FLS 6415 - Causal Inference for the Political Economy of Development Week 4 - The Effects of Social Organization & Lab Experiments

Jonathan Phillips

August 2017

How do social constructs like identity affect development?

- How do social constructs like identity affect development?
- How does social organization alter the impact of institutions?

- How do social constructs like identity affect development?
- How does social organization alter the impact of institutions?
- ► Big questions:
 - Does Weber's 'Protestant Ethic' explain western growth?
 - How do elites maintain their power over generations?
 - Why are some people more willing to pay taxes than others?

- 'Norms' and 'culture' structure interactions within and between social groups
- What is a social norm?

- 'Norms' and 'culture' structure interactions within and between social groups
- What is a social norm?
 - Social norms as 'informal institutions'
 - Social norms as mutual expectations/coordination devices
 - Social norms as punishments in game theory

What question is Whitt (2014) trying to answer?

- What question is Whitt (2014) trying to answer?
 - "This study examines the impact of ethnic violence on a basic norm of fairness."

- What question is Whitt (2014) trying to answer?
 - "This study examines the impact of ethnic violence on a basic norm of fairness."
 - Can social norms of fairness across ethnic divisions persist after ethnic violence?

- What question is Whitt (2014) trying to answer?
 - "This study examines the impact of ethnic violence on a basic norm of fairness."
 - Can social norms of fairness across ethnic divisions persist after ethnic violence?
 - Does ethnicity affect how Bosnians treat each other?



Why use a lab experiment?

- Why use a lab experiment?
 - We can't manipulate ethnicity directly
 - Lots of confounders affect real-world ethnic interactions: wealth, location, context etc.
 - Hard to measure outcomes of varying real-world interactions

What is the population?

- What is the population?
- What is the sample? / sampling procedure?

- What is the population?
- What is the sample? / sampling procedure?
- Does the sample (over-sampling) matter for inference?

What is the treatment?

- What is the treatment?
- What is the control?

- What is the treatment?
- ► What is the **control**?
- What is the treatment assignment mechanism?

- What is the treatment?
- What is the control?
- What is the treatment assignment mechanism?
- What is the lab/game context?

- What is the treatment?
- What is the control?
- What is the treatment assignment mechanism?
- What is the lab/game context?
 - Choice set
 - Information/instructions
 - History/sequence
 - Physical and social setting

- What is the treatment?
- What is the control?
- What is the treatment assignment mechanism?
- What is the lab/game context?
 - Choice set
 - Information/instructions
 - History/sequence
 - Physical and social setting
- What does the researcher control and what do they not control?

Subject ethnicity	Recipient R I	Mean to recipient R I	Recipient R2	Mean to recipient R2	SD	% Bias against R2	N	<i>t</i> -test for difference in means (RI = R2)
Bosniak	Bosniak	5.55	Croat	4.45	1.33	22.5	80	5.23****
	Bosniak	5.86	Serb	4.14	1.70	37.5	88	6.71***
	Croat	5.37	Serb	4.63	1.72	22.9	83	3.86***
Croat	Croat	6.05	Bosniak	3.95	1.92	31.3	64	6.19 ^{%%}
	Croat	5.81	Serb	4.19	1.71	31.5	73	5.72***
	Serb	5.08	Bosniak	4.92	1.93	19.7	66	0.48
Serb	Serb	5.68	Bosniak	4.32	1.77	36.8	76	4.74 ****
	Serb	5.57	Croat	4.43	1.57	28.0	75	4.45 ^{****}
	Croat	5.07	Bosniak	4.93	1.57	13.4	67	0.52

 Table 3. Experimental Allocations by Subject and Recipient Ethnicity.

****Significant at $p \leq .010$.

What are the results?

- What are the results?
- 2/3rds allocate 5:5
 - Is that large or small?

