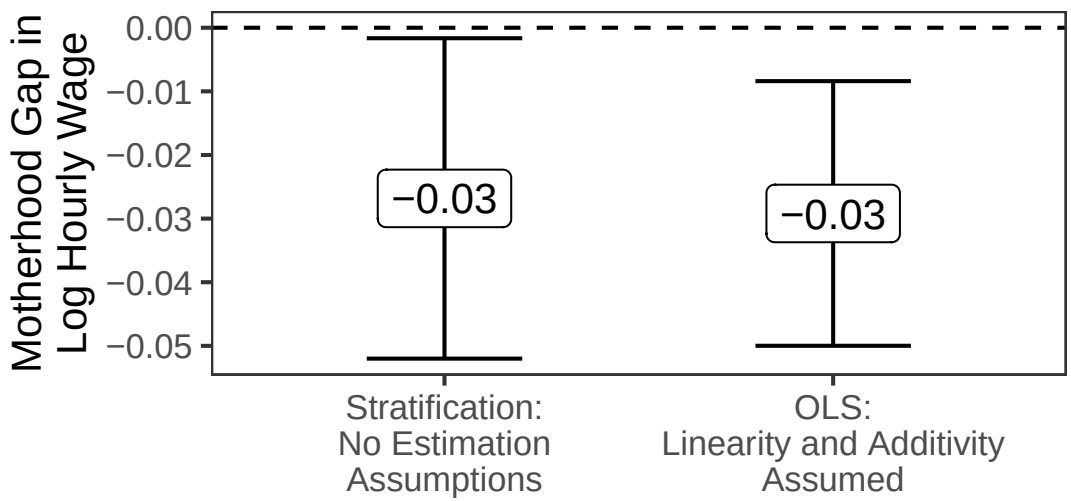
Pal and Waldfogel 2017

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\Big(\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Mother}) - \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Non-mother})\Big)$$



Pal and Waldfogel 2017

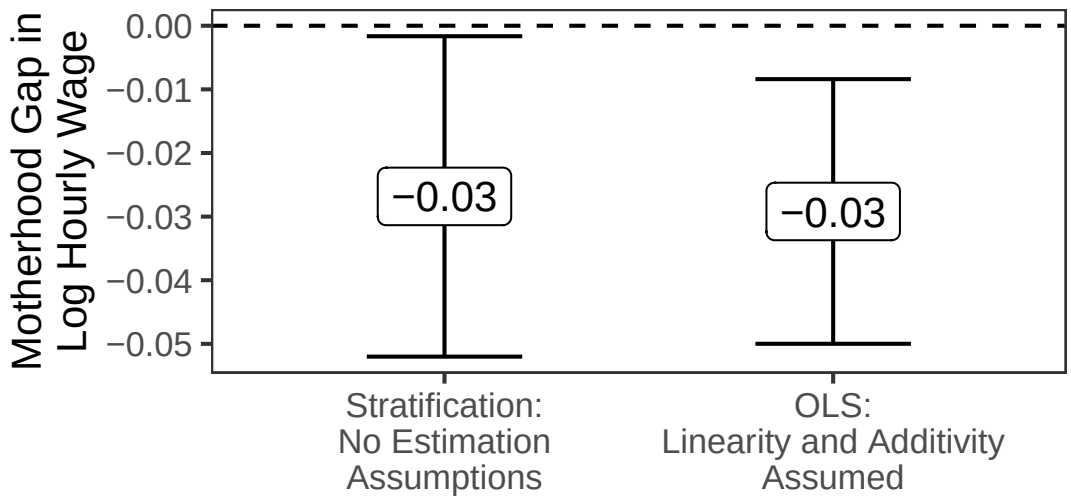
$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\Big(\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Mother}) - \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Non-mother})\Big)$$



