Interpreting and Critiquing Causal Evidence

Day 5 - Constructive Critiques

Jonathan Phillips

January 12, 2024

Section 1

Constructive Critiques

Being Constructive

- Effective critiques are essential to learning

Being Constructive

- Effective critiques are essential to learning
  - We have a scholarly obligation to point out errors in reasoning

Being Constructive

- Effective critiques are essential to learning
  - We have a scholarly obligation to point out errors in reasoning
  - We learn collectively by collaborating

Being Constructive

- Effective critiques are essential to learning
  - We have a scholarly obligation to point out errors in reasoning
  - We learn collectively by collaborating
  - We learn individually by thinking critically about others' work

Being Constructive

- Effective critiques are essential to learning
  - We have a scholarly obligation to point out errors in reasoning
  - We learn collectively by collaborating
  - We learn individually by thinking critically about others' work
- There is no research project that cannot be improved

Being Constructive

- But criticism can also be used as a weapon

Being Constructive

- But criticism can also be used as a weapon
  - To compete for attention/jobs

Being Constructive

- But criticism can also be used as a weapon
  - To compete for attention/jobs
  - To discourage colleagues

Being Constructive

- But criticism can also be used as a weapon
  - To compete for attention/jobs
  - To discourage colleagues
  - To assert status/hierarchy/superiority

Being Constructive

- But criticism can also be used as a weapon
  - To compete for attention/jobs
  - To discourage colleagues
  - To assert status/hierarchy/superiority
  - To destroy valuable research

Being Constructive

- But criticism can also be used as a weapon
  - To compete for attention/jobs
  - To discourage colleagues
  - To assert status/hierarchy/superiority
  - To destroy valuable research
  - To release our own frustrations

Being Constructive

- To avoid these risks, we must make our criticisms **constructive**

Being Constructive

- To avoid these risks, we must make our criticisms **constructive**
  1. In terms of style

Being Constructive

- To avoid these risks, we must make our criticisms **constructive**
  1. In terms of style
  2. In terms of content

Section 2

Constructive Style

Styles of Critique

- Your task is to convince the author to *improve* their work not to *abandon* it

Styles of Critique

- Your task is to convince the author to *improve* their work not to *abandon* it
- So they have to:
  - Understand your comment

Styles of Critique

- Your task is to convince the author to *improve* their work not to *abandon* it
- So they have to:
  - Understand your comment
  - Not take it as a personal 'ad hominem' attack

Styles of Critique

- Your task is to convince the author to *improve* their work not to *abandon* it
- So they have to:
  - Understand your comment
  - Not take it as a personal 'ad hominem' attack
  - Have options for how to respond

Styles of Critique

- Always remember your critique might be wrong!

Styles of Critique

- Always remember your critique might be wrong!
  - You always know the data less well than the author

Styles of Critique

- Always remember your critique might be wrong!
  - You always know the data less well than the author
  - Recognize the inherent challenges and constraints of implementing the research

Styles of Critique

- Always remember your critique might be wrong!
  - You always know the data less well than the author
  - Recognize the inherent challenges and constraints of implementing the research
- So phrase your comment in terms of ’as I understand your argument’

Styles of Critique

- Always remember your critique might be wrong!
  - You always know the data less well than the author
  - Recognize the inherent challenges and constraints of implementing the research
- So phrase your comment in terms of ’as I understand your argument’
- Or ’Could it be that something else is also happening?’

Styles of Critique

- **Be specific!** Which part of the research design is problematic?

Styles of Critique

- **Be specific!** Which part of the research design is problematic?
- **Be concrete!** Use an example/counterexample to communicate the risk

Styles of Critique

- **Be specific!** Which part of the research design is problematic?
- **Be concrete!** Use an example/counterexample to communicate the risk
- **Be objective!** We care about the research quality, not your personal opinion

Styles of Critique

- **Be specific!** Which part of the research design is problematic?
- **Be concrete!** Use an example/counterexample to communicate the risk
- **Be objective!** We care about the research quality, not your personal opinion
- **Suggest an alternative**

Styles of Critique

- Depersonalize criticism

Styles of Critique

- Depersonalize criticism
  - Instead of "you did it wrong...", refer to "in this type of research there is a risk..."

Styles of Critique

- Depersonalize criticism
  - Instead of "you did it wrong...", refer to "in this type of research there is a risk..."
  - "I feel like there might be some readers who did not understand..."

