FLS 6441 - Methods III: Explanation and Causation

Week 5 - Natural Experiments

Jonathan Phillips

April 2019

Classification of Research Designs

	Independence of Treatment Assignment?	**Researcher Controls Treatment Assignment?**
Controlled Experiments	✓	✓
Natural Experiments	✓	
Observational Studies		

Classification of Research Designs

		Independence of Treatment Assignment	**Researcher Controls Treatment Assignment?**
Controlled Experiments	Field Experiments	✓	✓
	Survey and Lab Experiments	✓	✓
Natural Experiments	Natural Experiments	✓	
	Instrumental Variables	✓	
	Discontinuities	✓	
Observational Studies	Difference-in-Differences		
	Controlling for Confounding		
	Matching		
	Comparative Cases and Process Tracing		

Section 1

Natural Experiments

Natural Experiments

Advantages:

- We don’t need to run our own experiment! (Too expensive, unethical or politically impossible)

Natural Experiments

Advantages:

- We don't need to run our own experiment! (Too expensive, unethical or politically impossible)
- Still have independence of potential outcomes from treatment

Natural Experiments

Advantages:

- We don’t need to run our own experiment! (Too expensive, unethical or politically impossible)
- Still have independence of potential outcomes from treatment
- Treatment may be more ’realistic’ than in a controlled experiment

Natural Experiments

Advantages:

- We don’t need to run our own experiment! (Too expensive, unethical or politically impossible)
- Still have independence of potential outcomes from treatment
- Treatment may be more ’realistic’ than in a controlled experiment

Disadvantages:

- We can never be sure randomization really worked

Natural Experiments

Advantages:

- We don’t need to run our own experiment! (Too expensive, unethical or politically impossible)
- Still have independence of potential outcomes from treatment
- Treatment may be more ’realistic’ than in a controlled experiment

Disadvantages:

- We can never be sure randomization really worked
- We don’t get to choose the treatments we want to evaluate, just ’discover’ them

Natural Experiments

Advantages:

- We don’t need to run our own experiment! (Too expensive, unethical or politically impossible)
- Still have independence of potential outcomes from treatment
- Treatment may be more ’realistic’ than in a controlled experiment

Disadvantages:

- We can never be sure randomization really worked
- We don’t get to choose the treatments we want to evaluate, just ’discover’ them
- We don’t get to choose the population and sample

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables
  2. **Causal Process Observations**

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables
  2. **Causal Process Observations**
     - Documents/code/video evidence

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables
  2. **Causal Process Observations**
     - Documents/code/video evidence
     - Interviews with eyewitnesses

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables
  2. **Causal Process Observations**
     - Documents/code/video evidence
     - Interviews with eyewitnesses
     - Verifying treatment assignment matches documents

Verifying Randomization

- If it's an important treatment, *someone* had an incentive to try and alter it
- The burden of proof is on us: How can we increase confidence that assignment was (as-if) random?
- Two strategies:
  1. Check balance on lots of variables
     - Especially variables that are potential omitted variables
  2. **Causal Process Observations**
     - Documents/code/video evidence
     - Interviews with eyewitnesses
     - Verifying treatment assignment matches documents
     - Identify risks of reverse causation, omitted variables, (Self-)selection

Verifying Randomization

- How does Snow argue that households' assignment to water company is as-if random?

Verifying Randomization

- How does Snow argue that households' assignment to water company is as-if random?
- How do we know that Brazil's municipal audits are random?

The Problem of Not Controlling Treatment Assignment

- “Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect.” (Sekhon and Titunik 2012)

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)
     - We can only compare those units *that were part of the original randomization*

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)
     - We can only compare those units *that were part of the original randomization*
  2. That the treatment is the factor we actually want to study