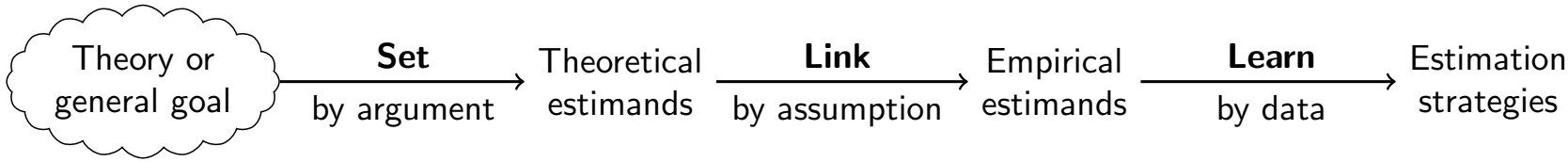
Pal and Waldfogel 2017

Setting the empirical estimand frees us to learn under more **credible** estimation assumptions

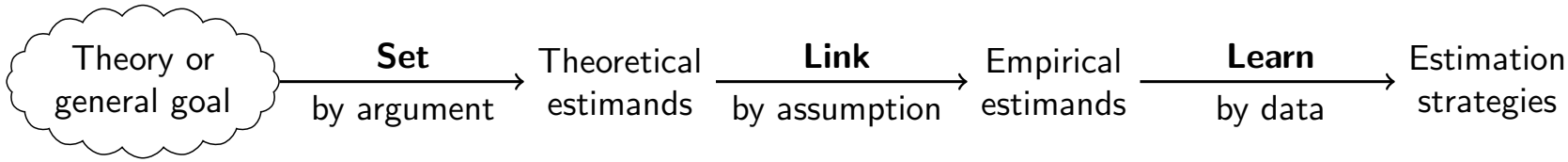
$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\Big(\hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Mother}) - \hat{\mathrm{E}}(Y \mid \vec{X} = \vec{x}_i, \text{Motherhood} = \text{Non-mother})\Big)$$



