FLS 6415 - Causal Inference for the Political Economy of Development

Week 2 - The Fundamentals

Jonathan Phillips

August 2017

The Economic Roots

- The proximate causes of growth and poverty reduction
  - Growth in the capital stock
  - Investment in human capital
  - Adopting/discovering new techniques/technologies
- But when do societies achieve these?

The Economic Roots

- Coase Theoreom: Markets allow efficient exchange, provided property rights defined

The Economic Roots

- Coase Theoreom: Markets allow efficient exchange, provided property rights defined
  - Including investment and technology transfer
  - Bargaining power doesn’t affect outcomes

The Economic Roots

- Coase Theoreom: Markets allow efficient exchange, provided property rights defined
  - Including investment and technology transfer
  - Bargaining power doesn't affect outcomes
- BUT transaction costs prevent complex contracting: measurement, enforcement, information
  - So externalities and other market failures persist

Political Economy of Development

- What are the historical approaches to development?

Political Economy of Development

- What are the historical approaches to development?
  1. ’Big Push’ capital investment (Marshall Plan)
  2. Remove government and market failures (Washington Consensus, ’Good Government’)
  3. Power and Politics determine policy

'Big Push' capital investment

- Market failures prevent investment and coordination
- So Government steps in to subsidise industry

’Big Push’ capital investment

- Market failures prevent investment and coordination
- So Government steps in to subsidise industry
- BUT **Government failures** prevent productivity gains
    - No enforcement of infant industries
    - Rent-seeking and corruption
    - Lack of information

Washington Consensus & Good Governance

- Government is the problem - So minimize government

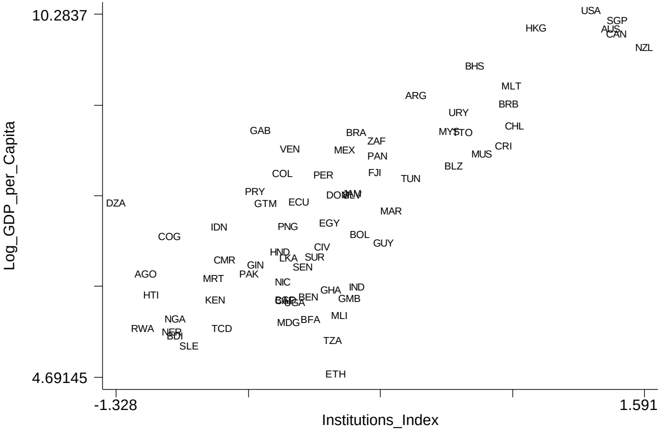
Washington Consensus & Good Governance

- Government is the problem - So minimize government
- Still a role to correct market failures

Washington Consensus & Good Governance

- Government is the problem - So minimize government
- Still a role to correct market failures
- Entails a minimum set of institutions that set the ‘rules’ and minimize transaction costs (Washington Consensus)

Figure 3: Logarithm of GDP per Capita in 1995 vs. Institutions Index



Washington Consensus & Good Governance

- Good Governance Institutions did not perform as expected:

Washington Consensus & Good Governance

- Good Governance Institutions did not perform as expected:
  - Government enforcement of property rights was not neutral
  - Reducing market failures was not enough to generate growth
  - The same institutions have different effects in different countries
  - Neither market systems nor democracies necessary for growth - East Asia

Power and Politics

- Need to understand actors' own incentives to make institutions work

Power and Politics

- Need to understand actors' own incentives to make institutions work
    - Implementation depends on a local political coalition
    - Effects depend on decentralized compliance and enforcement
    - Incentives depend on the **distribution of power**
    - Changing transaction costs produces new rents
    - Imposing institutions changes their effects

Power and Politics

- Need to understand actors' own incentives to make institutions work
  - Implementation depends on a local political coalition
  - Effects depend on decentralized compliance and enforcement
  - Incentives depend on the **distribution of power**
  - Changing transaction costs produces new rents
  - Imposing institutions changes their effects
- Promoting development means moving 'Beyond Good Governance', i.e. beyond institutions