- What are the results?
- 2/3rds allocate 5:5
 - Is that large or small?
- Higher allocations go to co-ethnics (in-group)
 - Croats are the most biased to their in-group
 - Serbs are the 'fairest'

- What are the results?
- 2/3rds allocate 5:5
 - Is that large or small?
- Higher allocations go to co-ethnics (in-group)
 - Croats are the most biased to their in-group
 - Serbs are the 'fairest'
- When forced to choose between two non-co-ethnics (out-group members)
 - Croats and Serbs show no preference
 - Bosniaks allocate more to Croats than Serbs
- Differences in allocations also correlated with attitudes
 - Caring about identity -> more bias
 - Fearing out-groups -> more bias
 - Bias -> Expectations of bias

How does Whitt interpret these findings?

How does Whitt interpret these findings?

- In-group solidarity
- Out-group aversion/punishment

How does Whitt interpret these findings?

- In-group solidarity
- Out-group aversion/punishment
- Could it represent anything else? What norm are we really measuring?
 - Deference/dependence?
 - Charity?
 - Fairness doesn't mean equality maybe compensating for average wealth/power

Are responses likely to be accurate?

- Are responses likely to be accurate?
 - NOT an incentive-compatible, behavioural measure
 - Anticipating what researchers wanted?
 - Did they really believe that people would receive the money?
 - Sequencing effects from previous experiments?
 - What if there was no clear focal point (5:5)?

What do the findings about recipient gender suggest?

- What do the findings about recipient gender suggest?
 - No difference in allocation
 - So do we really think there's no discrimination or unfairness to women?

Do we ever do this treatment in reality?

- Do we ever do this treatment in reality?
 - Where are institutions here? What institutional context does this correspond to?
 - Where are political actors/manipulations/discourse here?
- Do we ever do this treatment in reality?
 - Where are institutions here? What institutional context does this correspond to?
 - Where are political actors/manipulations/discourse here?
- How much can we generalize?

- Do we ever do this treatment in reality?
 - Where are institutions here? What institutional context does this correspond to?
 - Where are political actors/manipulations/discourse here?
- How much can we generalize?
- What are the policy implications?

Habyarimana et al (2007)

- ► Habyarimana et al (2007)
- Existing consensus: Ethnic diversity -> Less public goods provision

- ► Habyarimana et al (2007)
- Existing consensus: Ethnic diversity -> Less public goods provision
- But how?

- ► Habyarimana et al (2007)
- Existing consensus: Ethnic diversity -> Less public goods provision
- But how?
 - Preferences in-group fairness
 - Technology social networks permit identification and sanctioning
 - Strategy Selection choose to cooperate more often

- ► Lab-in-the-field
- Population: Ugandans
- Sample: 300 people in a diverse, low public-goods poor area
- **Treatment/Control**: Various Games
- Treatment assignment: Random assignment to co-ethnic/non-co-ethnic

- Preferences dictator game between self and two others
 - No bias towards co-ethnics

- > Preferences dictator game between self and two others
 - No bias towards co-ethnics
- Technology 1, productivity teamwork in a puzzle requiring communication
 - Co-ethnic teams don't perform any better

- ► Preferences dictator game between self and two others
 - No bias towards co-ethnics
- Technology 1, productivity teamwork in a puzzle requiring communication
 - Co-ethnic teams don't perform any better
- Technology 2, social networks Can you find a co-ethnic in the town faster than a non-co-ethnic?
 - Yes (43% vs 28% success)

- Preferences dictator game between self and two others
 - No bias towards co-ethnics
- Technology 1, productivity teamwork in a puzzle requiring communication
 - Co-ethnic teams don't perform any better
- Technology 2, social networks Can you find a co-ethnic in the town faster than a non-co-ethnic?
 - Yes (43% vs 28% success)
- Strategy Selection Does anonymity for the sender in the dictator game make a difference?
 - Yes offer more to co-ethnics when offerers believe they can be seen