The Problem of Not Controlling Treatment Assignment

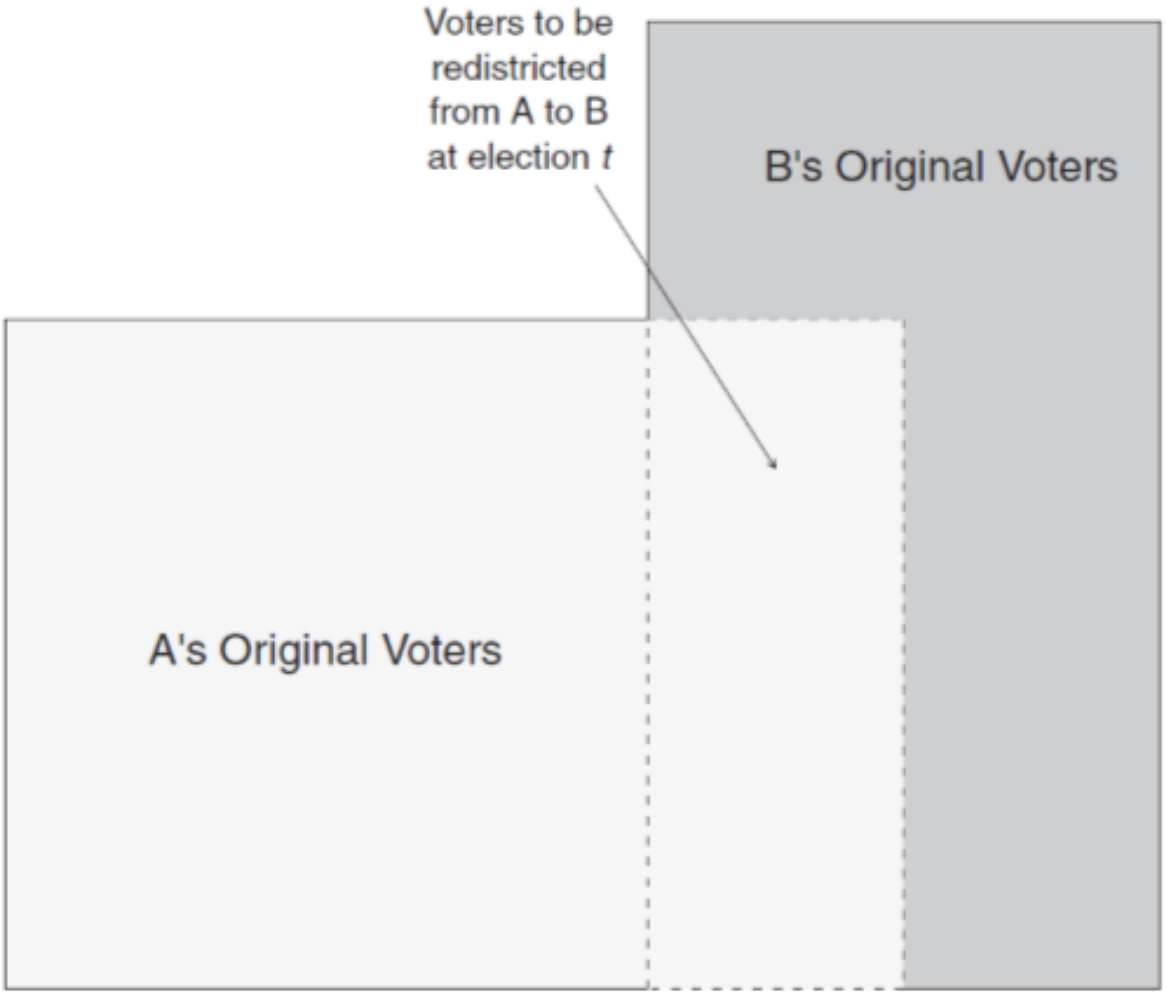
- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)
     - We can only compare those units *that were part of the original randomization*
  2. That the treatment is the factor we actually want to study
     - We have to 'interpret' the treatment

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)
     - We can only compare those units *that were part of the original randomization*
  2. That the treatment is the factor we actually want to study
     - We have to 'interpret' the treatment
     - Sometimes treatments are 'bundles'

The Problem of Not Controlling Treatment Assignment

- "Random assignment of the intervention is not sufficient to provide an unbiased estimate of the causal effect." (Sekhon and Titunik 2012)
- Treatment and control groups are defined *after* randomization - it's our responsibility to make sure:
  1. **These two groups actually are comparable** (POs are independent of treatment)
     - We can only compare those units *that were part of the original randomization*
  2. That the treatment is the factor we actually want to study
     - We have to 'interpret' the treatment
     - Sometimes treatments are 'bundles'
     - Sometimes treatments are 'repeated', creating interactions or changing expectations