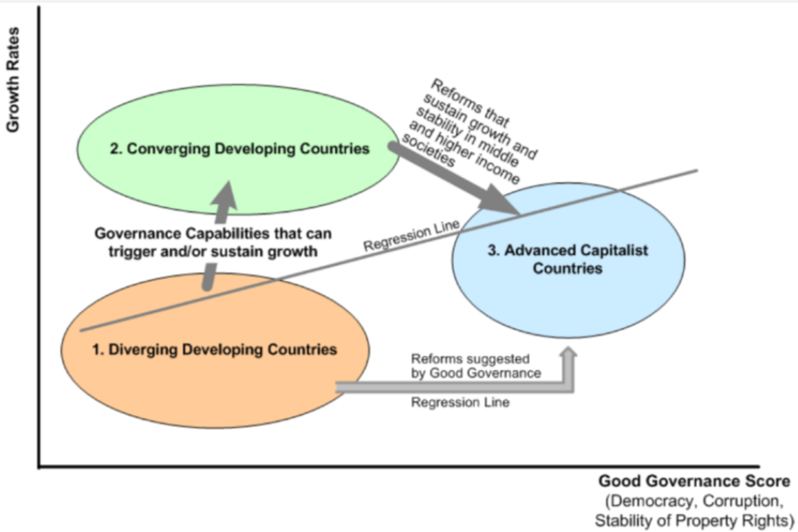
Causation and Institutions

- How good is the **causal evidence** on institutions' effects on growth?

Causation and Institutions

- How good is the **causal evidence** on institutions' effects on growth?
- Weak:
  - Reverse causation: Growth provides the surplus and incentives to finance institutions
  - Omitted variable: Distribution of power drives *both* growth and institutions
  - No causal strategy to analyze cross-country data
  - No compelling example case of an institutions-first approach to development

Causation and Institutions



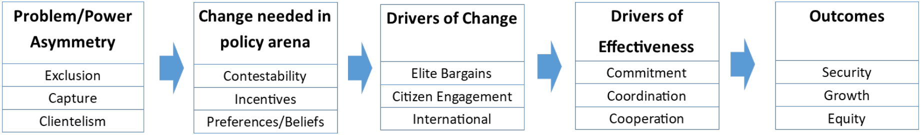
The World Bank's Approach

- Which approach does the World Bank WDR 2017 take?

The World Bank’s Approach

- Which approach does the World Bank WDR 2017 take?
  - Power matters - the opposite of the Coase Theorem
  - Institutions cannot be transplanted - the opposite of the Washington Consensus
  - Reform hard if we’re already in an equilibrium
  - Non-linear development process, eg. Brazil’s protests, Russia’s reforms
  - Promoting development means understanding and influencing the domestic policy arena

The World Bank’s Approach



The World Bank's Approach

- Where are institutions here?

The World Bank's Approach

- Where are institutions here?
  - Institutional rules still matter
  - But we need to understand them as part of a causal *mechanism*, not a generic treatment
  - Who has an incentive to promote an institution?
  - Who will comply with it?
  - Who will enforce it?
  - How does it interact with other instiutions?
- The focus is on institutions' actual causal effects in the local context, not just on the wording of the rules

The World Bank's Approach

- Does the Bank diagnose the development process correctly?

The World Bank's Approach

- Does the Bank diagnose the development process correctly?
- What are the 'policy implications' of this approach?

The World Bank’s Approach

- Does the Bank diagnose the development process correctly?
- What are the ’policy implications’ of this approach?
- What is the causal evidence supporting the approach?

