Unbiased but Inflated Causal Effects

Lunch Seminar

Vincent Bagilet & Léo Zabrocki
https://vincentbagilet.github.io/
https://lzabrocki.github.io/

Columbia University
Paris School of Economics

25/01/2022

Warning:

Warning:
This is Still Work in Progress!

Warning:

This is Still Work in Progress!

Generalization of our JMP

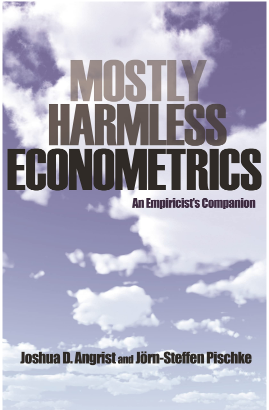
Causality in Observational Studies

Goal: overcoming confounders

How? Matching, IV, RDD, DiD, etc...

Use only part of the variation

Can decrease statistical power



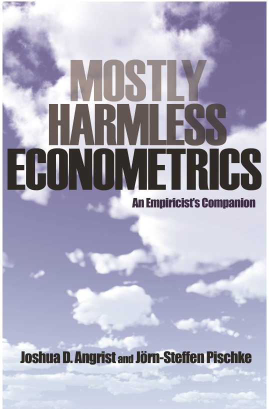
Causality in Observational Studies

Goal: overcoming confounders

How? Matching, IV, RDD, DiD, etc...

Use only part of the variation

Can decrease statistical power



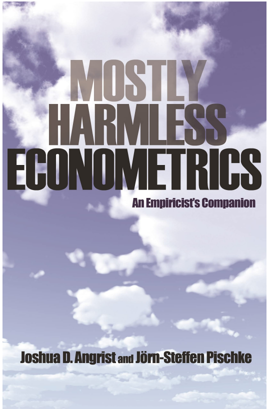
Causality in Observational Studies

Goal: overcoming confounders

How? Matching, IV, RDD, DiD, etc...

Use only part of the variation

Can decrease statistical power



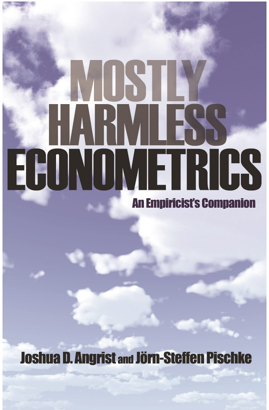
Causality in Observational Studies

Goal: overcoming confounders

How? Matching, IV, RDD, DiD, etc...

Use only part of the variation

Can decrease statistical power



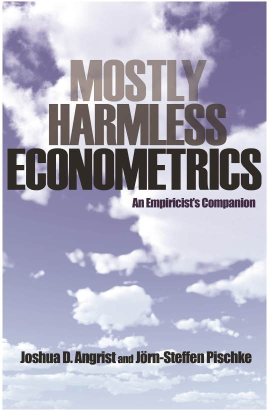
Causality in Observational Studies

Goal: overcoming confounders

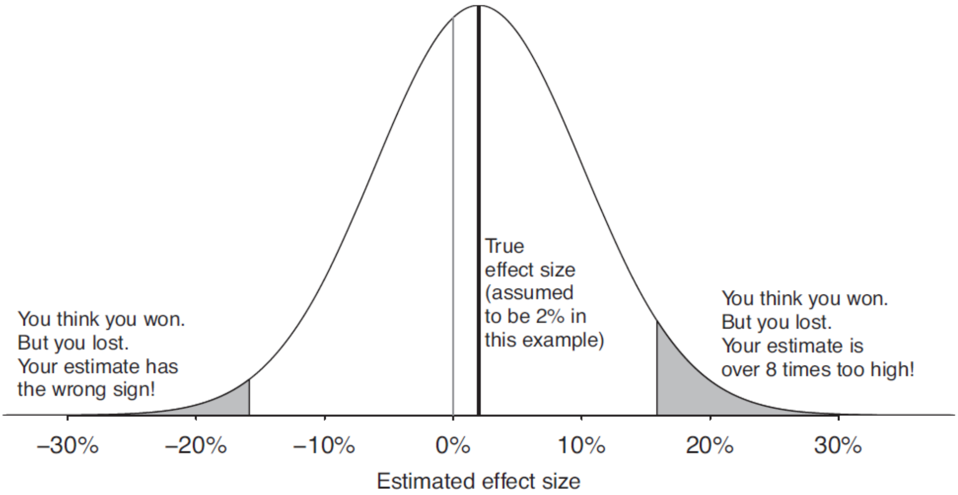
How? Matching, IV, RDD, DiD, etc...

Use only part of the variation

Can decrease statistical power



The Consequences of Low Power



Type M Error

For a given study design, the exaggeration ratio is:

$$\mathbb{E}\left[\left|\frac{\hat{\beta}}{\beta}\right| \middle| \text{statistical significance}\right]$$

Type M Error

For a given study design, the exaggeration ratio is:

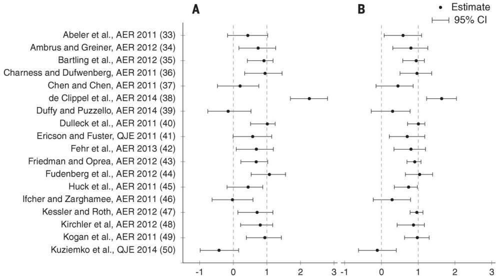
$$\mathbb{E}\left[\left|\frac{\hat{\beta}}{\beta}\right| \middle| \text{statistical significance}\right]$$

Replication Crisis

Low power partly explains the replication crisis

Often discussed for RCT

Example: Camerer *et al.* (2016)

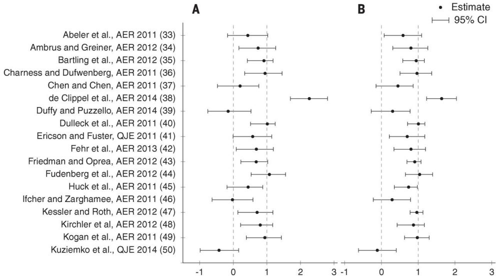


Replication Crisis

Low power partly explains the replication crisis

Often discussed for RCT

Example: Camerer *et al.* (2016)