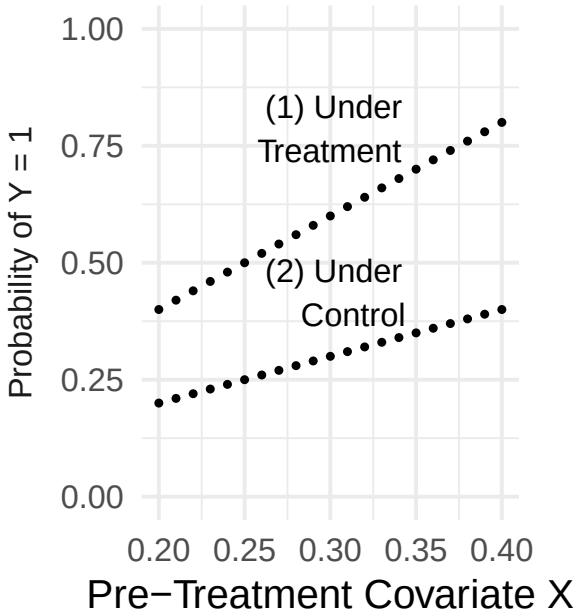
Pal and Waldfogel 2017

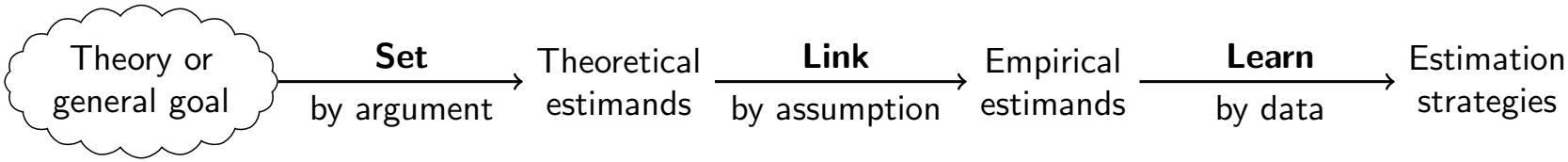


Precise estimands would **clarify confusion** in the literature

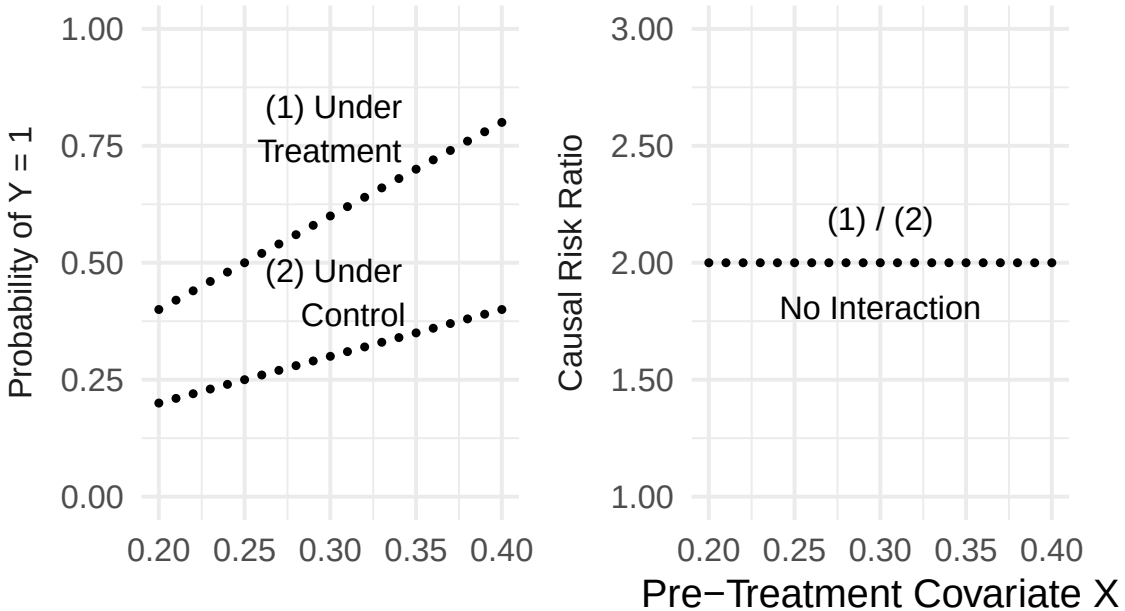


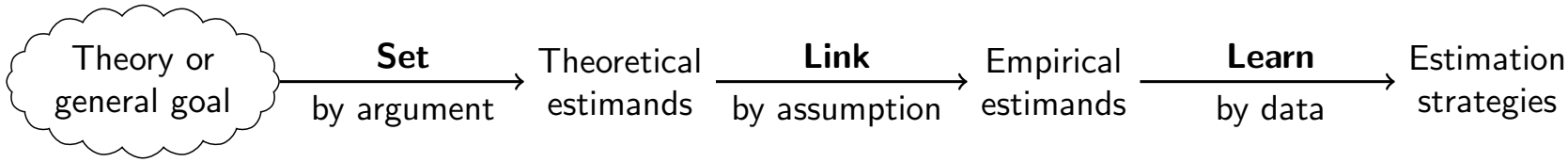
Precise estimands would **clarify confusion** in the literature