Causal Inference

- Specify treatment:

$$D_i = \begin{cases} 1, \text{ if treated} \\ 0, \text{ if not treated} \end{cases}$$

- Clearly define the **contrast**
- Beware of **compound treatments** - new policy may provide both training and funding
- Beware of **scale effects** - giving one person a ticket to jump the queue is different to giving everyone a ticket to jump the queue
- If this is a policy question, can you **replicate the treatment** in reality?
  - Is an NGO delivering aid in a randomized trial the same as a government delivering aid on a daily basis?

Potential Outcomes

- The causal effect of treatment is how the **same unit**'s outcome differs when it is treated and not treated

$$Y_{Di} = \begin{cases} Y_{1i} \text{ Potential Outcome if unit i treated} \\ Y_{0i} \text{ Potential Outcome if unit i not treated} \end{cases}$$

- Treatment Effect $= Y_{1i} - Y_{0i}$

Causal Inference

- What is **The Fundamental Problem of Causal Inference?**

Causal Inference

- What is **The Fundamental Problem of Causal Inference?**
  - No units can receive **both** treatment and control
  - So we can never observe both Y_1 and Y_0 for the same unit

Causal Inference

- What is **The Fundamental Problem of Causal Inference?**
  - No units can receive **both** treatment and control
  - So we can never observe both Y_1 and Y_0 for the same unit
  - The very best we can do is estimate the effect by comparing **across** units
  - That is why we are doing causal **inference**, not causal proof

Potential Outcomes

- To compare across units we need counterfactuals: **control** units that do not receive treatment
- Causal Inference is all about identifying a **plausible counterfactual**"
  - The potential outcomes of the control unit are the same as those of the treated unit

Potential Outcomes

- Which unit is a plausible counterfactual for unit A?

Plausible Counterfactuals

	Y_1	Y_0
A	5	2
B	5	2
C	5	4
D	7	2

Potential Outcomes

- But we can NEVER confirm if a unit is a plausible counterfactual
- We can only gather data on **observed outcomes**, Y_i

$$Y_i = \begin{cases} Y_{1i} \text{ if } D_i = 1 \\ Y_{0i} \text{ if } D_i = 0 \end{cases}$$

$$Y_i = D_i \cdot Y_{1i} + (1 - D_i) \cdot Y_{0i} \tag{1}$$

Estimating Causal Effects

- With complete information on potential outcomes, calculating treatment effects is trivial

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	4	7	3
B	0	9	5	5	4
C	0	4	4	4	0
D	1	4	3	4	1

Estimating Causal Effects

- With complete information on potential outcomes, calculating treatment effects is trivial

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	4	7	3
B	0	9	5	5	4
C	0	4	4	4	0
D	1	4	3	4	1
$E(Y_1) =$		6			
$E(Y_0) =$			4		

- $ATE = E(Y_1 - Y_0) = 8/4 = 2$
- $ATE = E(Y_1) - E(Y_0) = 6 - 4 = 2$

Estimating Causal Effects

- From observed outcomes can we calculate an Average Treatment Effect?

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	?	7	?
B	0	?	5	5	?
C	0	?	4	4	?
D	1	4	?	4	?

Estimating Causal Effects

- From observed outcomes can we calculate an Average Treatment Effect?

Calculating Treatment Effects

	D	Y_1	Y_0	Y_i	Real Effect, $Y_1 - Y_0$
A	1	7	?	7	?
B	0	?	5	5	?
C	0	?	4	4	?
D	1	4	?	4	?
$E(Y_1) =$		5.5			
$E(Y_0) =$			4.5		

Estimating Causal Effects

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$\begin{aligned} ATE = & \quad E(Y_1) - E(Y_0) && (2) \\ = & \quad 5.5 - 4.5 && (3) \\ = & \quad 1 && (4) \end{aligned}$$

- Half the true treatment effect

Estimating Causal Effects

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE = & \quad E(Y_1) - E(Y_0) \quad (2) \\
= & \quad 5.5 - 4.5 \quad (3) \\
= & \quad 1 \quad (4)
\end{aligned}
$$

- Half the true treatment effect
- Why?