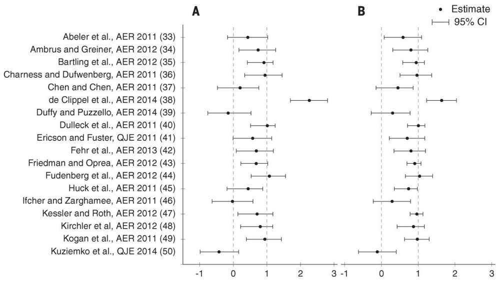


Replication Crisis

Low power partly explains the replication crisis

Often discussed for RCT

Example: Camerer *et al.* (2016)

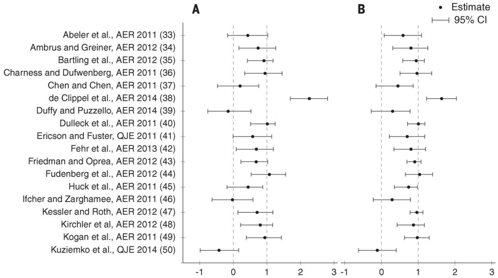


Replication Crisis

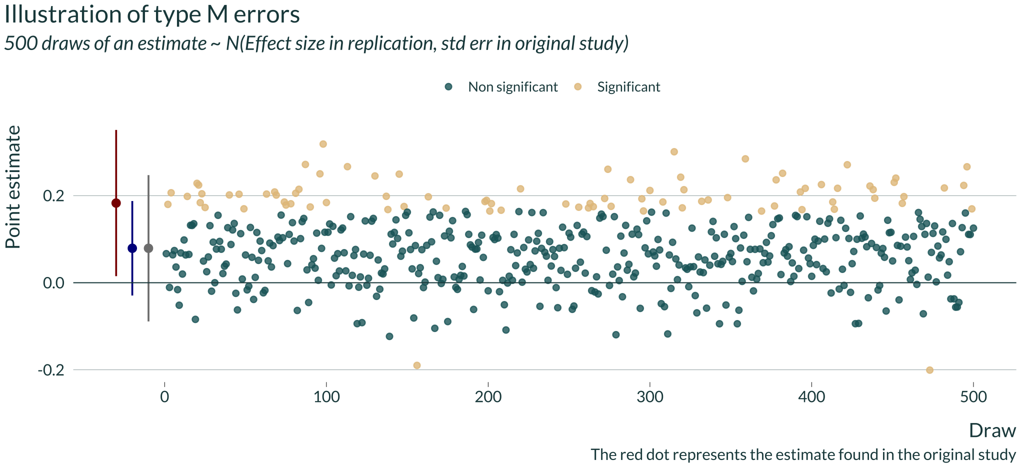
Low power partly explains the replication crisis

Often discussed for RCT

Example: Camerer *et al.* (2016)



Replication Crisis



Story of the Paper in One Picture

figures/meme/1.jpg

Story of the Paper in One Picture

figures/meme/2.jpg

Story of the Paper in One Picture

figures/meme/3.jpg

More Seriously...

To overcome confounding:

- Matching prunes units
- RDD only considers units close to the threshold
- IV only exploits variation explained by the instrument

But they throw out variation and could reduce power

More Seriously...

To overcome confounding:

Matching prunes units

RDD only considers units close to the threshold

IV only exploits variation explained by the instrument

But they throw out variation and could reduce power

More Seriously...

To overcome confounding:

Matching prunes units

RDD only considers units close to the threshold

IV only exploits variation explained by the instrument

But they throw out variation and could reduce power

More Seriously...

To overcome confounding:

Matching prunes units

RDD only considers units close to the threshold

IV only exploits variation explained by the instrument

But they throw out variation and could reduce power

More Seriously...

To overcome confounding:

Matching prunes units

RDD only considers units close to the threshold

IV only exploits variation explained by the instrument

But they throw out variation and could reduce power

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding
- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding
- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding

- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding
- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding
- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

What we Do

We relate two existing literatures

Using simulations based on published research:

- We show how the tuning of key parameters of causal inference methods can overcome confounding
- ...but also produce inflated estimates if they are too stringent

Discuss a list of solutions to take this issue into account

Contributions

1. Contribute to the bias-variance debate
   (Imbens & Kalyanaraman 2012, Deaton & Cartwright 2018, Hernán & Robins 2020)

2. Revisit this trade-off through the lens of publication bias
   (Ionnadis 2008, Gelman & Carlin 2014, Brodeur et al. 2020)

3. Add another stone to the literature on the quality of inference
   (Black et al. 2021, Broderick et al. 2021, Young 2021)

Contributions

1. Contribute to the bias-variance debate
   (Imbens & Kalyanaraman 2012, Deaton & Cartwright 2018, Hernán & Robins 2020)

2. Revisit this trade-off through the lens of publication bias
   (Ionnadis 2008, Gelman & Carlin 2014, Brodeur et al. 2020)

3. Add another stone to the literature on the quality of inference
   (Black et al. 2021, Broderick et al. 2021, Young 2021)

Contributions

1. Contribute to the bias-variance debate
   (Imbens & Kalyanaraman 2012, Deaton & Cartwright 2018, Hernán & Robins 2020)
2. Revisit this trade-off through the lens of publication bias
   (Ionnadis 2008, Gelman & Carlin 2014, Brodeur et al. 2020)
3. Add another stone to the literature on the quality of inference
   (Black et al. 2021, Broderick et al. 2021, Young 2021)

Contributions

1. Contribute to the bias-variance debate
   (Imbens & Kalyanaraman 2012, Deaton & Cartwright 2018, Hernán & Robins 2020)
2. Revisit this trade-off through the lens of publication bias
   (Ionnadis 2008, Gelman & Carlin 2014, Brodeur et al. 2020)
3. Add another stone to the literature on the quality of inference
   (Black et al. 2021, Broderick et al. 2021, Young 2021)

Unbiased but Inflated Causal Effects

GitHub

Hi and welcome!