Voters to be
redistricted
from A to B
at election t
B's Original Voters
A's Original Voters

The Problem of Not Controlling Treatment Assignment

	A's Original Voters	**Switched Voters**	**B's Original Voters**
2000 election context		Same	Same
Duration of exposure to incumbent in district B		4 years	10 years
1996 and prior election context	Same	Same	

The Problem of Not Controlling Treatment Assignment

	A's Original Voters vs. Switched Voters	**B's Original Voters vs. Switched Voters**
Potential Outcomes Independent of Treatment Assignment?	Yes	No
What is 'Treatment'?	Different election context, different candidates	Difference in duration of exposure to incumbent

Section 2

Randomized Natural Experiments

Ferraz and Finan (2008)

- Do voters punish corrupt politicians?

Ferraz and Finan (2008)

- Do voters punish corrupt politicians?
- Corruption is hard to manipulate (ethically)

Ferraz and Finan (2008)

- Do voters punish corrupt politicians?
- Corruption is hard to manipulate (ethically)
- We can also look at voters' *information* about corruption

Ferraz and Finan (2008)

- **Population:** Brazilian municipalities with population less than 450,000
- **Sample:** 373 Municipalities with audits either side of 2004 elections and first-term mayors
- **Treatment:** CGU Audit before election
- **Control:** Audit after election
- **Treatment Assignment Mechanism:** Randomized (Caixa)
- **Outcome:** Vote Share for the Incumbent

Ferraz and Finan (2008)

- Methodology
  - $VS_{ms} = \alpha + \beta \text{Audited Early}_{ms} + X_{ms} + FE_s + \epsilon_{ms}$

Ferraz and Finan (2008)

- Methodology
  - $VS_{ms} = \alpha + \beta \text{Audited Early}_{ms} + X_{ms} + FE_s + \epsilon_{ms}$
  - Result: No Effect

Ferraz and Finan (2008)

- The importance of a theoretical model:

Ferraz and Finan (2008)

- The importance of a theoretical model:
- Treatment is the release of information, but the *theory* they seek to test is when voters learn something about candidates

Ferraz and Finan (2008)

- The importance of a theoretical model:
- Treatment is the release of information, but the *theory* they seek to test is when voters learn something about candidates
- So we need treatment and control groups reflecting the theory

Ferraz and Finan (2008)

- The importance of a theoretical model:
- Treatment is the release of information, but the *theory* they seek to test is when voters learn something about candidates
- So we need treatment and control groups reflecting the theory
- Voters'*priors* about the candidate's corruption vary

Ferraz and Finan (2008)

- The importance of a theoretical model:
- Treatment is the release of information, but the *theory* they seek to test is when voters learn something about candidates
- So we need treatment and control groups reflecting the theory
- Voters'*priors* about the candidate's corruption vary
- And the *content* of the information varies

Ferraz and Finan (2008)

- The importance of a theoretical model:
- Treatment is the release of information, but the *theory* they seek to test is when voters learn something about candidates
- So we need treatment and control groups reflecting the theory
- Voters'*priors* about the candidate's corruption vary
- And the *content* of the information varies
- It's the interaction of expectations and information content that matters

Ferraz and Finan (2008)

- Methodology
  - So expected results are *conditional on content of the audit report*

Ferraz and Finan (2008)

- Methodology
  - So expected results are *conditional on content of the audit report*
  - $VS_{ms} = \alpha + \beta \text{Audited Early}_{ms} + \beta_2 \text{Corruption}_{ms} + \beta_3 \text{Audited Early}_{ms} * \text{Corruption}_{ms} + X_{ms} + \text{FE}_s + \epsilon_{ms}$

Ferraz and Finan (2008)

- Results

Ferraz and Finan (2008)