Styles of Critique

- Depersonalize criticism
  - Instead of "you did it wrong...", refer to "in this type of research there is a risk..."
  - "I feel like there might be some readers who did not understand..."
- If in doubt, use the feedback sandwich:

Styles of Critique

- Depersonalize criticism
  - Instead of "you did it wrong...", refer to "in this type of research there is a risk..."
  - "I feel like there might be some readers who did not understand..."
- If in doubt, use the feedback sandwich:
  1. Something positive/encouraging
  2. Critique
  3. Something positive/encouraging

Styles of Critique

- Finally:

Styles of Critique

- Finally:
  - Is the comment really necessary?

Styles of Critique

- Finally:
  - Is the comment really necessary?
  - If it is a minor issue, is there a better way to communciate it?

Styles of Critique

- Finally:
  - Is the comment really necessary?
  - If it is a minor issue, is there a better way to communciate it?
  - If you have not fully understood, take time to invest in understanding it before commenting

Section 3

Constructive Content

Strengthening Causal Arguments

1. Multiple tests of theory

Strengthening Causal Arguments

1. Multiple tests of theory
2. Multiple methods

Strengthening Causal Arguments

1. Multiple tests of theory
2. Multiple methods
3. Uncovering ’hidden’ units

Strengthening Causal Arguments

1. Multiple tests of theory
2. Multiple methods
3. Uncovering ’hidden’ units
4. Heterogeneity tests

Strengthening Causal Arguments

1. Multiple tests of theory
2. Multiple methods
3. Uncovering ’hidden’ units
4. Heterogeneity tests
5. Placebo tests

Strengthening Causal Arguments

1. Multiple tests of theory
2. Multiple methods
3. Uncovering ’hidden’ units
4. Heterogeneity tests
5. Placebo tests
6. Investigating Mechanisms

1. Multiple Tests

- Learning requires testing theories with evidence

1. Multiple Tests

- Learning requires testing theories with evidence
- Competing theories have **multiple distinct implications**

1. Multiple Tests

- Learning requires testing theories with evidence
- Competing theories have **multiple distinct implications**
- We should test *all* of these implications
- "The dog that didn’t bark"

1. Multiple Tests

- Learning requires testing theories with evidence
- Competing theories have **multiple distinct implications**
- We should test *all* of these implications
- "The dog that didn’t bark"
  - If a theory has multiple implications, and we find one of these did not occur in reality, we have evidence against our theory

1. Multiple Tests

- Learning requires testing theories with evidence
- Competing theories have **multiple distinct implications**
- We should test *all* of these implications
- "The dog that didn't bark"
  - If a theory has multiple implications, and we find one of these did not occur in reality, we have evidence against our theory
- **Critical tests**: Ideally we want to focus on those tests that 'separate' theories, telling us which one is true

1. Multiple Tests

- For example, Deaton argues poor health causes low income

1. Multiple Tests

- For example, Deaton argues poor health causes low income
  - But the critique is of possible reverse causation: Low income causes poor health.

1. Multiple Tests

- For example, Deaton argues poor health causes low income
  - But the critique is of possible reverse causation: Low income causes poor health.
- Additional tests to support his argument include:

1. Multiple Tests

- For example, Deaton argues poor health causes low income
  - But the critique is of possible reverse causation: Low income causes poor health.
- Additional tests to support his argument include:
  1. Whether the relationship falls after retirement

1. Multiple Tests

- For example, Deaton argues poor health causes low income
  - But the critique is of possible reverse causation: Low income causes poor health.
- Additional tests to support his argument include:
  1. Whether the relationship falls after retirement
  2. Whether the relationship is weaker among women, who on average work fewer hours

1. Multiple Tests

- For example, Deaton argues poor health causes low income
  - But the critique is of possible reverse causation: Low income causes poor health.
- Additional tests to support his argument include:
  1. Whether the relationship falls after retirement
  2. Whether the relationship is weaker among women, who on average work fewer hours
  3. Whether the relationship holds even for diseases which could easily be cured with more income

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions
- And helps us identify *all* of the implications of our theory to test

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions
- And helps us identify *all* of the implications of our theory to test
- For example, the Acemoglu-Robinson model predicts democratization is more likely when:
  1. Inequality is moderately high

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions
- And helps us identify *all* of the implications of our theory to test
- For example, the Acemoglu-Robinson model predicts democratization is more likely when:
  1. Inequality is moderately high
  2. The poor are less frequently able to coordinate

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions
- And helps us identify *all* of the implications of our theory to test
- For example, the Acemoglu-Robinson model predicts democratization is more likely when:
  1. Inequality is moderately high
  2. The poor are less frequently able to coordinate
  3. Revolution is more costly