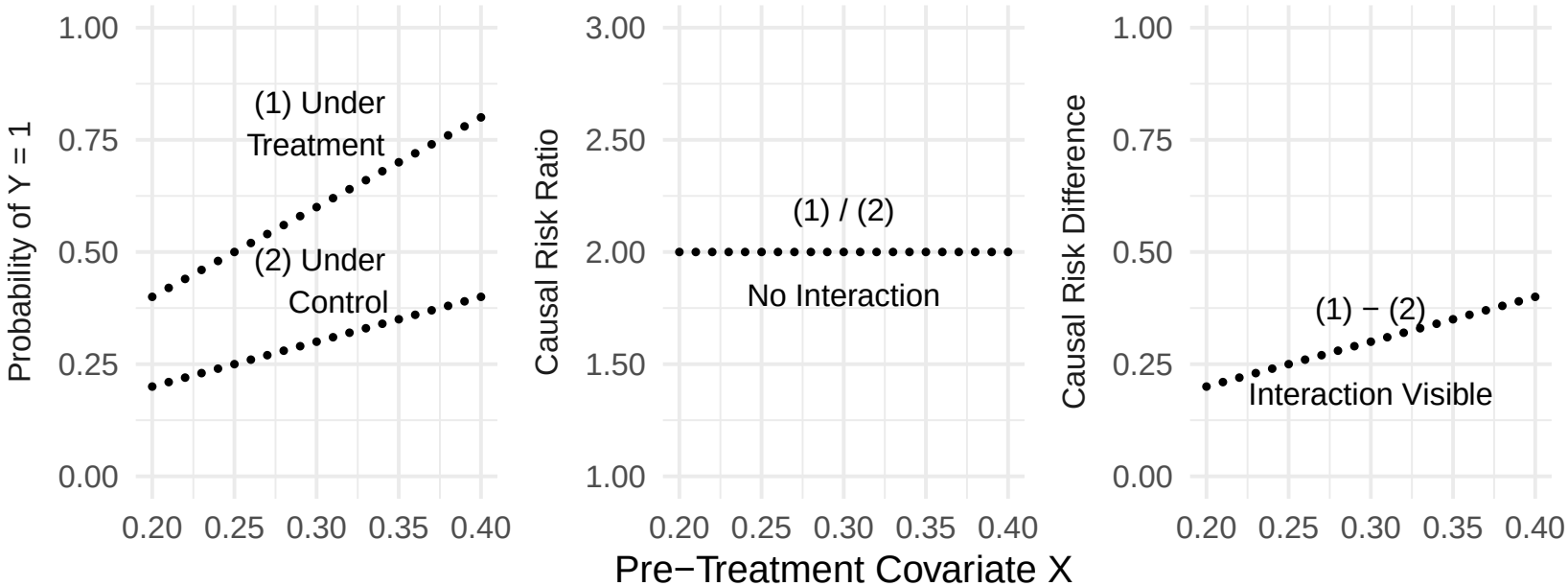


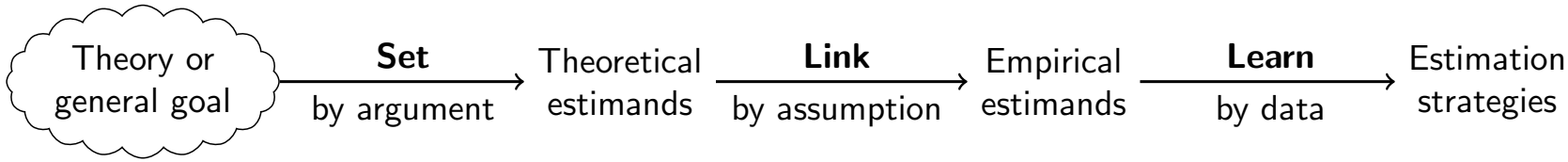
Precise estimands would **clarify confusion** in the literature



