

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 Ex- periments	✓	✓
Natural Experi- ments	✓	
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

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
 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

The Problem of not picking your own treatment

- ▶ “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 picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:

The Problem of not picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:
 1. These two groups actually are comparable (that they have potential outcomes that are independent of treatment assignment)
 - ▶ In general, we can only compare those units *that were part of the original randomization*

The Problem of not picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:
 1. These two groups actually are comparable (that they have potential outcomes that are independent of treatment assignment)
 - ▶ In general, we can only compare those units *that were part of the original randomization*
 2. That the treatment is the factor we actually care about

The Problem of not picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:
 1. These two groups actually are comparable (that they have potential outcomes that are independent of treatment assignment)
 - ▶ In general, we can only compare those units *that were part of the original randomization*
 2. That the treatment is the factor we actually care about
 - ▶ We have to 'interpret' the treatment

The Problem of not picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:
 1. These two groups actually are comparable (that they have potential outcomes that are independent of treatment assignment)
 - ▶ In general, we can only compare those units *that were part of the original randomization*
 2. That the treatment is the factor we actually care about
 - ▶ We have to 'interpret' the treatment
 - ▶ Sometimes treatments are 'bundles'

The Problem of not picking your own treatment

- ▶ “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 constructed *after* randomization - it's our responsibility to make sure:
 1. These two groups actually are comparable (that they have potential outcomes that are independent of treatment assignment)
 - ▶ In general, we can only compare those units *that were part of the original randomization*
 2. That the treatment is the factor we actually care about
 - ▶ We have to 'interpret' the treatment
 - ▶ Sometimes treatments are 'bundles'
 - ▶ Sometimes treatments are 'repeated', creating interactions or changing expectations

FIGURE 1. Illustration of Redistricting Research Designs

- District A
- District B

Voters to be
redistricted
from A to B
at election t

B's Original Voters

A's Original Voters

One-time Redistricting Design
Old Neighbors as counterfactual
New Voters vs. Old Neighbors

Old Neighbors

The Problem of not picking your own treatment

- ▶ Treatments may be bundled

The Problem of not picking your own treatment

- ▶ Treatments may be bundled
- ▶ What theoretical hypothesis do we actually want to test?

Section 2

Randomized

Ferraz and Finan (2008)

- Does accountability also work for negative politician performance like corruption?

Ferraz and Finan (2008)

- ▶ Does accountability also work for negative politician performance like corruption?
- ▶ But corruption is hard to manipulate ethically

Ferraz and Finan (2008)

- ▶ Does accountability also work for negative politician performance like corruption?
- ▶ But corruption is hard to manipulate ethically
- ▶ What is the inferential problem of using observational data on corruption?

Ferraz and Finan (2008)

- ▶ Does accountability also work for negative politician performance like corruption?
- ▶ But corruption is hard to manipulate ethically
- ▶ What is the inferential problem of using observational data on corruption?
- ▶ We can also look at voters' *information* about corruption

Ferraz and Finan (2008)

- ▶ Does accountability also work for negative politician performance like corruption?
- ▶ But corruption is hard to manipulate ethically
- ▶ What is the inferential problem of using observational data on corruption?
- ▶ We can also look at voters' *information* about corruption
- ▶ What is the inferential problem of using information on 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

- $IncumbVoteShare_{ms} = \alpha + \beta AuditedEarly_{ms} + X_{ms} + FE_s + \epsilon_{ms}$

Ferraz and Finan (2008)

- ▶ Methodology

- ▶ $IncumbVoteShare_{ms} = \alpha + \beta AuditedEarly_{ms} + X_{ms} + FE_s + \epsilon_{ms}$
 - ▶ NO EFFECT

Ferraz and Finan (2008)

- ▶ The importance of a theoretical model:
 1. The content of the information released varies
 2. People's expectations/priors vary
 3. For reports to have an effect, voters must receive it through the media
- ▶ 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)