- Preferences dictator game between self and two others
 - No bias towards co-ethnics
- Technology 1, productivity teamwork in a puzzle requiring communication
 - Co-ethnic teams don't perform any better
- Technology 2, social networks Can you find a co-ethnic in the town faster than a non-co-ethnic?
 - Yes (43% vs 28% success)
- Strategy Selection Does anonymity for the sender in the dictator game make a difference?
 - Yes offer more to co-ethnics when offerers believe they can be seen

 Conclusion: Norms and Networks allow co-ethnics to provide more public goods

- Conclusion: Norms and Networks allow co-ethnics to provide more public goods
 - ...But where are the public goods here?
 - Are public goods organized by voluntary contributions or coercive central authority?
 - Is this true of all parts of Kampala? Uganda? All ethnic groups?

- Past inequality -> Expectations -> Poor's Behaviour -> Current Inequality
- Effect of the Indian caste system on development through expectations

- Past inequality -> Expectations -> Poor's Behaviour -> Current Inequality
- Effect of the Indian caste system on development through expectations
- Why use a lab experiment?

- Past inequality -> Expectations -> Poor's Behaviour -> Current Inequality
- Effect of the Indian caste system on development through expectations
- Why use a lab experiment?
 - Can't manipulate caste directly
 - Want to isolate differences in social treatment, not differences in opportunities
 - Focus manipulation on social treatment of caste

• **Population**: Indians with a caste (all religions)

- Population: Indians with a caste (all religions)
- ► Sample: High School Boys in one village in Uttar Pradesh

- Population: Indians with a caste (all religions)
- ► Sample: High School Boys in one village in Uttar Pradesh
- Treatment: Others' awareness of caste

- Population: Indians with a caste (all religions)
- ► Sample: High School Boys in one village in Uttar Pradesh
- Treatment: Others' awareness of caste
- Control: Lack of awareness of caste

- Population: Indians with a caste (all religions)
- **Sample**: High School Boys in one village in Uttar Pradesh
- Treatment: Others' awareness of caste
- Control: Lack of awareness of caste
- Outcome measure: Performance in completing 15 mazes; Rewards allocated by a local person

- Results
 - Caste unknown: No difference in performance between castes

- Results
 - Caste unknown: No difference in performance between castes
 - Caste revealed: Low-caste perform much worse (42% less earnings while high-caste earn 6% more)

- Results
 - Caste unknown: No difference in performance between castes
 - Caste revealed: Low-caste perform much worse (42% less earnings while high-caste earn 6% more)
 - Castes segregated: Low-caste perform even worse

- Results
 - Caste unknown: No difference in performance between castes
 - Caste revealed: Low-caste perform much worse (42% less earnings while high-caste earn 6% more)
 - Castes segregated: Low-caste perform even worse
 - No discretion in rewards: No difference in performance

► Interpretation:

- ► Interpretation:
 - Low caste expect to be judged poorly
 - So the reward is less motivating
 - So they try less

- Issues:
 - Is the effect due to other boys knowing caste, or the evaluator?
 - Could this be a lab-specific effect?
 - Any bias in the type of low-caste boys who go to school?
 - What was the political context of caste in the village?
 - What about the very different south of the country?

Causal Inference

- Causal Inference
- Why lab experiments?

- Causal Inference
- Why lab experiments?
 - Treatments we cannot administer in reality
 - Outcome measurements that are hard to take in reality
 - Random treatment assignment not permitted in reality

• Treatment Assignment: Same as a Field Experiment

- **Treatment Assignment**: Same as a Field Experiment
- Treatment: Not a manipulation of real world political or economic processes, but establishing controlled 'lab' conditions

- **Treatment Assignment**: Same as a Field Experiment
- Treatment: Not a manipulation of real world political or economic processes, but establishing controlled 'lab' conditions
 - The advantage: Control over context helps isolate mechanisms
 - The disadvantage: Can we generalize to the real world from this artificial context?
- Problems generalizing from the lab:
 - Hawthorne effect: Lab context influences behaviour, social desirability bias
 - Context effects: The real-world always provides more information, more history
 - **Process effects**: People care *how* decisions are made
 - Selection effects: Actors in specific roles are rarely representative samples, 'WEIRD' or pro-social lab subjects