Estimating Causal Effects

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$\begin{aligned} ATE = & E(Y_1) - E(Y_0) & (2) \\ = & 5.5 - 4.5 & (3) \\ = & 1 & (4) \end{aligned}$$

- Half the true treatment effect
- Why?
  - The units that got treated had lower Y_1
  - The units that were controls had higher Y_0

Estimating Causal Effects

- If we use the control units as counterfactuals...
- Average Treatment Effect:

$$
\begin{aligned}
ATE &= E(Y_1) - E(Y_0) && (2)\\
&= 5.5 - 4.5 && (3)\\
&= 1 && (4)
\end{aligned}
$$

- Half the true treatment effect
- Why?
  - The units that got treated had lower Y_1
  - The units that were controls had higher Y_0
  - The ’stand-in’ counterfactuals were wrong

Estimating Causal Effects

- So how can we ensure we have **plausible** counterfactuals?
  - (A control unit with the same potential outcomes)
- The bias in units' potential outcomes depends on which units get treated and which ones don't
- We need to understand the **treatment assignment mechanism**

Estimating Causal Effects

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1

Estimating Causal Effects

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1
$E(Y_1) =$			4			
$E(Y_0) =$				4.5		

Estimating Causal Effects

- Imagine a treatment assignment mechanism where all women get treated

Treatment Assignment by Covariate

	X	D	Y_1	Y_0	Y_i	Real Effect
A	Man	0	7	4	4	3
B	Man	0	9	5	5	4
C	Woman	1	4	4	4	0
D	Woman	1	4	3	4	1
$E(Y_1) =$			4			
$E(Y_0) =$				4.5		

- ATE = 4 - 4.5 = -0.5
- This is **confounding** or an **omitted variable** - another variable affects both treatment and potential outcomes

Estimating Causal Effects

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1

Estimating Causal Effects

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1
$E(Y_1) =$		8			
$E(Y_0) =$			3.5		

Estimating Causal Effects

- Imagine a treatment assignment mechanism where people get to *choose* their treatment

Treatment Assignment by Self-Selection

	D	Y_1	Y_0	Y_i	Real Effect
A	1	7	4	7	3
B	1	9	5	9	4
C	0	4	4	4	0
D	0	4	3	3	1
$E(Y_1) =$		8			
$E(Y_0) =$			3.5		

- ATE = 8 - 3.5 = 4.5
- This is **self-selection bias** - treatment is affected by potential outcomes

Estimating Causal Effects

- Depending on the treatment assignment mechanism we get a range of Average Treatment Effects:

Comparing Average Treatment Effects

Treated Units	**ATE**
Real Effect for all units	2
A & D	1
Women	-0.5
Self-selection	4.5

Estimating Causal Effects

- We can identify the source of these biases in potential outcomes:

Estimating Causal Effects

- We can identify the source of these biases in potential outcomes:

$$\underbrace{E(Y_i|D=1) - E(Y_i|D=0)}_{\text{Observed Effect}} \quad (5)$$

Estimating Causal Effects

- We can identify the source of these biases in potential outcomes:

$$\underbrace{E(Y_i|D=1)-E(Y_i|D=0)}_{\text{Observed Effect}}=\underbrace{E(Y_{1i}-Y_{0i})}_{\text{Real ATE}}$$
$$\underbrace{+\frac{1}{2}\Big[E(Y_{1i}|D=1)-E(Y_{1i}|D=0)\Big]}_{\text{Imbalance on }Y_1}\underbrace{+\frac{1}{2}\Big[E(Y_{0i}|D=1)-E(Y_{0i}|D=0)\Big]}_{\text{Imbalance on }Y_0} \tag{6}$$