1. Multiple Tests

- Formal models are particularly useful for developing multiple tests
- We can debate theory all day long as it's based on our intuition and experience
- Formal theory forces us to be explicit about our assumptions
- And helps us identify *all* of the implications of our theory to test
- For example, the Acemoglu-Robinson model predicts democratization is more likely when:
  1. Inequality is moderately high
  2. The poor are less frequently able to coordinate
  3. Revolution is more costly
  4. When the poor care less about the future

2. Multiple Methods

- Our methodologies' assumptions are often impossible to test in quantitative data

2. Multiple Methods

- Our methodologies' assumptions are often impossible to test in quantitative data
- But qualitative evidence can help justify our assumptions

2. Multiple Methods

- Our methodologies' assumptions are often impossible to test in quantitative data
- But qualitative evidence can help justify our assumptions
- This means using multiple (mixed) methods

2. Multiple Methods

- Our methodologies' assumptions are often impossible to test in quantitative data
- But qualitative evidence can help justify our assumptions
- This means using multiple (mixed) methods
- These are all "Causal Process Observations" (Collier et al 2010)

2. Multiple Methods

- Examples of Causal Process Observations:

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?
  - Is selection bias possible in the data? How was the dataset constructed?

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?
  - Is selection bias possible in the data? How was the dataset constructed?
  - Was randomization of a field experiment successful? We can interview people about the process, gather documents

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?
  - Is selection bias possible in the data? How was the dataset constructed?
  - Was randomization of a field experiment successful? We can interview people about the process, gather documents
  - Are there spillovers (violations of SUTVA)? We can conduct a survey and find out how people interacted

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?
  - Is selection bias possible in the data? How was the dataset constructed?
  - Was randomization of a field experiment successful? We can interview people about the process, gather documents
  - Are there spillovers (violations of SUTVA)? We can conduct a survey and find out how people interacted
  - Is a regression discontinuity threshold enforced neutrally? Or was the threshold chosen to make sure a particular unit passed?

2. Multiple Methods

- Examples of Causal Process Observations:
  - What is the treatment assignment mechanism? Who decides when teachers get evaluated?
  - Is selection bias possible in the data? How was the dataset constructed?
  - Was randomization of a field experiment successful? We can interview people about the process, gather documents
  - Are there spillovers (violations of SUTVA)? We can conduct a survey and find out how people interacted
  - Is a regression discontinuity threshold enforced neutrally? Or was the threshold chosen to make sure a particular unit passed?
  - Can people sort/migrate across a discontinuity? We can use administrative data on migration rates to assess if these differences might be large enough to explain our results

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas
- Qualitative evidence helps:

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas
- Qualitative evidence helps:
  - To verify that slaves usually sailed from the nearest port, and not a different country

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas
- Qualitative evidence helps:
  - To verify that slaves usually sailed from the nearest port, and not a different country
  - To inform the need for extra controls, eg. for legal system, natural resources

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas
- Qualitative evidence helps:
  - To verify that slaves usually sailed from the nearest port, and not a different country
  - To inform the need for extra controls, eg. for legal system, natural resources
  - To identify the direction of the bias - reverse causation is less of a problem because he shows the richest ethnic groups were most affected by the slave trade

2. Multiple Methods

- Nunn (2008) asks whether the slave trade explains underdevelopment in parts of Africa, using an instrumental variable for distance to the Americas
- Qualitative evidence helps:
  - To verify that slaves usually sailed from the nearest port, and not a different country
  - To inform the need for extra controls, eg. for legal system, natural resources
  - To identify the direction of the bias - reverse causation is less of a problem because he shows the richest ethnic groups were most affected by the slave trade
  - To argue in support of the exclusion restriction for the instrumental variable: that plantations were set up in the Carribean because of the climate, not because they were near the supply of slaves in West Africa

3. Uncovering ’Hidden’ Units

- What if our sample size is small, there is no experimental option and selection bias is a concern?