This website gathers code and additional material for the paper "Unbiased but Inflated Causal Effects" by Vincent Bagilet and Léo Zabrocki.

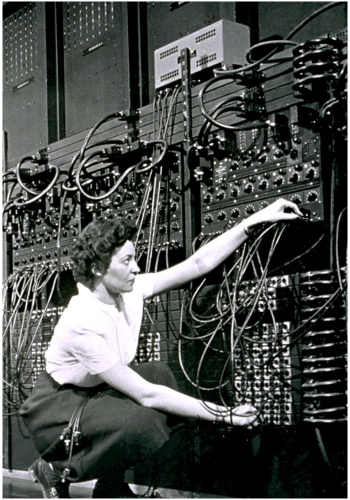
The website is under construction and the analysis is still at a preliminary stage.

Simulating the Power Issue of Causal Inference Methods

General Approach

For each causal inference method:

- Take a real research example
- Simulate from scratch its DGP
- Vary the key parameter for overcoming confounding



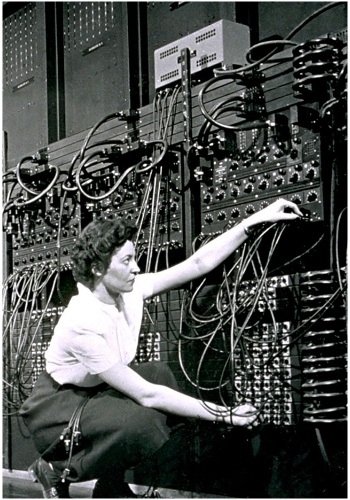
General Approach

For each causal inference method:

Take a real research example

Simulate from scratch its DGP

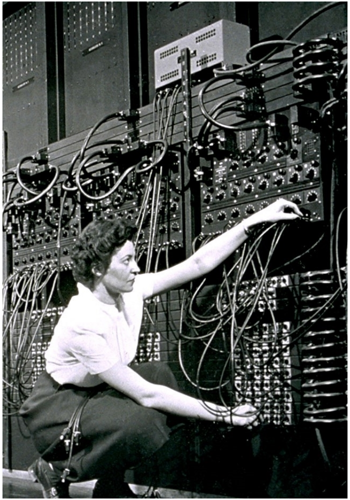
Vary the key parameter for overcoming confounding



General Approach

For each causal inference method:

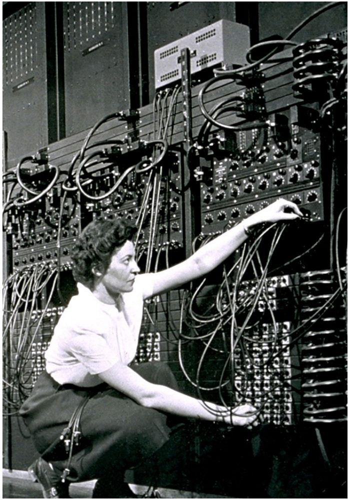
- Take a real research example
- Simulate from scratch its DGP
- Vary the key parameter for overcoming confounding



General Approach

For each causal inference method:

- Take a real research example
- Simulate from scratch its DGP
- Vary the key parameter for overcoming confounding



Matching

Case-study:

Non-randomized labor training program

25% units take the treatment

ATT = + 100 euros

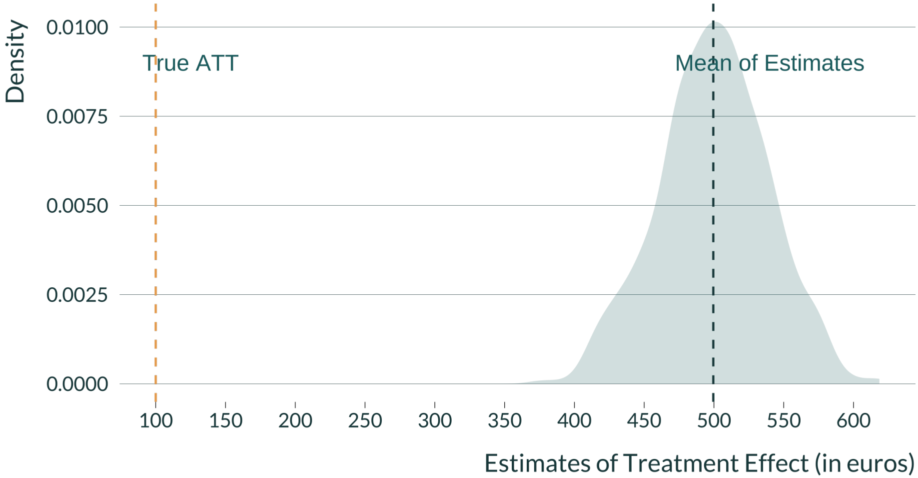
Propensity score matching

Matching

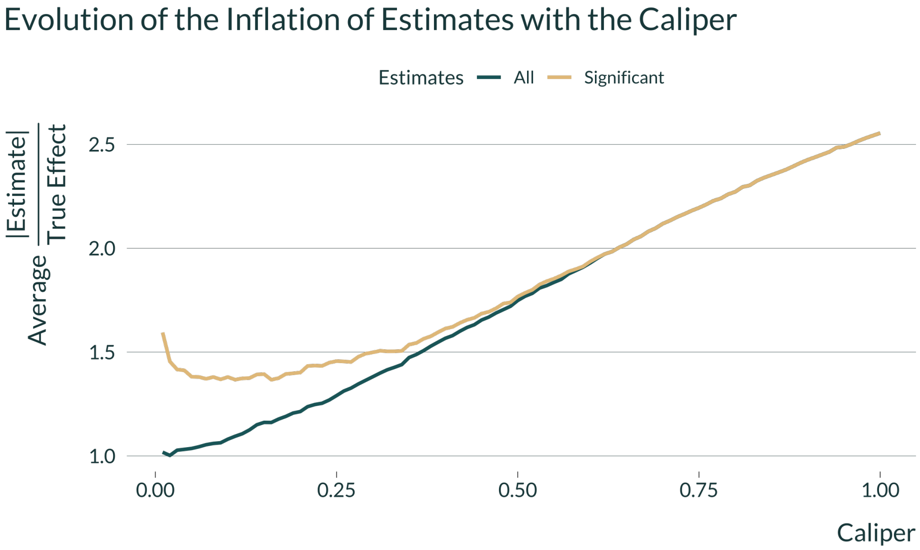
figures/2.simulations/graph_true_ps.png

Matching

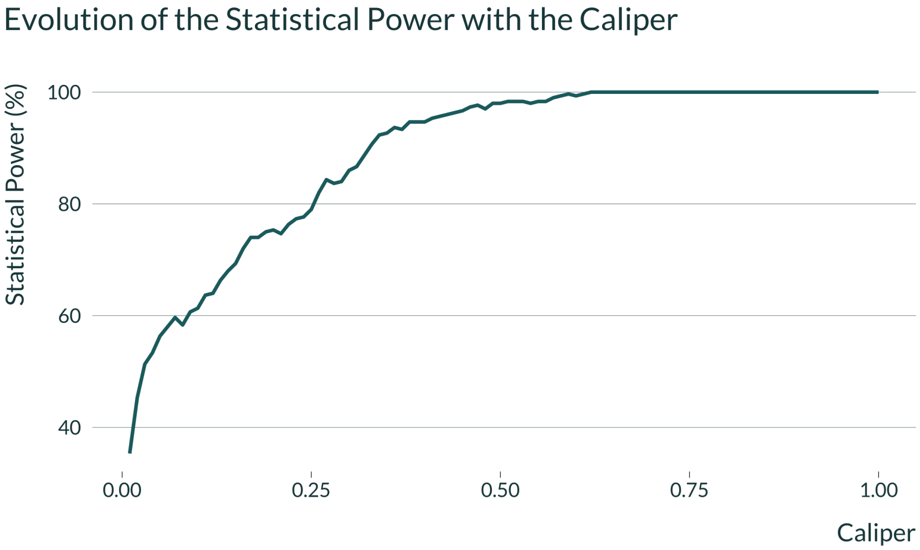
Distribution of Outcome Regression Estimates



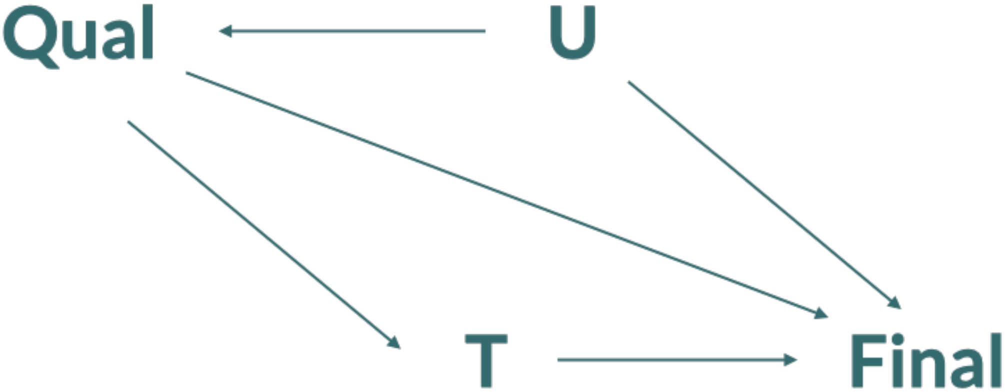
Matching


Evolution of the Inflation of Estimates with the Caliper
Estimates
All
Significant
Average |Estimate| / True Effect
2.5
2.0
1.5
1.0
0.00
0.25
0.50
0.75
1.00
Caliper

Matching


Evolution of the Statistical Power with the Caliper
Statistical Power (%)
100
80
60
40
0.00
0.25
0.50
0.75
1.00
Caliper

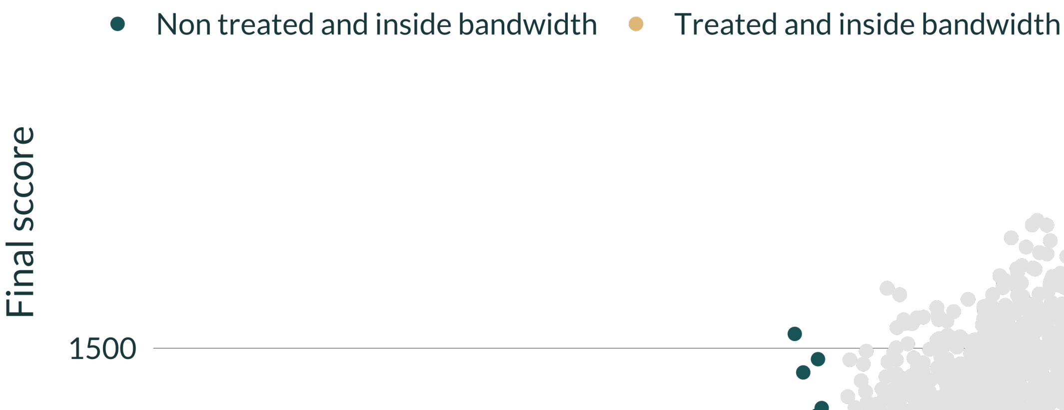
RDD



RDD

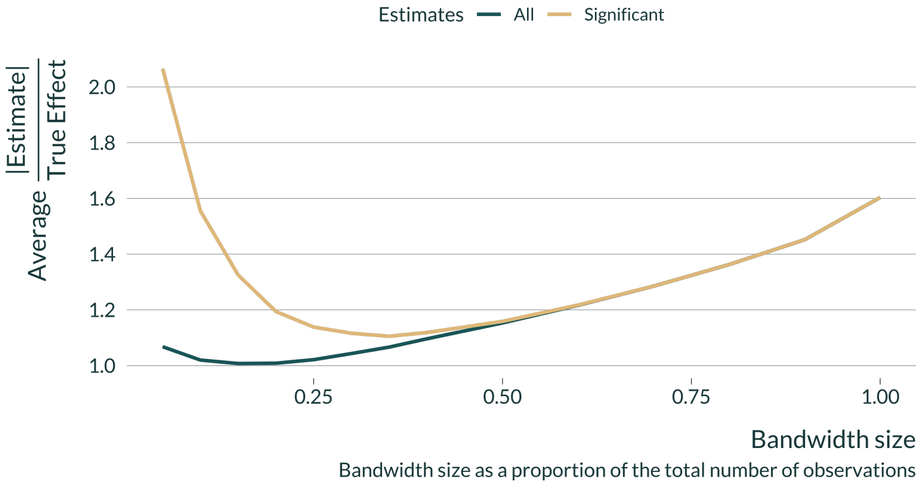
Final scores against qualification scores