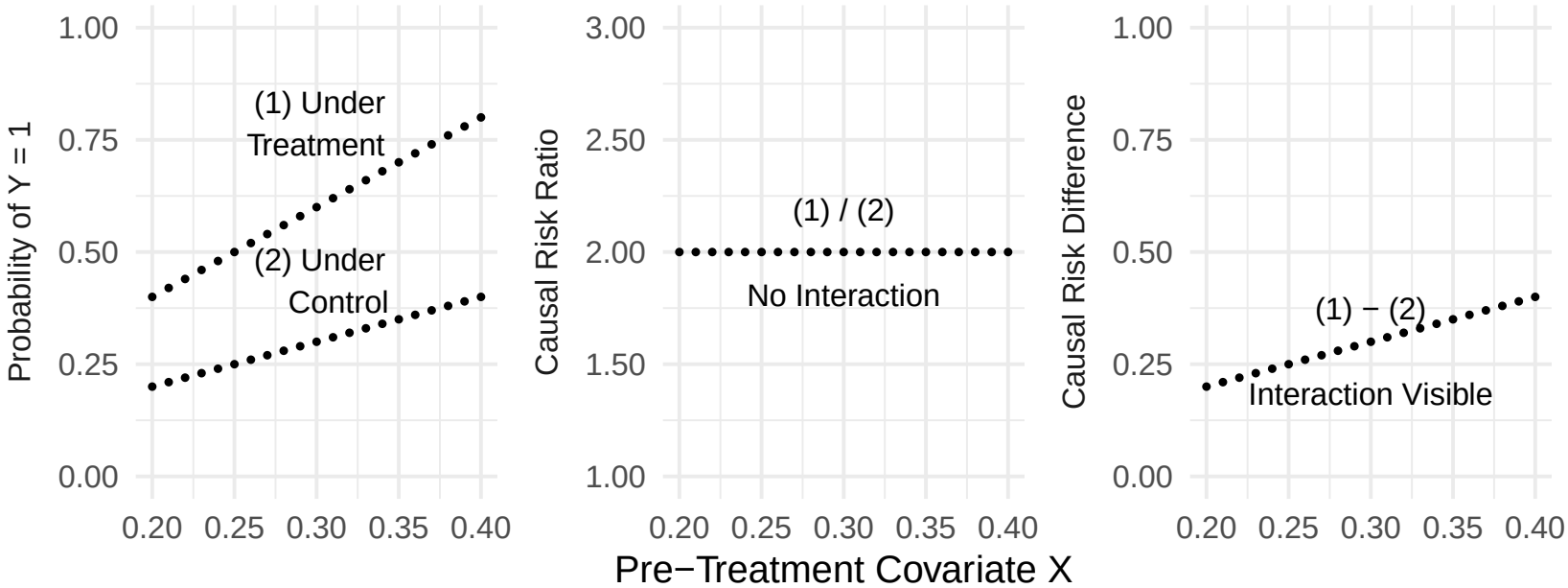


Precise estimands would **clarify confusion** in the literature



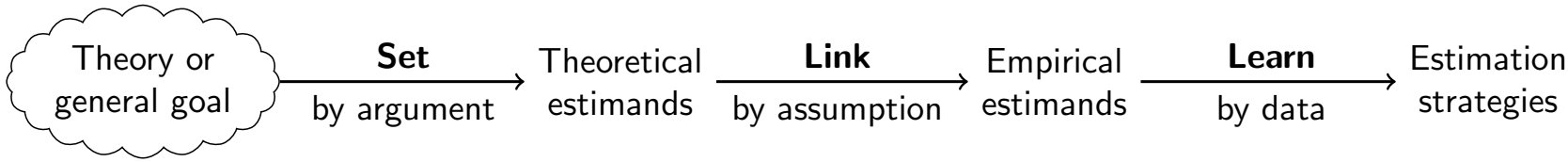


Precise estimands would **clarify confusion** in the literature



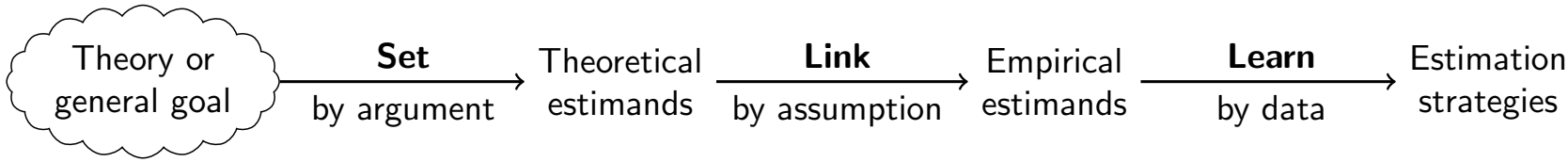
It depends on the **estimand**

Lundberg, Johnson, and Stewart. Setting the Target: Precise Estimands and the Gap Between Theory and Empirics



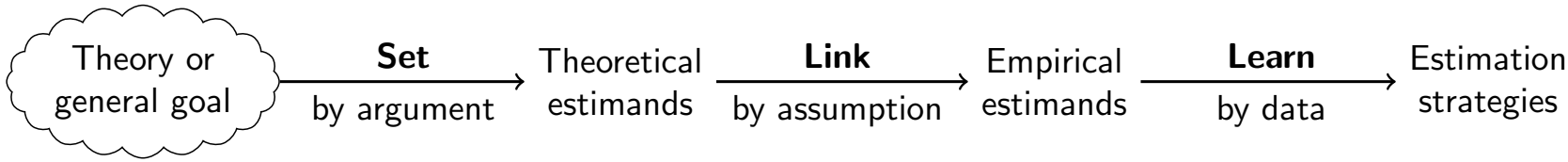
The **target tautology** is widespread

$\rightarrow$ goals are defined by procedures done to data



The **target tautology** is widespread
→ goals are defined by procedures done to data