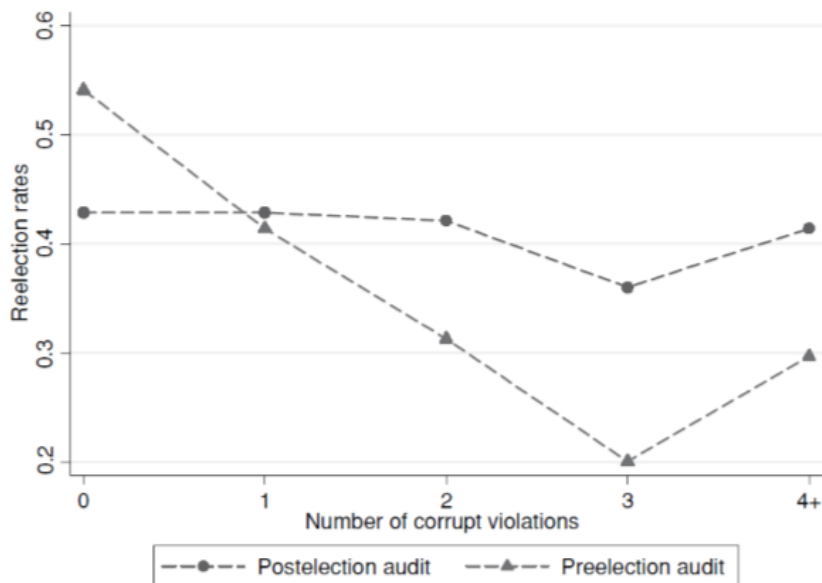
► Methodology

- So expected results are *conditional on content of the audit report*
- $IncumbVoteShare_{ms} =$
 $\alpha + \beta_1 AuditedEarly_{ms} + \beta_2 Corruption_{ms} +$
 $\beta_3 AuditedEarly_{ms} * Corruption_{ms} + X_{ms} + FE_s + \epsilon_{ms}$

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



Ferraz and Finan (2008)

- ▶ Did randomization work?
- ▶ Excludability: Is treatment the same in pre/post-election audits?
- ▶ Are corruption measures exogenous?

Section 3

Non-Randomized

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"
- ▶ There are good reasons to be skeptical

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"
- ▶ There are good reasons to be skeptical
- ▶ Humans are *strategic* - they make decisions based on how they affect the world, i.e. based on 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"
- ▶ There are good reasons to be skeptical
- ▶ Humans are *strategic* - they make decisions based on how they affect the world, i.e. based on potential outcomes
- ▶ But sometimes they are trying to alter outcomes *different to the potential outcomes we care about*

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"
- ▶ There are good reasons to be skeptical
- ▶ Humans are *strategic* - they make decisions based on how they affect the world, i.e. based on 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

- ▶ 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"
- ▶ There are good reasons to be skeptical
- ▶ Humans are *strategic* - they make decisions based on how they affect the world, i.e. based on 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

- ▶ 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"
- ▶ There are good reasons to be skeptical
- ▶ Humans are *strategic* - they make decisions based on how they affect the world, i.e. based on 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 theory and qualitative evidence of the treatment assignment mechanism

Posner (2004)

- How can we achieve causal inference without randomization?

Posner (2004)

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

Posner (2004)

- ▶ 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?

Posner (2004)