Illustration of the definition of the bandwidth and the treatment

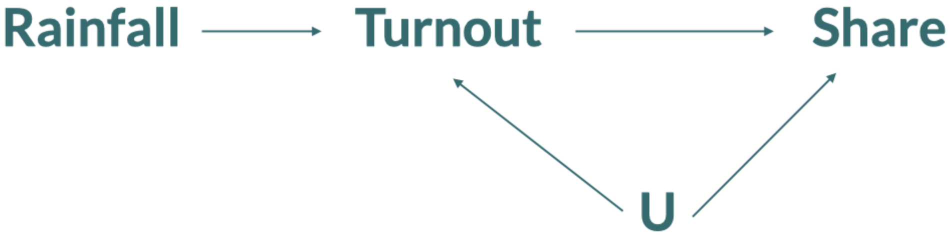


RDD

Evolution of bias with bandwith size, conditional on significativity



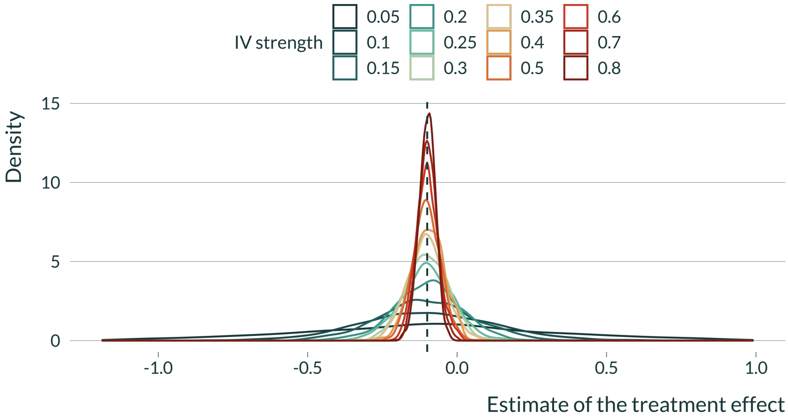
IV



IV

Distribution of the estimates of the treatment effect

Comparison across IV strengths



For readability, extreme estimates are filtered out
The vertical line represents the true effect

IV

figures/2.simulations/graph_results_IV

IV

□figures/2.simulations/fstat_bias_simpl

Event-Study Designs

What is the trade-off for event-study designs?

Even in large sample, a low number of observations affected by exogenous shocks could decrease the power

Event-Study Designs

What is the trade-off for event-study designs?

Even in large sample, a low number of observations affected by exogenous shocks could decrease the power

Event-Study Designs

What is the trade-off for event-study designs?

Even in large sample, a low number of observations affected by exogenous shocks could decrease the power

Cross-Cutting Issues

All designs could be affected:

- Proportion of treated units
- Distribution of the outcomes
- Outliers
- ...

Cross-Cutting Issues

All designs could be affected:

- Proportion of treated units
- Distribution of the outcomes
- Outliers
- ...

Cross-Cutting Issues

All designs could be affected:

- Proportion of treated units
- Distribution of the outcomes
- Outliers
- ...

Cross-Cutting Issues

All designs could be affected:

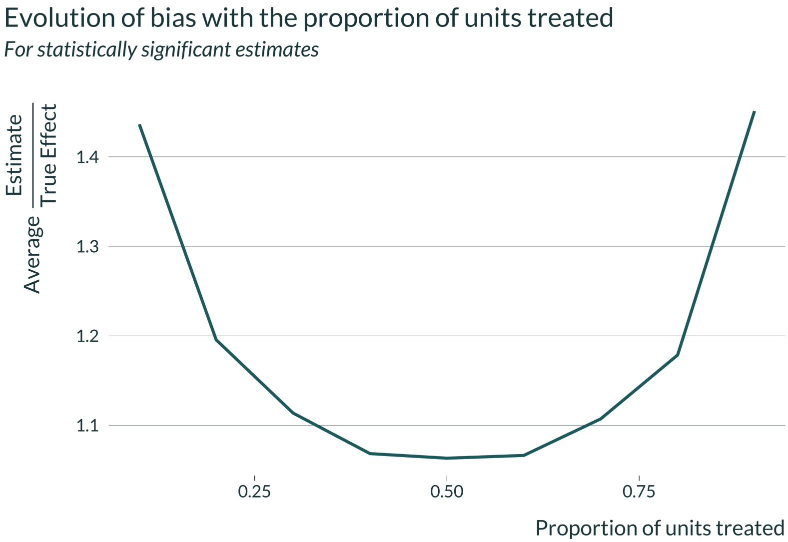
- Proportion of treated units
- Distribution of the outcomes
- Outliers
- ...

Cross-Cutting Issues

All designs could be affected:

- Proportion of treated units
- Distribution of the outcomes
- Outliers
- ...

Proportion of Treated Units



How to Address the Power Issue?

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Keeping the Issue in Mind at All Stages

1. What to do before embarking on an observational study?

   Evaluate the power of the study by simulating its DGP

2. What to do after analyzing the results?

   Compute what would be the power for alternative effect sizes

   Assess if the analysis suffers from confounding

3. Solving the issue?

   Methods for de-biasing inflated estimates [in progress]
   (Drysdale 2020, van Zwet & Cator 2021)

Prospective Power Simulations

Use statistical formulas (see AEA RCT Registry)

But not very flexible

Better & recommenced approach is to simulate the DGP
(EGAP, Black et al. 2021, Gelman et al. 2021)

But how? Very few guidance exist!