NB: For equal-sized treatment and control groups

Estimating Causal Effects

- Disaggregating the Self-Selection Bias:

$$\frac{(7+9-4-3)}{2} = \frac{(7+9+4+4-4-5-4-3)}{4} + \frac{1}{2}\Big[\frac{(7+9)}{2} - \frac{(4+4)}{2}\Big] + \frac{1}{2}\Big[\frac{(4+5)}{2} - \frac{(4+3)}{2}\Big]$$
$$4.5 = {\color{red}2} + 2 + \frac{1}{2} \quad (7)$$

Treatment Assignment Mechanisms

- The rest of the course is mostly about the types of treatment assignment mechanisms that **avoid these biases** and provide plausible counterfactuals

Treatment Assignment Mechanisms

1. **Controlled Experiments** where we **control** the treatment assignment
    - Field Experiments
    - Survey Experiments
    - Lab Experiments

Treatment Assignment Mechanisms

2. **Natural Experiments** where the assignment mechanism creates balanced potential outcomes
   - Randomized Natural Experiments
   - Regression Discontinuities
   - Instrumental Variables

Treatment Assignment Mechanisms

3. **Observable Studies:** Where 'helpful' treatment assignments might not be available
    - No historical examples of natural experiments
    - Not feasible or ethical to run a field experiment

- One alternative way of making potential outcomes comparable is to **selectively use Observable Data**
    - Difference-in-differences
    - Controlling for confouding variables
    - Matching

Treatment Assignment Mechanisms

Analysis Types and Assumptions

Week		Researcher Controls Treatment Assignment?	Treatment Assignment Independent of Potential Outcomes	SUTVA	Additional Assumptions
	Controlled Experiments				
1	Field Experiments	✓	✓	✓	
2	Survey and Lab Experiments	✓	✓	✓	Controlled Environment for treatment exposure
	Natural Experiments				
3	Randomized Natural Experiments	X	✓	✓	Compliance with Randomization
4	Instrumental Variables	X	✓	✓	First stage and Exclusion Restriction (Instrument explains treatment but not outcome)
5	Regression Discontinuity	X	✓	✓	Continuity of covariates; No manipulation; No compounding discontinuities
	Observational Studies				
6	Difference-in-Differences	X	X	✓	No Time-varying confounders; Parallel Trends
7	Controlling for Confounding	X	X	✓	Blocking all Back-door paths
8	Matching	X	X	✓	Overlap in sample characteristics

Treatment Assignment Mechanisms

4. **Small-N studies:** With few units available we can at least avoid some key biases:

- Comparative Case Studies
- Process Tracing

Measuring Causal Effects

- We can measure causal effects for different groups: men, women...
- And by treatment status:
  - Crucial where the treated population has different potential outcomes to the control population

Measuring Causal Effects

- We can measure causal effects for different groups: men, women...
- And by treatment status:
  - Crucial where the treated population has different potential outcomes to the control population
- Average Treatment Effect = $E(Y_1 - Y_0)$
- Average Treatment Effect on the Treated = $E(Y_1 - Y_0 | D = 1)$
- Average Treatment Effect on the Untreated = $E(Y_1 - Y_0 | D = 0)$

Measuring Causal Effects

- We can measure causal effects for different groups: men, women...
- And by treatment status:
  - Crucial where the treated population has different potential outcomes to the control population
- Average Treatment Effect = $E(Y_1 - Y_0)$ What would happen if we treated everyone
- Average Treatment Effect on the Treated = $E(Y_1 - Y_0|D = 1)$ What happened to those who were actually treated
- Average Treatment Effect on the Untreated = $E(Y_1 - Y_0|D = 0)$ What would happen if we *extended* treatment to others

Measuring Causal Effects

- We can also (depending on the data) calculate quantile effects: what is the effect of going to university for those who are at the bottom 10th percentile of effects?
- We can NEVER identify individual causal effects
- Even measuring the same person before and after treatment
  - The time will be different
  - You may have learned something
  - Measurement may itself have an effect
- Remember that average causal effects are *net* effects
  - Some people can be hurt while others benefit