3. Uncovering ’Hidden’ Units

- What if our sample size is small, there is no experimental option and selection bias is a concern?
- One strategy is to identify more units:

3. Uncovering ’Hidden’ Units

- What if our sample size is small, there is no experimental option and selection bias is a concern?
- One strategy is to identify more units:
  - Those units are often ’hidden’, either because we did not think about them or there is no data on them initially

3. Uncovering ’Hidden’ Units

- What if our sample size is small, there is no experimental option and selection bias is a concern?
- One strategy is to identify more units:
  - Those units are often ’hidden’, either because we did not think about them or there is no data on them initially
  - We can expand our dataset and adjust our research question

3. Uncovering ’Hidden’ Units

- Lewis (2016) wanted to ask whether ethnicity affected the formation of rebel groups

3. Uncovering ’Hidden’ Units

- Lewis (2016) wanted to ask whether ethnicity affected the formation of rebel groups
- But there is a selection/survival bias in the data - we only have data on the groups that *succeeded*

3. Uncovering 'Hidden' Units

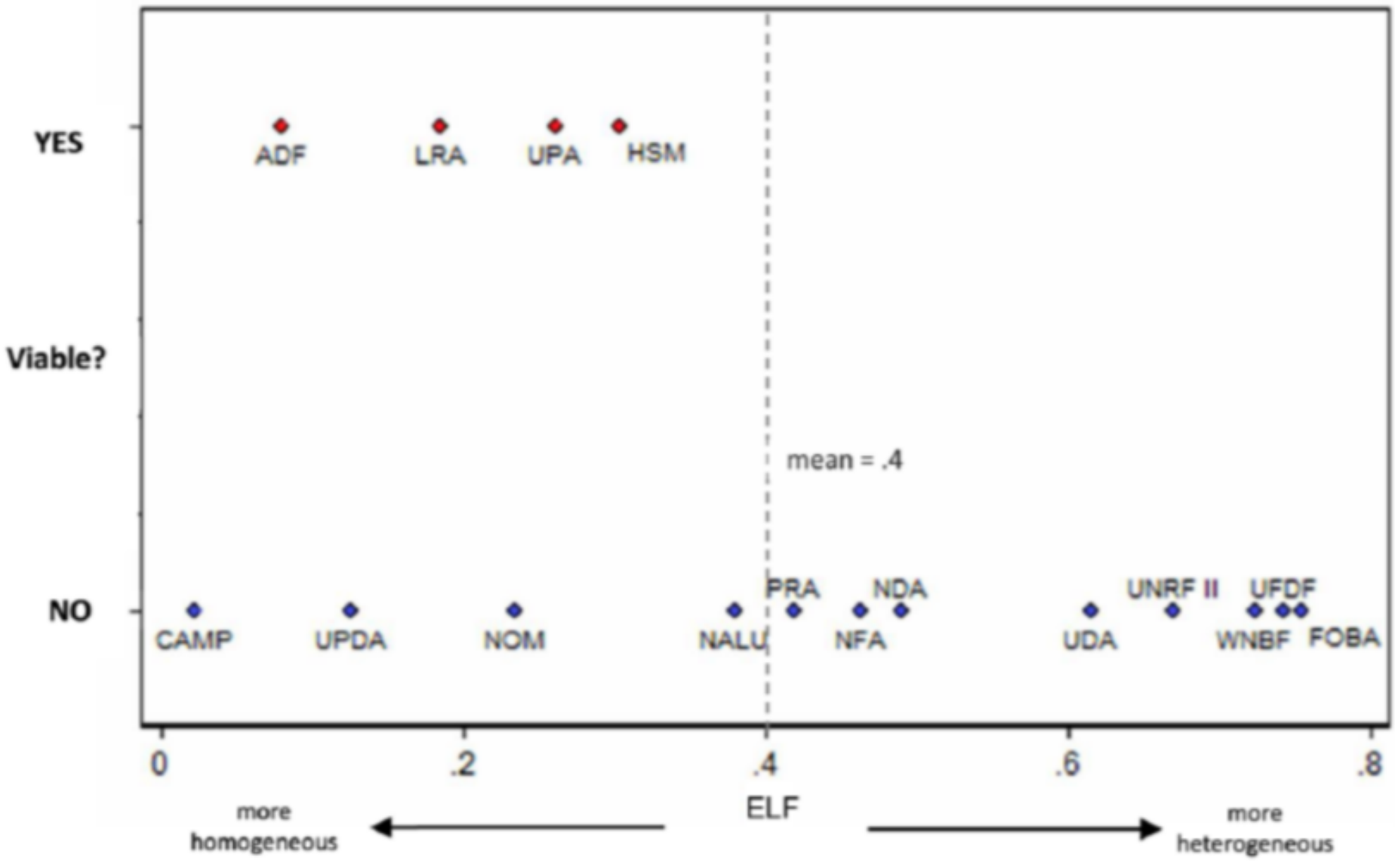
- Lewis (2016) wanted to ask whether ethnicity affected the formation of rebel groups
- But there is a selection/survival bias in the data - we only have data on the groups that *succeeded*
- She collected data from Uganda on *all* rebel groups