Prospective Power Simulations

Use statistical formulas (see AEA RCT Registry)

But not very flexible

Better & recommenced approach is to simulate the DGP
(EGAP, Black et al. 2021, Gelman et al. 2021)

But how? Very few guidance exist!

Prospective Power Simulations

Use statistical formulas (see AEA RCT Registry)

But not very flexible

Better & recommenced approach is to simulate the DGP
(EGAP, Black et al. 2021, Gelman et al. 2021)

But how? Very few guidance exist!

Prospective Power Simulations

Use statistical formulas (see AEA RCT Registry)

But not very flexible

Better & recommenced approach is to simulate the DGP
(EGAP, Black et al. 2021, Gelman et al. 2021)

But how? Very few guidance exist!

Prospective Power Simulations

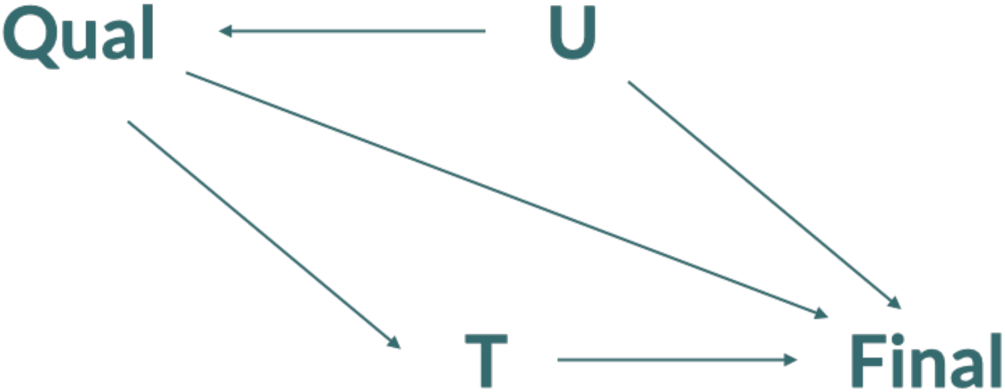
Use statistical formulas (see AEA RCT Registry)

But not very flexible

Better & recommenced approach is to simulate the DGP
(EGAP, Black et al. 2021, Gelman et al. 2021)

But how? Very few guidance exist!

Prospective Power Simulations



The Model

N the number of students

$U_i \sim N(0, \sigma_u^2)$ the unobserved ability

$Qualification_i = H_i + \sigma U_i^2$ where $H \sim N(\mu_h, \sigma_h^2)$

$\epsilon_i \sim N(0, \sigma_e^2)$

$Final_i = \alpha + \beta T_i + \gamma Qualification_i + U_i^2 + \epsilon_i$

The Model

N the number of students

$U_i \sim N(0, \sigma_u^2)$ the unobserved ability

$Qualification_i = H_i + \sigma U_i^2$ where $H \sim N(\mu_h, \sigma_h^2)$

$\epsilon_i \sim N(0, \sigma_e^2)$

$Final_i = \alpha + \beta T_i + \gamma Qualification_i + U_i^2 + \epsilon_i$

The Model

N the number of students

$U_i \sim N(0, \sigma_u^2)$ the unobserved ability

$Qualification_i = H_i + \sigma U_i^2$ where $H \sim N(\mu_h, \sigma_h^2)$

$\epsilon_i \sim N(0, \sigma_e^2)$

$Final_i = \alpha + \beta T_i + \gamma Qualification_i + U_i^2 + \epsilon_i$

The Model

N the number of students

$U_i \sim N(0, \sigma_u^2)$ the unobserved ability

$Qualification_i = H_i + \sigma U_i^2$ where $H \sim N(\mu_h, \sigma_h^2)$

$\epsilon_i \sim N(0, \sigma_e^2)$

$Final_i = \alpha + \beta T_i + \gamma Qualification_i + U_i^2 + \epsilon_i$

The Model

N the number of students

$U_i \sim N(0, \sigma_u^2)$ the unobserved ability

$Qualification_i = H_i + \sigma U_i^2$ where $H \sim N(\mu_h, \sigma_h^2)$

$\epsilon_i \sim N(0, \sigma_e^2)$

$Final_i = \alpha + \beta T_i + \gamma Qualification_i + U_i^2 + \epsilon_i$

Workflow


SET
PARAMETER
VALUES
GENERATE DATA
ESTIMATE MODEL
REPEAT *N* TIMES

```r
generate_data_rdd <- function(N, sigma_u, mu_h, sigma_h, sigma_e,
    alpha, beta, gamma, delta, q_c) {

  data <- tibble(id = 1:N) %>%
    mutate(
      u = rnorm(nrow(.), 0, sigma_u),
      qual = rnorm(nrow(.), mu_h, sigma_h) + delta*u^2,
      e = rnorm(nrow(.), 0, sigma_e),
      treated = qual < quantile(qual, q_c),
      final0 = alpha + gamma*qual + delta*u^2 + e,
      final1 = final0 + beta,
      final = final0 + beta*treated,
      q_c = q_c
    )

  return(data)
}
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm 0.2 \nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. ± 0.2 $\nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm$0.2 $\nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm$0.2 $\nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm 0.2 \nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm$0.2 $\nearrow$ in test scores

Does this estimate make sense?

`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Same previous example

$\hat{\beta}$ = 0.5 S.D. $\pm 0.2 \nearrow$ in test scores

Does this estimate make sense?

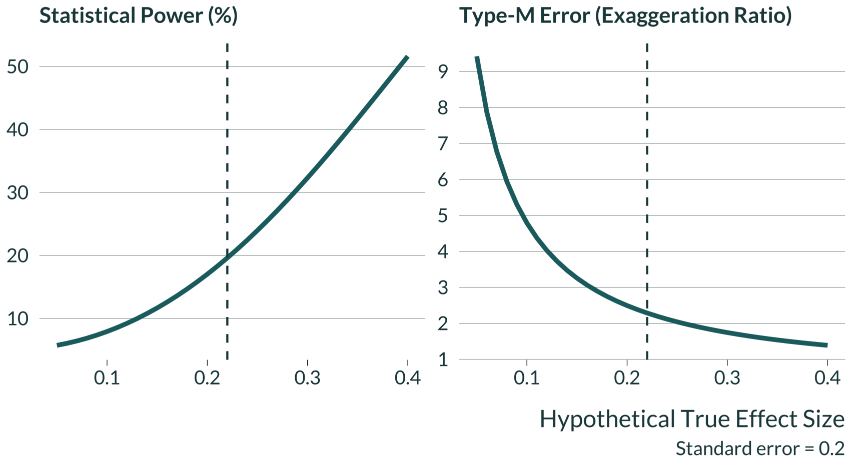
`retrodesign` implents the power analysis

Only guesses about true effect sizes + S.E. are needed