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

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

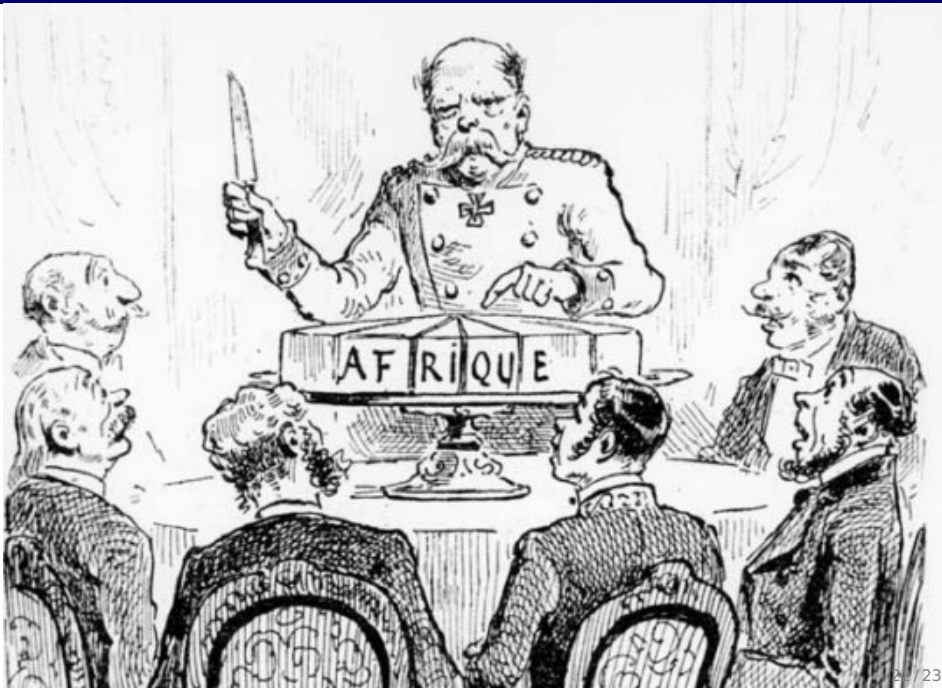


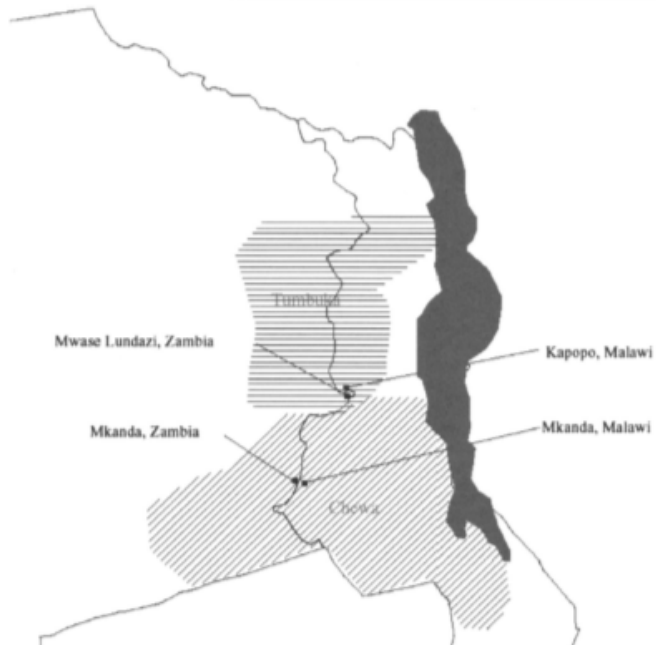
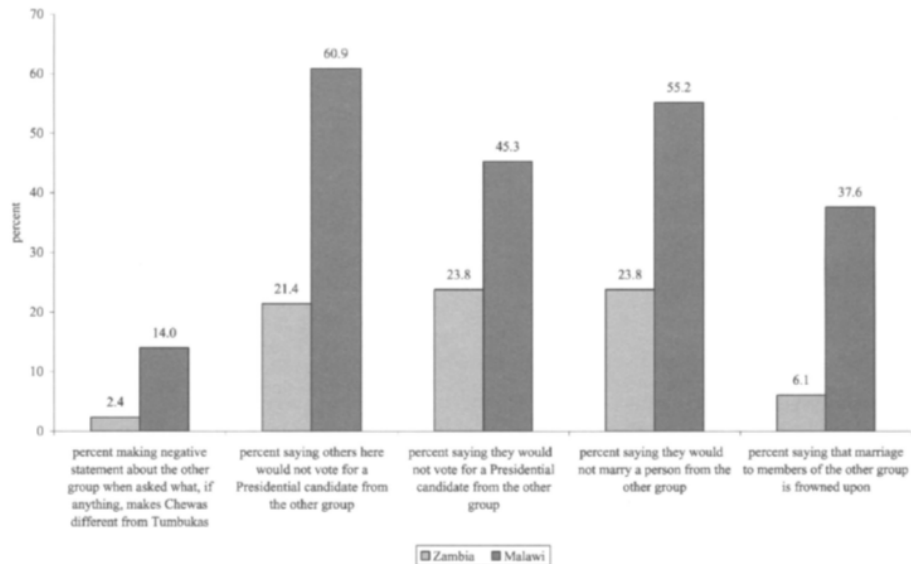
FIGURE 1. Research Sites

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

Posner (2004)

- What is treatment here?

Posner (2004)

- What is treatment here? What is Posner interested in?

Posner (2004)

- ▶ What is treatment here? What is Posner interested in?
- ▶ How do we know that's the treatment? Posner decided that being 'assigned' to Zambia means being in relatively small ethnic groups

Posner (2004)

- ▶ What is treatment here? What is Posner interested in?
- ▶ How do we know that's the treatment? Posner decided that being 'assigned' to Zambia means being in relatively small ethnic groups
- ▶ But lots of things are different about Zambia!