The lab differs from the field

- The lab differs from the field
 - The stakes
 - The norms (specific norms of being an experimental subject)
 - The degree of scrutiny
 - The sample of individuals
 - The degree of anonymity

 Levitt and List 2006 argue lab experiments are *inherently* flawed because the decisions we want to measure are likely to change depending on the degree of **scrutiny**

- Levitt and List 2006 argue lab experiments are *inherently* flawed because the decisions we want to measure are likely to change depending on the degree of **scrutiny**
- "You tip more when you're on a date"
- Social norms are activated, eg. treating one-shot games like repeated games
- Scrutiny alters who wants to make a decision as well as the decision they make
- Subjets use cues (heuristics) to draw on 'similar' situations from the real world

 Many studies find more cooperation in the lab than in the real world

- Many studies find more cooperation in the lab than in the real world
 - Scrutiny increases cooperation
 - Anonymity reduces cooperation
 - That's interesting in itself! We can manipulate the degree of scrutiny/anonymity etc.

 Lab experiments may be generalizable where norms/morality is less important

- Lab experiments may be generalizable where norms/morality is less important
 - ▶ ???

Survey Experiments

- ► Treatment occurs *within* the survey questionnaire
 - Different versions of the questionnaire randomly applied
 - Not a field experiment: Still an artificial context
 - Not a lab experiment: People not brought to a single location or interacting

- How do people make choices between many options?
- Treatments are often 'bundles', but which aspect matters most?

Hainmueller et al 2013 - How do attitudes to immigrants depend on immigrant characteristics?

- Hainmueller et al 2013 How do attitudes to immigrants depend on immigrant characteristics?
- Vary education, profession, language, gender, national origin, etc.
- Profiles
 - Attributes
 - Values
- Randomize attribute order to prevent bias

Jens Hainmueller et al.

Please read the descriptions of the potential immigrants carefully. Then, please indicate which of the two immigrants you would personally prefer to see admitted to the United States.

	Immigrant 1	Immigrant 2	
Prior Trips to the U.S.	Entered the U.S. once before on a tourist visa	Entered the U.S. once before on a tourist visa	
Reason for Application	Reunite with family members already in U.S.	Reunite with family members already in U.S.	
Country of Origin	Mexico	Iraq	
Language Skills	During admission interview, this applicant spoke fluent English	During admission interview, this applicant spoke fluent English	
Profession	Child care provider	Teacher	
Job Experience	One to two years of job training and experience	Three to five years of job training and experience	
Employment Plans	Does not have a contract with a U.S. employer but has done job interviews	Will look for work after arriving in the U.S.	
Education Level	Equivalent to completing two years of college in the U.S. College degree in the U.S.		
Gender	Female	Male	

	Immigrant 1	Immigrant 2	
If you had to choose between them, which of these two immigrants should be given priority to come to the United States to live?	0	0	

On a scale from 1 to 7, where 1 indicates that the United States should absolutely not admit the immigrant and 7 indicates that the United States should definitely admit the immigrant, how would you rate immigrant 1?

Absolutely Not Admit						Definitely Admit
1	2	3	4	5	0	/
0	0	0	0	0	0	0

Using the same scale, how would you rate immigrant 2?

Absolutely Not Admit						Definitely Admit
1	2	3	4	5	6	/
0	0	0	0	0	0	0



choice outcomes hereafter. Second, in "rating-based conjoint analysis," respondents give a numerical rating to each profile which represents their degree of preference for the profile. This format is preferred by some analysts who contend that such ratings provide more direct, finely grained information about respondents' preferences. We call this latter type of outcome a rating outcome.