```
retro_design(0.3, 0.2)
```

Retrospective Power Simulations

Evolution of power and type-M error with true effect size

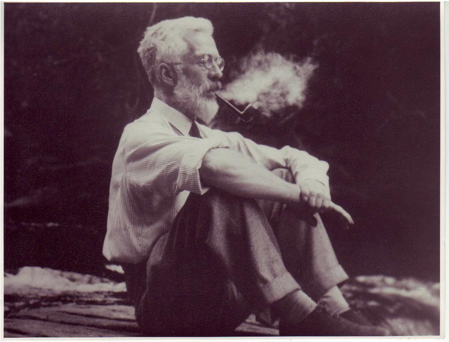


Quantitative Bias Analysis

Old but hot topic:

- Cornfield et al. (1959)
- Rosenbaum (2002, 2010)
- Cinelli & Hazlett (2020)

Quantify how hidden biases of various magnitudes can alter conclusions

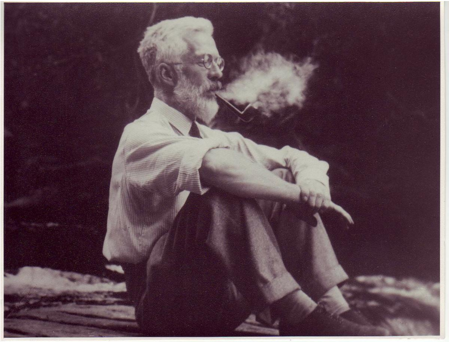


Quantitative Bias Analysis

Old but hot topic:

- Cornfield et al. (1959)
- Rosenbaum (2002, 2010)
- Cinelli & Hazlett (2020)

Quantify how hidden biases of various magnitudes can alter conclusions



Quantitative Bias Analysis

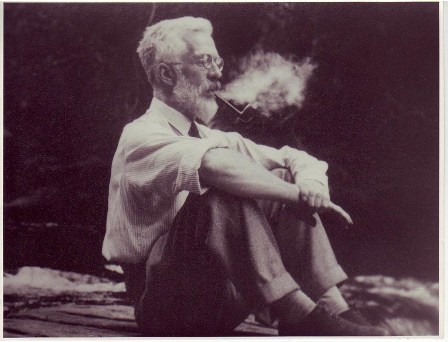
Old but hot topic:

Cornfield et al. (1959)

Rosenbaum (2002, 2010)

Cinelli & Hazlett (2020)

Quantify how hidden biases of various magnitudes can alter conclusions

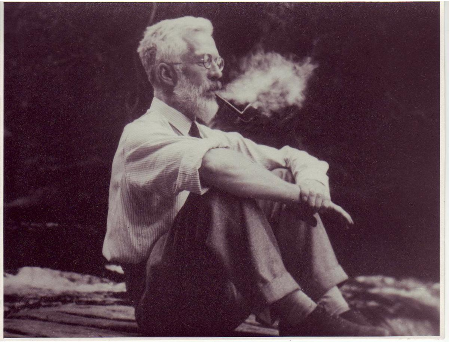


Quantitative Bias Analysis

Old but hot topic:

- Cornfield et al. (1959)
- Rosenbaum (2002, 2010)
- Cinelli & Hazlett (2020)

Quantify how hidden biases of various magnitudes can alter conclusions

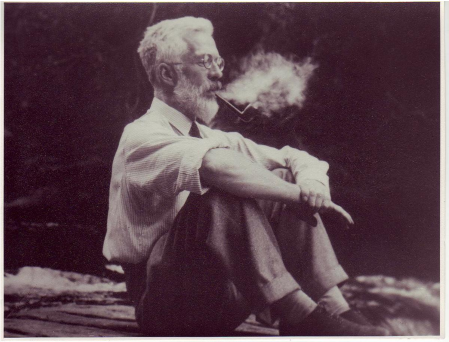


Quantitative Bias Analysis

Old but hot topic:

- Cornfield et al. (1959)
- Rosenbaum (2002, 2010)
- Cinelli & Hazlett (2020)

Quantify how hidden biases of various magnitudes can alter conclusions



De-Biasing the Estimate

Forthcoming work

De-Biasing the Estimate

Forthcoming work

Concluding Thoughts

Bias in Observational Studies

King and Zeng (2007):

$$\text{Bias} = \text{Omitted Variable Bias} + \text{Post-Treatment Bias} + \text{Interpolation Bias} + \text{Extrapolation Bias}$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$\text{Bias} = \text{Omitted Variable Bias} + \text{Post-Treatment Bias} + \text{Interpolation Bias} + \text{Extrapolation Bias}$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$\text{Bias} = \text{Omitted Variable Bias} + \text{Post-Treatment Bias} + \text{Interpolation Bias} + \text{Extrapolation Bias}$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$
\begin{aligned}
\text{Bias} = \; & \text{Omitted Variable Bias} + \\
& \text{Post-Treatment Bias} + \\
& \text{Interpolation Bias} + \\
& \text{Extrapolation Bias}
\end{aligned}
$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$\begin{aligned} \text{Bias} = \ & \text{Omitted Variable Bias} + \\ & \text{Post-Treatment Bias} + \\ & \text{Interpolation Bias} + \\ & \text{Extrapolation Bias} \end{aligned}$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$
\begin{aligned}
\text{Bias} = \; & \text{Omitted Variable Bias} + \\
& \text{Post-Treatment Bias} + \\
& \text{Interpolation Bias} + \\
& \text{Extrapolation Bias}
\end{aligned}
$$

... We still need to think about power

Bias in Observational Studies

King and Zeng (2007):

$$\begin{aligned}\text{Bias} = \ & \text{Omitted Variable Bias} + \\ & \text{Post-Treatment Bias} + \\ & \text{Interpolation Bias} + \\ & \text{Extrapolation Bias}\end{aligned}$$

... We still need to think about power

General Recommendations

Change attitude towards statistically insignificant results
(Ziliak and McCloskey 2008, Wasserstein and Lazar 2016, McShane et al. 2019)

Avoid dichotomizing evidence using the 5% significance threshold
(Greenland 2017)

Replicate studies with similar designs!
(Christensen 2019)

Embrace uncertainty by interpreting CI's width
(Amrhein et al. 2019, Gelman et al. 2020, Romer 2020)

General Recommendations

Change attitude towards statistically insignificant results
(Ziliak and McCloskey 2008, Wasserstein and Lazar 2016, McShane et al. 2019)