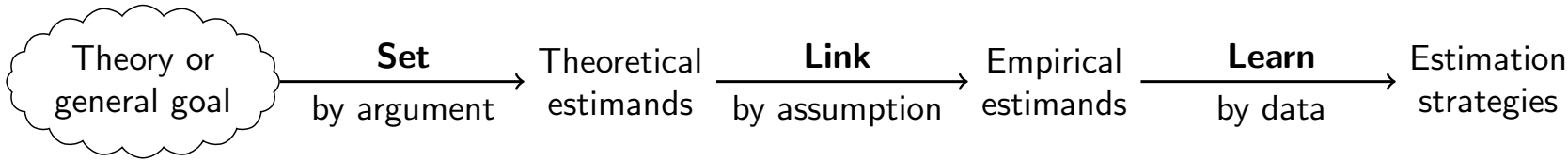
Instead, we should **set the target** nonparametrically
→ could be approximated by many candidate procedures



The **target tautology** is widespread
$\rightarrow$ goals are defined by procedures done to data

Instead, we should **set the target** nonparametrically
$\rightarrow$ could be approximated by many candidate procedures

Estimands bring **benefits** to everyone.

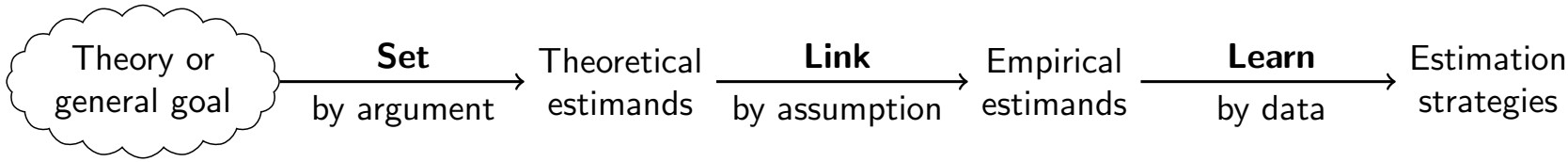


The **target tautology** is widespread
→ goals are defined by procedures done to data

Instead, we should **set the target** nonparametrically
→ could be approximated by many candidate procedures

Estimands bring **benefits** to everyone.

- Estimands help the **analyst** to make methodological choices

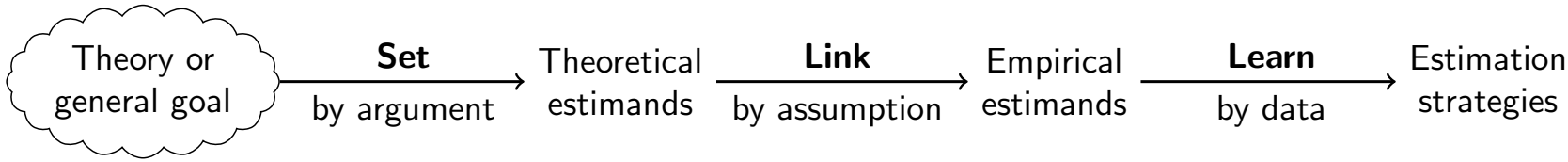


The **target tautology** is widespread
→ goals are defined by procedures done to data

Instead, we should **set the target** nonparametrically
→ could be approximated by many candidate procedures

Estimands bring **benefits** to everyone.