3. Uncovering 'Hidden' Units

- Lewis (2016) wanted to ask whether ethnicity affected the formation of rebel groups
- But there is a selection/survival bias in the data - we only have data on the groups that *succeeded*
- She collected data from Uganda on *all* rebel groups
  - Expanding the sample from 1-4 (in most datasets) to 15

3. Uncovering 'Hidden' Units

- Lewis (2016) wanted to ask whether ethnicity affected the formation of rebel groups
- But there is a selection/survival bias in the data - we only have data on the groups that *succeeded*
- She collected data from Uganda on *all* rebel groups
  - Expanding the sample from 1-4 (in most datasets) to 15
  - Showing that ethnicity does *not* affect rebel group formation, but may affect their success


YES
Viable?
NO
ADF
LRA
UPA
HSM
mean = .4
PRA
NDA
UNRF II
UFDF
CAMP
UPDA
NOM
NALU
NFA
UDA
WNBF
FOBA
0
.2
.4
.6
.8
ELF
more homogeneous
more heterogeneous

4. Heterogeneity Tests

- Sometimes the implications of theory are very precise

4. Heterogeneity Tests

- Sometimes the implications of theory are very precise
- Treatment is likely to have affected subgroups to different degrees

4. Heterogeneity Tests

- Sometimes the implications of theory are very precise
- Treatment is likely to have affected subgroups to different degrees
- We can use heterogeneity tests to disaggregate the effect to each subgroup and compare

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption
- Their theory is that this is produced by 'electoral accountability'

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption
- Their theory is that this is produced by 'electoral accountability'
- They provide evidence for this specific theory by:

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption
- Their theory is that this is produced by 'electoral accountability'
- They provide evidence for this specific theory by:
  - Subsetting the data to only those municipalities with local radio stations which broadcast the findings of corruption and showing the effect is much *stronger*

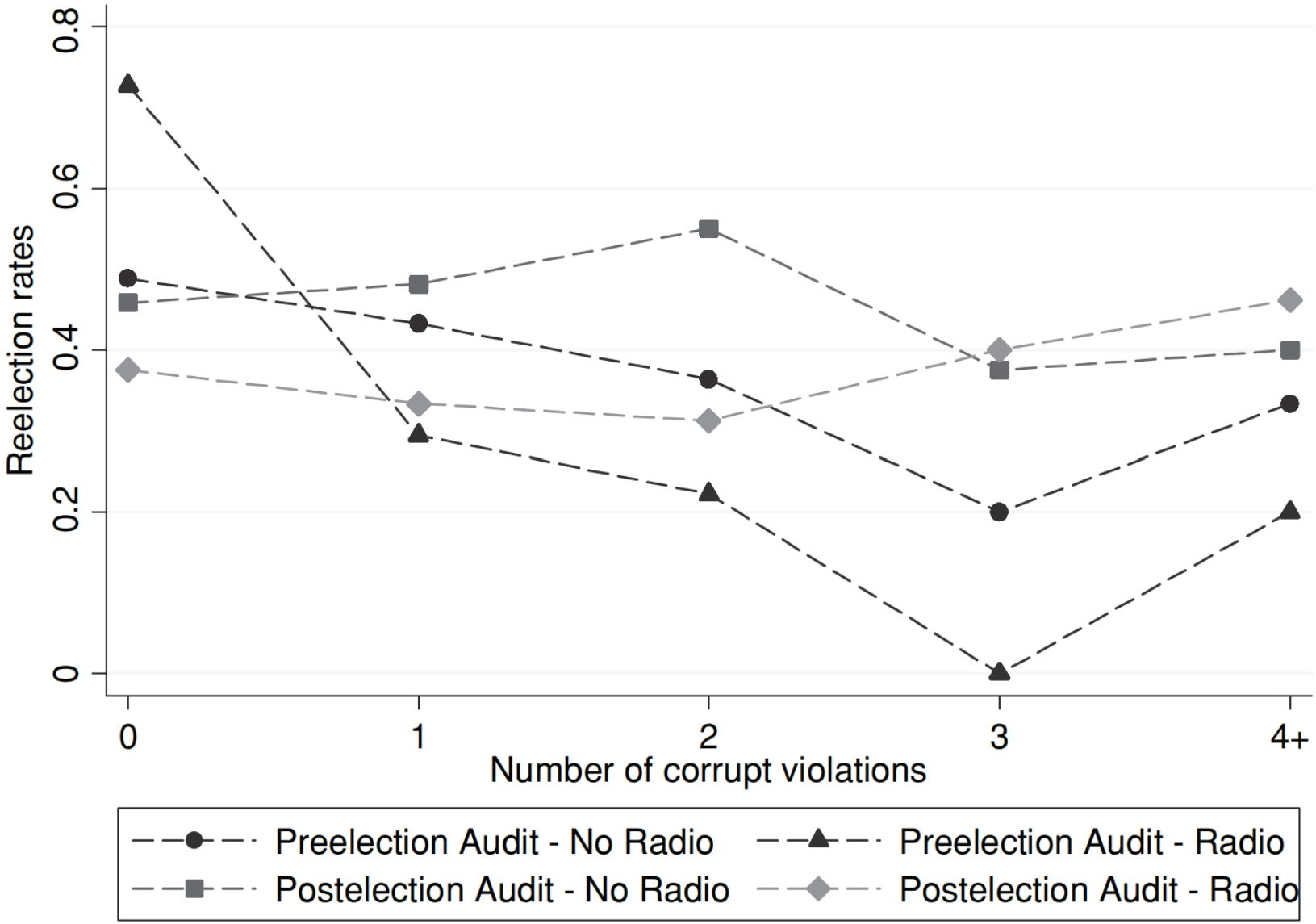
4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption
- Their theory is that this is produced by 'electoral accountability'
- They provide evidence for this specific theory by:
  - Subsetting the data to only those municipalities with local radio stations which broadcast the findings of corruption and showing the effect is much *stronger*
  - Subsetting the data to only those municipalities with mayors in their first-term who face re-election incentives, showing corruption is *lower*