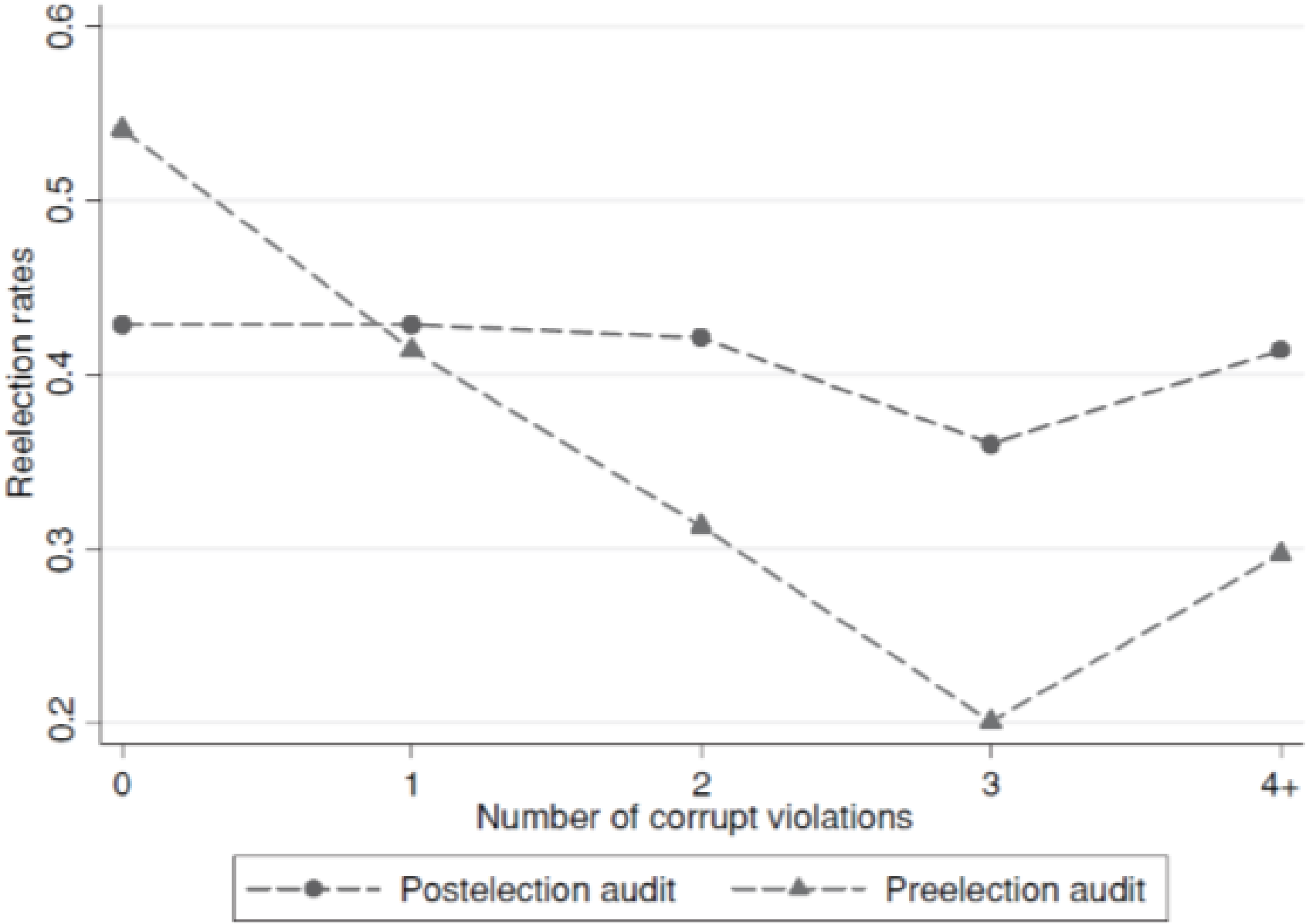
- Results
  - Strong corruption information (2 violations) reduces re-election by 7% points

Ferraz and Finan (2008)

- Results
  - Strong corruption information (2 violations) reduces re-election by 7% points
  - Stronger corruption information (3 violations) reduces re-election by 14% points

Ferraz and Finan (2008)

- Results
  - Strong corruption information (2 violations) reduces re-election by 7% points
  - Stronger corruption information (3 violations) reduces re-election by 14% points
  - Strong corruption information (2 violations) with local radio reduces re-election by 11% points


Reelection rates
0.2
0.3
0.4
0.5
0.6
0
1
2
3
4+
Number of corrupt violations
Postelection audit
Preelection audit

Section 3

Non-Randomized Natural Experiments

Non-Randomized Natural Experiments

- How can we achieve causal inference without randomization?