- Estimands help the **analyst** to make methodological choices
- Estimands help the **reviewer** to pinpoint the step where the link to theory breaks down

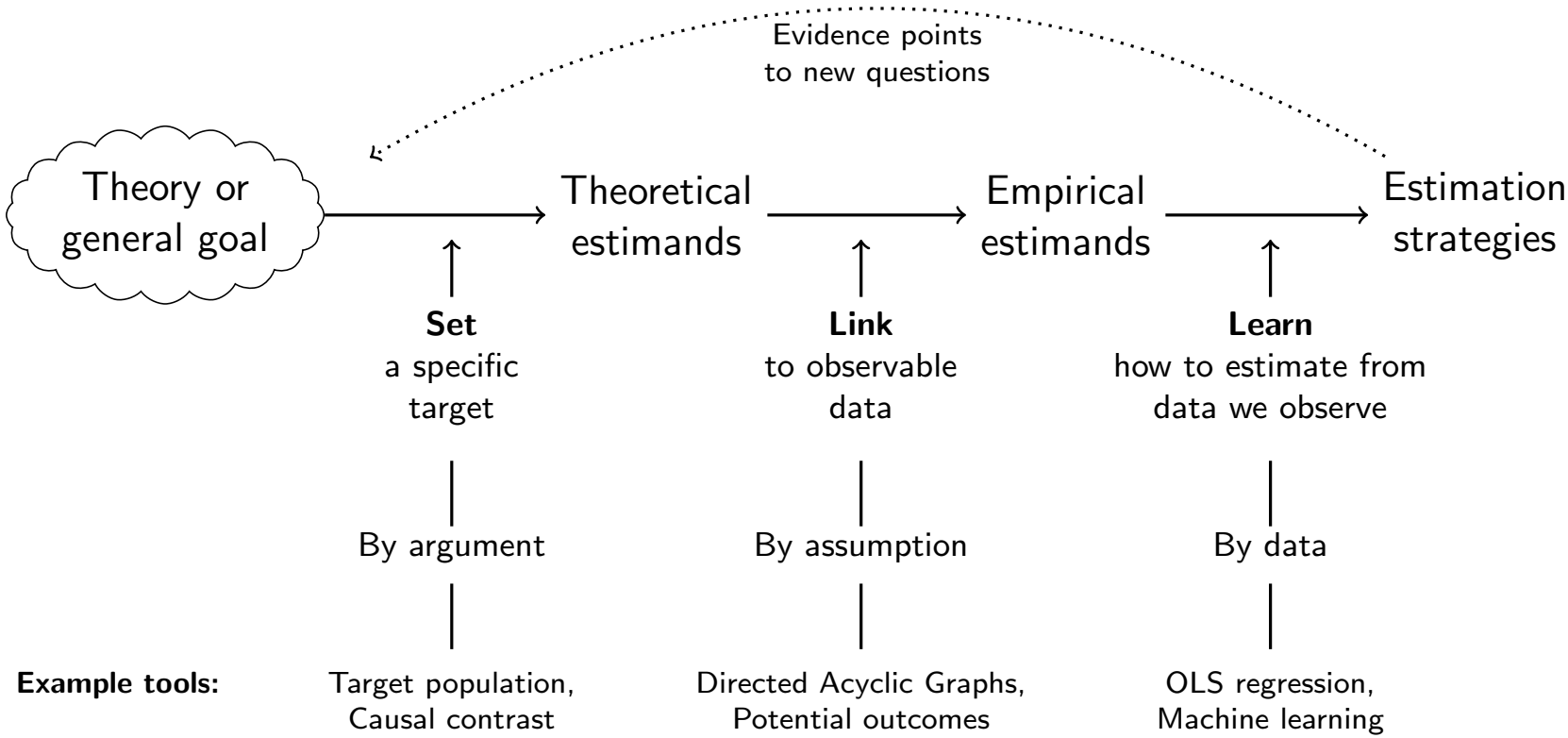


The **target tautology** is widespread
→ goals are defined by procedures done to data

Instead, we should **set the target** nonparametrically
→ could be approximated by many candidate procedures

Estimands bring **benefits** to everyone.

- Estimands help the **analyst** to make methodological choices
- Estimands help the **reviewer** to pinpoint the step where the link to theory breaks down
- Estimands help the **community** to clarify the contribution of each paper



Paper on SocArxiv.
Replication code on GitHub.
Slides on GitHub.