4. Heterogeneity Tests

- For example, Ferraz and Finan (2008) ask how random audits affect corruption rates
- They find that audits significantly reduce corruption
- Their theory is that this is produced by 'electoral accountability'
- They provide evidence for this specific theory by:
  - Subsetting the data to only those municipalities with local radio stations which broadcast the findings of corruption and showing the effect is much *stronger*
  - Subsetting the data to only those municipalities with mayors in their first-term who face re-election incentives, showing corruption is *lower*
- What other theory would be consistent with *all* of this evidence?

4. Heterogeneity Tests



5. Placebo tests

- Our theory has very precise implications, and we normally test the 'positive' version

5. Placebo tests

- Our theory has very precise implications, and we normally test the 'positive' version
- But we can also test the 'non-predictions' of our theory, when there should *not* be an effect

5. Placebo tests

- Our theory has very precise implications, and we normally test the 'positive' version
- But we can also test the 'non-predictions' of our theory, when there should *not* be an effect
- If we found an effect where there should *not* be one, we might think something is weird in our data/methodology and have less confidence in our main result
  - Placebo Treatments
  - Placebo Outcomes
  - Placebo Populations

5. Placebo tests

- For example, with a regression discontinuity on close elections we expect a ’jump’ effect when elections are tied (winning margin=0)

5. Placebo tests

- For example, with a regression discontinuity on close elections we expect a ’jump’ effect when elections are tied (winning margin=0)
- We expect there *not* to be a ’jump’ effect when winning margin=10%

5. Placebo tests

- For example, with a regression discontinuity on close elections we expect a ’jump’ effect when elections are tied (winning margin=0)
- We expect there *not* to be a ’jump’ effect when winning margin=10%
- So we can apply our regression discontinuity again and measure the effect at winning margin=10%

5. Placebo tests

- For example, with a regression discontinuity on close elections we expect a ’jump’ effect when elections are tied (winning margin=0)
- We expect there *not* to be a ’jump’ effect when winning margin=10%
- So we can apply our regression discontinuity again and measure the effect at winning margin=10%
- If we still find an effect, there might be something wrong with our data/method