Avoid dichotomizing evidence using the 5% significance threshold
(Greenland 2017)

Replicate studies with similar designs!
(Christensen 2019)

Embrace uncertainty by interpreting CI's width
(Amrhein et al. 2019, Gelman et al. 2020, Romer 2020)

General Recommendations

Change attitude towards statistically insignificant results
(Ziliak and McCloskey 2008, Wasserstein and Lazar 2016, McShane et al. 2019)

Avoid dichotomizing evidence using the 5% significance threshold
(Greenland 2017)

Replicate studies with similar designs!
(Christensen 2019)

Embrace uncertainty by interpreting CI's width
(Amrhein et al. 2019, Gelman et al. 2020, Romer 2020)

General Recommendations

Change attitude towards statistically insignificant results
(Ziliak and McCloskey 2008, Wasserstein and Lazar 2016, McShane et al. 2019)

Avoid dichotomizing evidence using the 5% significance threshold
(Greenland 2017)

Replicate studies with similar designs!
(Christensen 2019)

Embrace uncertainty by interpreting CI's width
(Amrhein et al. 2019, Gelman et al. 2020, Romer 2020)

General Recommendations

Change attitude towards statistically insignificant results
(Ziliak and McCloskey 2008, Wasserstein and Lazar 2016, McShane et al. 2019)

Avoid dichotomizing evidence using the 5% significance threshold
(Greenland 2017)

Replicate studies with similar designs!
(Christensen 2019)

Embrace uncertainty by interpreting CI's width
(Amrhein et al. 2019, Gelman et al. 2020, Romer 2020)

A Concrete Example

Banerjee, Duflo, & Sharma (AER: Insights 2021):

> *Treated households are 0.2 SD more food secure than the control group by 18 months and 0.25 SD by year three ($p < 0.01$). This effect grows to 0.4 SD by year seven and remains at 0.13 SD by year ten ($p < 0.05$).*

Alternative phrasing:

> *Our data are compatible with effect sizes ranging from + 0.004 SD to + 0.25 SD.*

A Concrete Example

Banerjee, Duflo, & Sharma (AER: Insights 2021):

> *Treated households are 0.2 SD more food secure than the control group by 18 months and 0.25 SD by year three ($p < 0.01$). This effect grows to 0.4 SD by year seven and remains at 0.13 SD by year ten ($p < 0.05$).*

Alternative phrasing:

> *Our data are compatible with effect sizes ranging from + 0.004 SD to + 0.25 SD.*

A Concrete Example

Banerjee, Duflo, & Sharma (AER: Insights 2021):

> *Treated households are 0.2 SD more food secure than the control group by 18 months and 0.25 SD by year three ($p < 0.01$). This effect grows to 0.4 SD by year seven and remains at 0.13 SD by year ten ($p < 0.05$).*

Alternative phrasing:

> *Our data are compatible with effect sizes ranging from + 0.004 SD to + 0.25 SD.*

A Concrete Example

Banerjee, Duflo, & Sharma (AER: Insights 2021):

> *Treated households are 0.2 SD more food secure than the control group by 18 months and 0.25 SD by year three ($p < 0.01$). This effect grows to 0.4 SD by year seven and remains at 0.13 SD by year ten ($p < 0.05$).*

Alternative phrasing:

> *Our data are compatible with effect sizes ranging from + 0.004 SD to + 0.25 SD.*

A Concrete Example

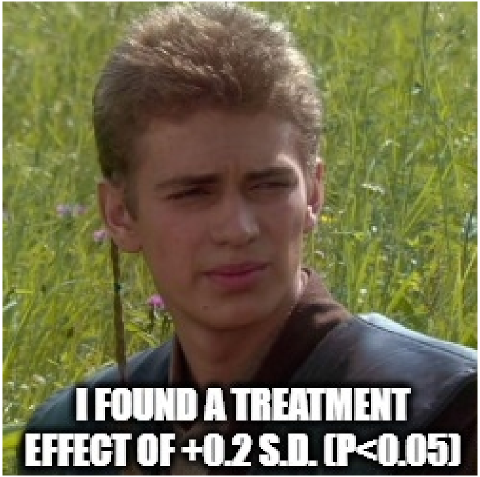
Banerjee, Duflo, & Sharma (AER: Insights 2021):

> *Treated households are 0.2 SD more food secure than the control group by 18 months and 0.25 SD by year three ($p < 0.01$). This effect grows to 0.4 SD by year seven and remains at 0.13 SD by year ten ($p < 0.05$).*

Alternative phrasing:

> *Our data are compatible with effect sizes ranging from + 0.004 SD to + 0.25 SD.*

Padmé's Take Home Message



Padmé's Take Home Message

Padmé’s Take Home Message

Padmé’s Take Home Message

Thank You for Your Time!