Non-Randomized Natural Experiments

- How can we achieve causal inference without randomization?
- Our assumption is always "The Treatment Assignment Mechanism is independent of potential outcomes"

Non-Randomized Natural Experiments

- How can we achieve causal inference without randomization?
- Our assumption is always "The Treatment Assignment Mechanism is independent of potential outcomes"
- Can we find real-world treatment assignments that ignored potential outcomes?

Non-Randomized Natural Experiments

- How can we achieve causal inference without randomization?
- Our assumption is always "The Treatment Assignment Mechanism is independent of potential outcomes"
- Can we find real-world treatment assignments that ignored potential outcomes?
  - "As good as random", "As-if random"

Non-Randomized Natural Experiments

- There are good reasons to be skeptical: Humans are *strategic* and anticipate potential outcomes

Non-Randomized Natural Experiments

- There are good reasons to be skeptical: Humans are *strategic* and anticipate potential outcomes
- But sometimes they are trying to alter outcomes *different to the potential outcomes we care about*

Non-Randomized Natural Experiments

- There are good reasons to be skeptical: Humans are *strategic* and anticipate potential outcomes
- But sometimes they are trying to alter outcomes *different to the potential outcomes we care about*
  - If these outcomes are not correlated with (/'orthogonal to'/'independent of') our own potential outcomes, we might be okay

Non-Randomized Natural Experiments

- There are good reasons to be skeptical: Humans are *strategic* and anticipate potential outcomes
- But sometimes they are trying to alter outcomes *different to the potential outcomes we care about*
  - If these outcomes are not correlated with (/'orthogonal to'/'independent of') our own potential outcomes, we might be okay
  - But we cannot test this

Non-Randomized Natural Experiments

- There are good reasons to be skeptical: Humans are *strategic* and anticipate potential outcomes
- But sometimes they are trying to alter outcomes *different to the potential outcomes we care about*
  - If these outcomes are not correlated with (/'orthogonal to'/'independent of') our own potential outcomes, we might be okay
  - But we cannot test this
  - We have to rely on qualitative evidence of the treatment assignment mechanism

Posner (2004)

- **Hypothesis:** Cultural differences become political cleavages when the cultural groups are large portions of the population

Posner (2004)

- **Hypothesis:** Cultural differences become political cleavages when the cultural groups are large portions of the population
- **Treatment:** Smaller country (relative to size of ethnic group)

Posner (2004)

- **Hypothesis:** Cultural differences become political cleavages when the cultural groups are large portions of the population
- **Treatment:** Smaller country (relative to size of ethnic group)
- **Control:** Larger country

Posner (2004)

- **Hypothesis:** Cultural differences become political cleavages when the cultural groups are large portions of the population
- **Treatment:** Smaller country (relative to size of ethnic group)
- **Control:** Larger country
- **Potential Outcomes:** Degree of political conflict between ethnic groups in larger/smaller countries

Posner (2004)

- **Hypothesis:** Cultural differences become political cleavages when the cultural groups are large portions of the population
- **Treatment:** Smaller country (relative to size of ethnic group)
- **Control:** Larger country
- **Potential Outcomes:** Degree of political conflict between ethnic groups in larger/smaller countries
- **Treatment Assignment Mechanism:** African borders that cross ethnic group boundaries

Posner (2004)

- African colonial borders assigned people to be ’Zambian’ or ’Malawian’.

Posner (2004)

- African colonial borders assigned people to be 'Zambian' or 'Malawian'.
- Straight lines drawn with a ruler in Berlin

Posner (2004)

- African colonial borders assigned people to be ’Zambian’ or ’Malawian’.
- Straight lines drawn with a ruler in Berlin
- Little knowledge of local geography or populations

Posner (2004)

- African colonial borders assigned people to be 'Zambian' or 'Malawian'.
- Straight lines drawn with a ruler in Berlin
- Little knowledge of local geography or populations
- Zambia-Malawi border defined by geography: by the watershed of the hills

Posner (2004)

- African colonial borders assigned people to be 'Zambian' or 'Malawian'.
- Straight lines drawn with a ruler in Berlin
- Little knowledge of local geography or populations
- Zambia-Malawi border defined by geography: by the watershed of the hills
- Splitting the Chewa and Tumbuka groups