5. Placebo tests

FIGURE 4. Local Linear Regression: Difference in Resistance Activity

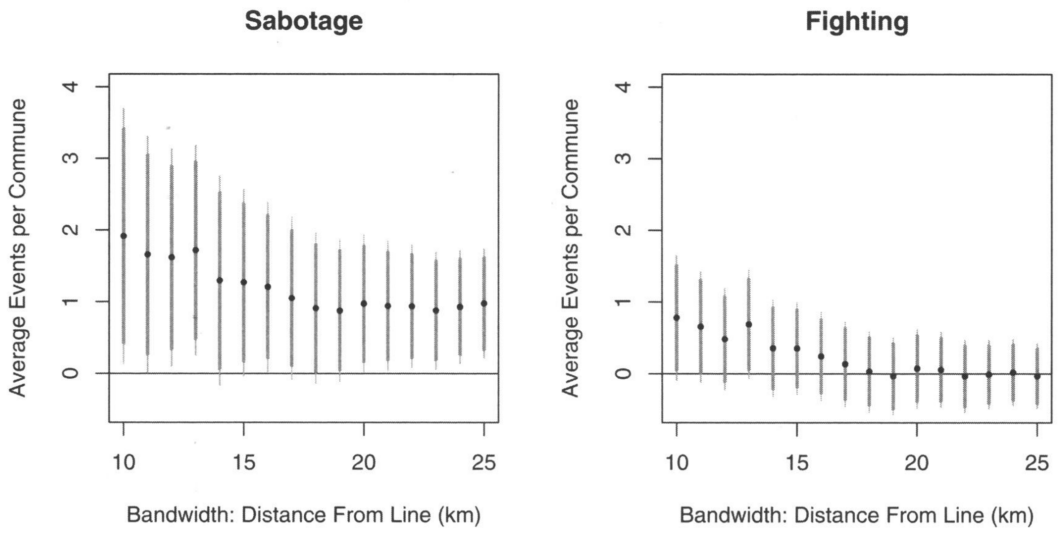
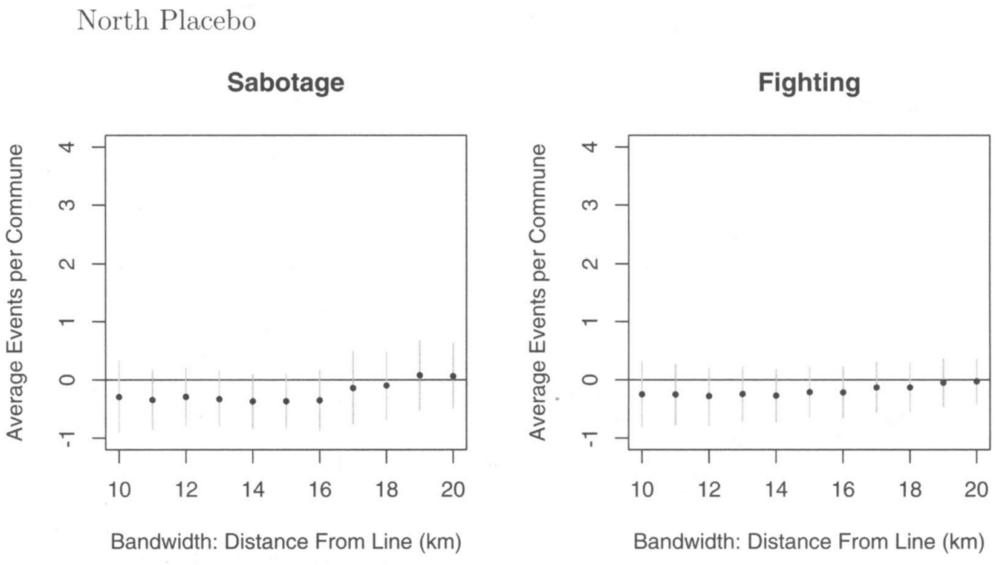


FIGURE 5. Robustness Check: Placebo



5. Placebo tests

- The same with difference-in-differences

5. Placebo tests

- The same with difference-in-differences
- If we were estimating the effect of a treatment that applied to some units on 5th August 2012, we expect no effect on 3rd July 2009

5. Placebo tests

- The same with difference-in-differences
- If we were estimating the effect of a treatment that applied to some units on 5th August 2012, we expect no effect on 3rd July 2009
  - Or on 4th August 2012

5. Placebo tests

- The same with difference-in-differences
- If we were estimating the effect of a treatment that applied to some units on 5th August 2012, we expect no effect on 3rd July 2009
  - Or on 4th August 2012
  - Or on 6th August 2012

5. Placebo tests

- The same with difference-in-differences
- If we were estimating the effect of a treatment that applied to some units on 5th August 2012, we expect no effect on 3rd July 2009
  - Or on 4th August 2012
  - Or on 6th August 2012
- The more tightly the data are consistent *only* with your theory, the more credible is your theory

5. Placebo tests

- Placebo tests also work for small-N studies (Glynn and Ichino 2012)