Assumptions for all Analyses

- Because we have to compare across units, how those units interact is crucial. We always assume:
  - 'Units do not interfere with each other' = **SUTVA** = Stable Unit Treatment Value Assumption
    - My potential outcomes do not depend on your treatment status: $Y_{1i}, Y_{0i} \perp D_j$
    - But: merit awards spillovers, immunization...
- Always justify SUTVA with our knowledge of how the data was generated

Statistical Inference

- Inference is about how we learn from a **sample** about a **population**
  - Our sample must be representative of that population if we are to make inference
  - REMEMBER: A random *sample* is different from random *treatment*.
    - A random sample allows us to make inference from the sample to the population
    - Random treatment (next week) allows us to make inference about counterfactuals
- Since all our results are based on comparison, they will change as we make more comparisons
- So we want to understand not just our ‘best guess’ of the causal effect, but our **confidence**
  - How do we measure uncertainty?

Statistical Inference

- Statistical significance depends on:
  - Sampling uncertainty - how well does our sample approximate the population?
  - Fundamental uncertainty - potential outcomes are not fixed, but are themselves distributions
  - Measurement uncertainty - did we precisely measure Y_i?
- We could use a simple t-test for difference in means
- Or standard regression tools

Statistical Inference

- **How much** can we learn from a causal analysis?
- **Internal Validity:** Have we succeeded in causal inference about our sample?
- **External Validity:** How much does our data tell us about the real world?
  - Would the same thing happen in another country? Next year?
  - Look out for variation in treatment, context, spillovers, learning etc.
  - *How* a treatment is introduced might also have an effect
- Any generalization requires assumptions

Causal Mechanisms

- We will try to identify abstract, portable processes
  - **Causal Mechanisms**
- **Portable:** If the weather affects election turnout ONLY in Acre, is that a useful causal mechanism?
- **Abstract:** If unions are good at mobilizing support, but so are churches, the mechanism is collective action, not union organization
- We still need to define the **scope conditions** in which we think this causal mechanism will operate as expected

Causal Mechanisms

- Examples of Causal Mechanisms:
  - Citizens
    - Electoral Accountability
    - Client Power
    - Collective Action
    - Social Trust/Sanctioning
    - Wealth Effects
  - Elites
    - Violence/Coercion
    - Brokerage/Patronage
    - Persuasion/Framing
    - Incumbency Power
  - Institutions
    - Power Devolution/Median Voter
    - Network Effects
    - Evolutionary Selection
    - Conversion/Layering/Drift/Replacement

Causal Mechanisms

- Examples of Causal Mechanisms:
  - Citizens
    - Electoral Accountability - Class 5
    - Client Power - Class 6
    - Collective Action - Class 11
    - Social Trust/Sanctioning - Class 4
    - Wealth Effects
  - Elites
    - Violence/Coercion - Class 8
    - Brokerage/Patronage - Class 9
    - Persuasion/Framing
    - Incumbency Power - Class 7
  - Institutions
    - Power Devolution/Median Voter - Class 3
    - Network Effects
    - Evolutionary Selection
    - Conversion/Layering/Drift/Replacement - Class 12

Reproducible Research

- The big problem: Give 5 researchers the same data and the same method and you'll get 5 different answers
- Replicating someone else's results is a minimum requirement, but it's hard
    - Manual data processing
    - No documentation of data processing
    - Errors unseen
    - Updates not consistent
    - Copy-paste errors
- Our research must be **reproducible**
    - Always generate the same results
    - Easily diagnose errors
    - Easily collaborate

Reproducible Research

- Principles of Reproducible Research
  1. Never touch the raw data
  2. Write code in a script
  3. Directly produce output documentation
  4. Every result comes from your code
  5. Comment and explain your code
  6. Manipulate data using clear rules, not individual items
  7. No cut-and-paste (more than twice)