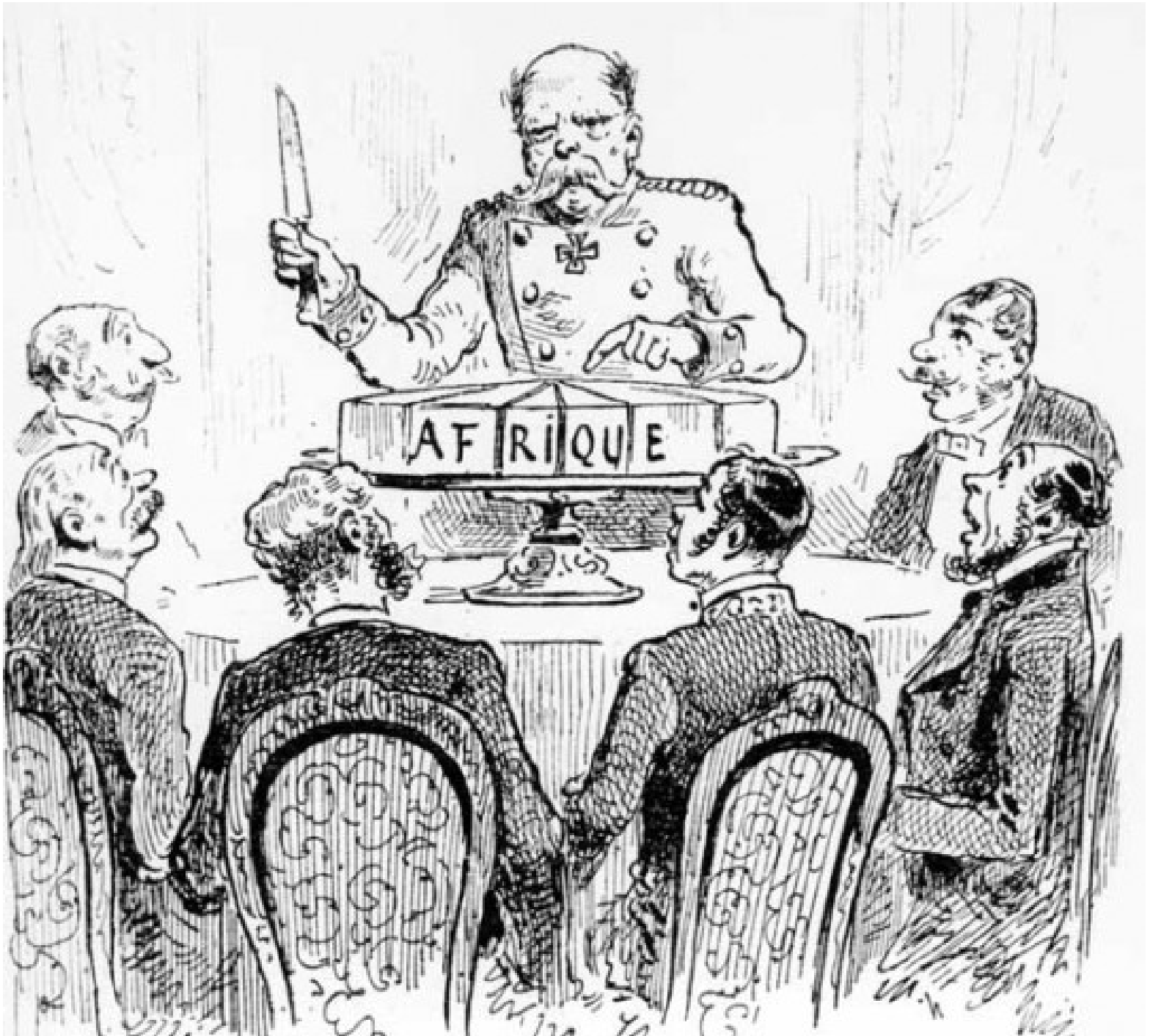

AFRIQUE

FIGURE 1. Research Sites

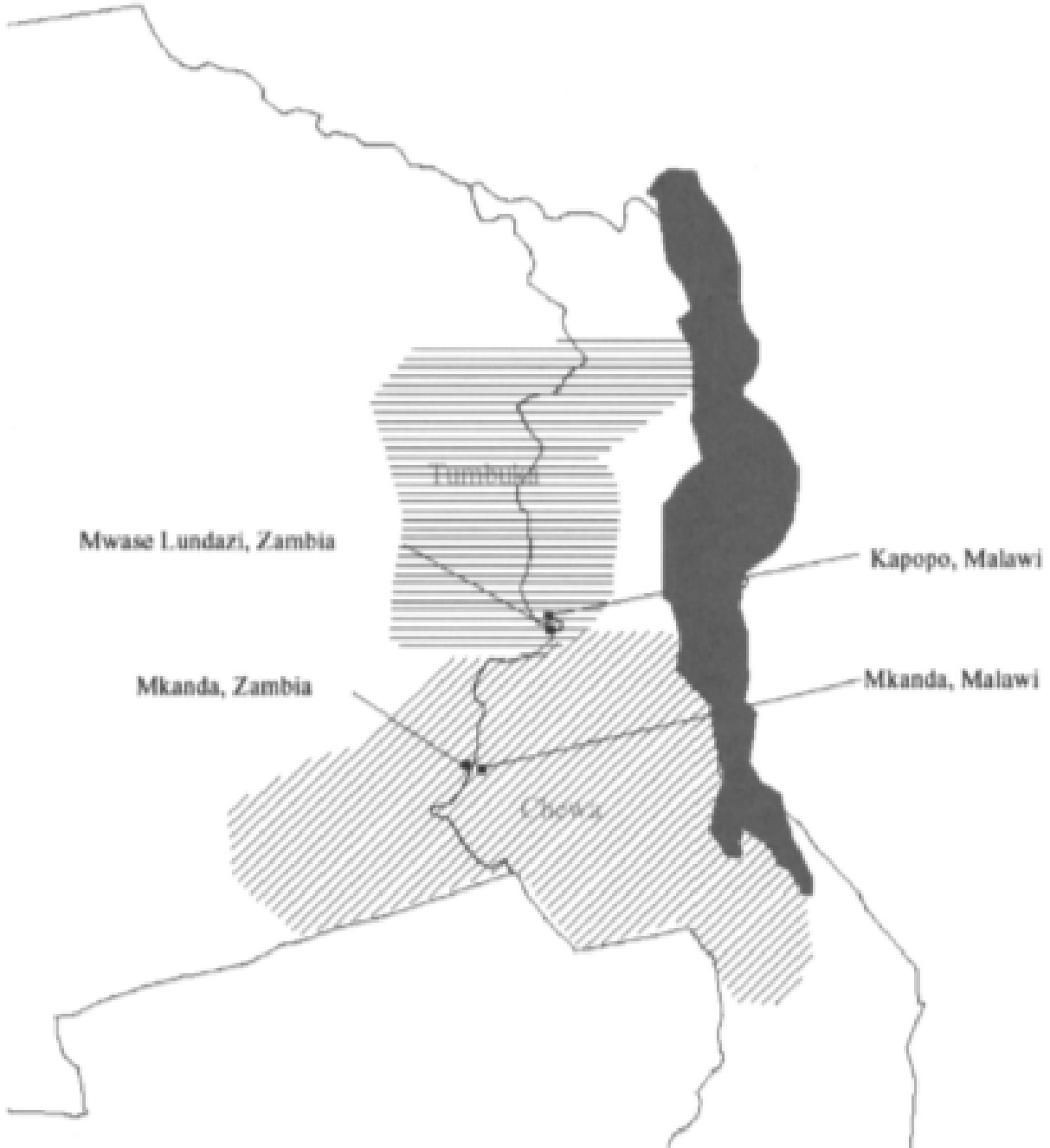
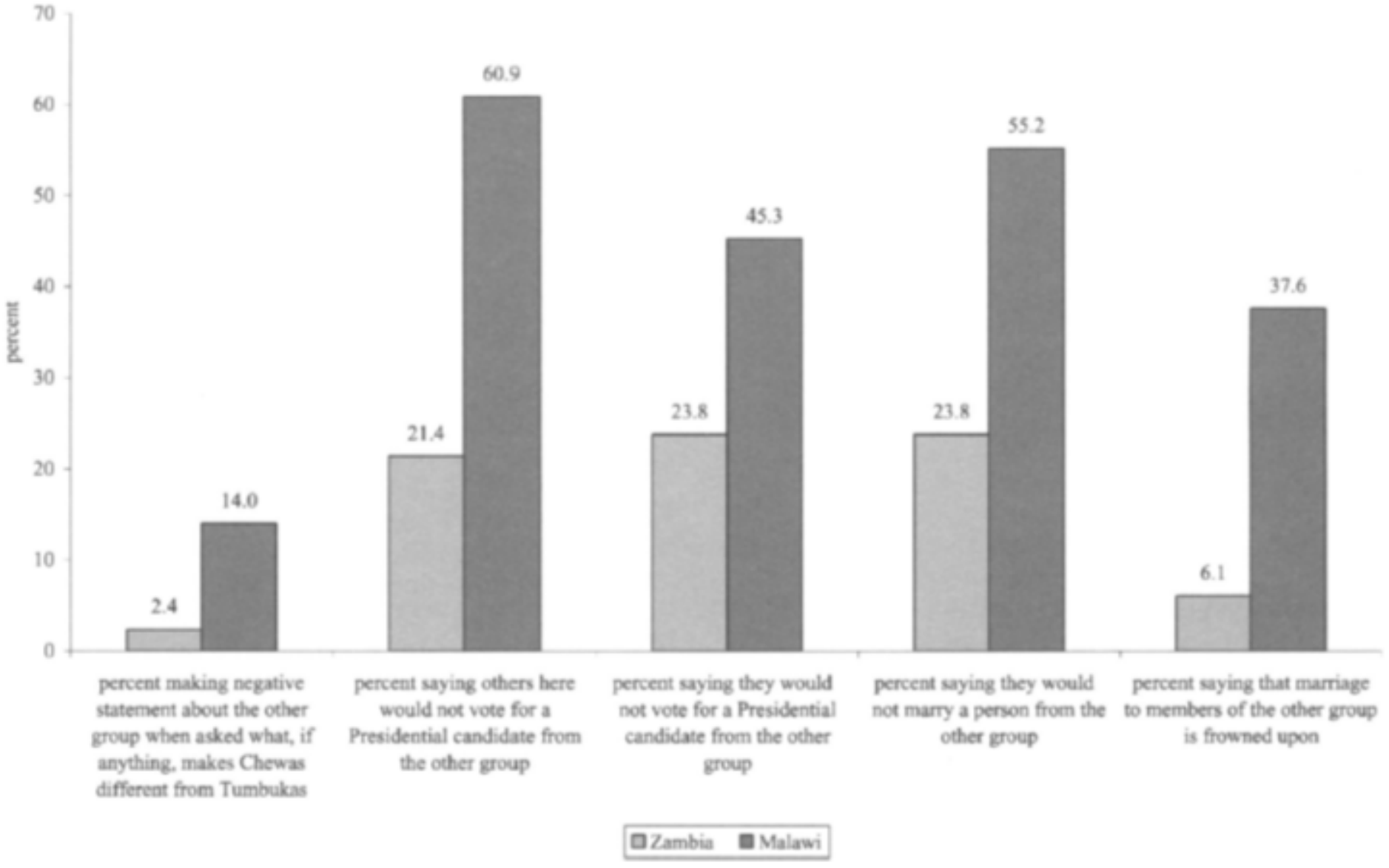


FIGURE 2. Chewa—Tumbuka Relations in Zambia and Malawi Compared



Posner (2004)

- What is treatment here?

Posner (2004)

- What is treatment here? Being in Zambia/Malawi

Posner (2004)

- What is treatment here? Being in Zambia/Malawi
- What is Posner interested in?

Posner (2004)

- What is treatment here? Being in Zambia/Malawi
- What is Posner interested in? Large ethnic groups relative to country size

Posner (2004)

- What is treatment here? Being in Zambia/Malawi
- What is Posner interested in? Large ethnic groups relative to country size
- But lots of things are different about Zambia!

Posner (2004)

- What is treatment here? Being in Zambia/Malawi
- What is Posner interested in? Large ethnic groups relative to country size
- But lots of things are different about Zambia!
- Eg. Zambia is *much* richer than Malawi due to copper revenues - maybe politics doesn't need to be as conflictual