5. Placebo tests

- Placebo tests also work for small-N studies (Glynn and Ichino 2012)
- We want to assess the effect of presidentialism on reducing party cohesion

5. Placebo tests

- Placebo tests also work for small-N studies (Glynn and Ichino 2012)
- We want to assess the effect of presidentialism on reducing party cohesion
- A good comparison is between the USA (presidential) and Canada (parliamentary)

5. Placebo tests

- Placebo tests also work for small-N studies (Glynn and Ichino 2012)
- We want to assess the effect of presidentialism on reducing party cohesion
- A good comparison is between the USA (presidential) and Canada (parliamentary)
- But we also gain confidence if we can show that other similar parliamentary systems have cohesive parties (Britain, Australia, etc.)

6. Mechanisms

- Often we talk as though we are testing ’treatments’

6. Mechanisms

- Often we talk as though we are testing 'treatments'
  - But that leaves an empty black box between treatment and outcome

6. Mechanisms

- Often we talk as though we are testing 'treatments'
  - But that leaves an empty black box between treatment and outcome
- Really we want to test **theories**, which include a clear mechanism connecting the treatment and the outcome

6. Mechanisms

- Often we talk as though we are testing 'treatments'
  - But that leaves an empty black box between treatment and outcome
- Really we want to test **theories**, which include a clear mechanism connecting the treatment and the outcome
- To show that a specific theory is operating, we want to trace every step of the mechanism

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*
- Habyarimana et al (2007) asked "why?"

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*
- Habyarimana et al (2007) asked "why?"
  - Preferences
  - Technology
  - Strategy selection

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*
- Habyarimana et al (2007) asked "why?"
  - Preferences
  - Technology
  - Strategy selection
- They designed laboratory games to test exactly each mechanism

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*
- Habyarimana et al (2007) asked "why?"
  - Preferences
  - Technology
  - Strategy selection
- They designed laboratory games to test exactly each mechanism
- Eg. To test if there is an ethnic 'technology' that helps co-ethnics, they asked Ugandans to find a specific person in a neighbourhood, and paid them a reward if they did

6. Mechanisms

- For example, multiple studies show a clear treatment effect: high ethnic diversity reduces public goods provision
  - But these studies had *no theory*
- Habyarimana et al (2007) asked "why?"
  - Preferences
  - Technology
  - Strategy selection
- They designed laboratory games to test exactly each mechanism
- Eg. To test if there is an ethnic 'technology' that helps co-ethnics, they asked Ugandans to find a specific person in a neighbourhood, and paid them a reward if they did
  - Co-ethnics found their target 43% of the time, non-co-ethnics only 28% of the time

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis
- We turn our single case into **multiple** observations - usually over time

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis
- We turn our single case into **multiple** observations - usually over time
- We test if those process observations are consistent with our theory

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis
- We turn our single case into **multiple** observations - usually over time
- We test if those process observations are consistent with our theory
- But what happened to counterfactuals here?

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis
- We turn our single case into **multiple** observations - usually over time
- We test if those process observations are consistent with our theory
- But what happened to counterfactuals here?
- We're substituting assumptions/theory for a counterfactual

6. Mechanisms

- Process Tracing is one way of demonstrating which mechanism connected the treatment and the outcome
- Within-case analysis
- We turn our single case into **multiple** observations - usually over time
- We test if those process observations are consistent with our theory
- But what happened to counterfactuals here?
- We're substituting assumptions/theory for a counterfactual
- Provides evidence for our specific case; generalization is hard

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:
  - There were only 10 minutes until the polling stations closed

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:
  - There were only 10 minutes until the polling stations closed
  - Only about 20% would have heard the announcements

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:
  - There were only 10 minutes until the polling stations closed
  - Only about 20% would have heard the announcements
  - Around half were Bush voters, who may also have stayed home

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:
  - There were only 10 minutes until the polling stations closed
  - Only about 20% would have heard the announcements
  - Around half were Bush voters, who may also have stayed home
  - Voters still had a reason to vote for other offices

6. Mechanisms

- Brady (2004) provides an example of a type of process tracing to evaluate the plausibility of a difference-in-differences research design
- Difference-in-differences analysis suggested media announcements that Al Gore won Florida in 2000 caused 10,000 Gore voters to stay at home, allowing Bush to win.
- But:
  - There were only 10 minutes until the polling stations closed
  - Only about 20% would have heard the announcements
  - Around half were Bush voters, who may also have stayed home
  - Voters still had a reason to vote for other offices
- Brady estimates that at most 224 people did not vote due to the media announcements