6

Causal Inference in Conjoint Analysis

Gender: female male		-	•	
Education: no formal 4th grade 8th grade high school hivo-year college college degree graduate degree			 -	
Language: fluent English broken English tried English but unable used interpreter	_	~		
Origin: Germany France Mexico Philippines Poland India China Sudan Sudan Somalia Iraq		<u>۔</u> ==		
Profession: janitor waiter child care provider gardener financial analyst construction worker teacher comsture programmer nurse research scientist doctor		=		-
Job experience: none 1-2 years 3-5 years 5+ years			•	.
Job plans: contract with employer interviews with employer will look for work no plans to look for work	-	•	-	•
Application reason: reunite with family seek better job escape persecution				
Prior trips to U.S.: never once as tourist many times as tourist six months with family once w/o authorization		_ _	· =	
	2 Cha	ange in Pr(Immigrant Pre	0 ferred for Admis	sion to U.S.)

Fig. 3 Effects of immigrant attributes on preference for admission. This plot shows estimates of the effects of the randomly assigned immigrant attributes on the probability of being preferred for admission to the United States. Estimates are based on the regression estimators with clustered standard errors, bars represent 95% confidence intervals. The points without horizontal bars denote the attribute value that is the reference category for each attribute.

21

a

20, 2013

 Estimating results uses a simple regression of respondent choices on profile attribute-values

- Estimating results uses a simple regression of respondent choices on profile attribute-values
- But each specific profile arises too rarely to make comparisons of individual attribute-values
 - So this is **not** an Average Treatment Effect
 - Eg. the effect of gender when age, language etc. are held constant
 - It is an Average Marginal Component Effect
 - Eg. the effect of gender averaging across all possibilities of age, language, etc.

- How realistic are the responses?
 - Not a behavioural measure; nothing at stake
 - Social desirability bias
 - Not like real-world preference-formation process
- Hainmueller et al 2014 compare conjoint responses to a Swiss referendum
- Citizens voted on specific naturalization applicants (Really!)



Figure S11: Effects of Applicant Attributes on Opposition to Naturalization Request (Unweighted Survey Sample)

Figure shows point estimates (dots) and corresponding, cluster-robust 95 % confidence intervals (horizontal lines) from ordinary least squares regressions. The dots on the zero line without confidence intervals denote the reference category for each applicant attribute.

- But note the conjoint method still hugely under-estimated the overall rejection rate
- ▶ 21% versus 37% in reality

Survey experiments are also valuable for measurement

- Survey experiments are also valuable for measurement
 - Most survey responses are biased to impress the researcher
 - Social desirability bias has differential effects across respondents and topics

- Survey experiments are also valuable for measurement
 - Most survey responses are biased to impress the researcher
 - Social desirability bias has differential effects across respondents and topics
- List experiments make individual responses invisible to the researcher

- Survey experiments are also valuable for measurement
 - Most survey responses are biased to impress the researcher
 - Social desirability bias has differential effects across respondents and topics
- List experiments make individual responses invisible to the researcher
- Knowing this, hopefully the respondent answers more accurately
- Gonzalez-Ocantos et al (2010) list experiment on vote-buying

Im going to hand you a card that mentions various activities, and I would like for you to tell me if they were carried out by candidates or activists during the last electoral campaign. Please, do not tell me which ones, only HOW MANY:

- they put up campaign posters or signs in your neighborhood/city;
- they visited your home;
- they placed campaign advertisements on television or radio;
- they threatened you to vote for them.

Im going to hand you a card that mentions various activities, and I would like for you to tell me if they were carried out by candidates or activists during the last electoral campaign. Please, do not tell me which ones, only HOW MANY:

- they put up campaign posters or signs in your neighborhood/city;
- they visited your home;
- they gave you a gift or did you a favor;
- they placed campaign advertisements on television or radio;
- they threatened you to vote for them.

- Nicaragua 2008 municipal elections
- Direct Question: Have you received a gift or favour in exchange for your vote?
 - ► 3%

- Nicaragua 2008 municipal elections
- Direct Question: Have you received a gift or favour in exchange for your vote?
 - ▶ 3%

List experiment:

- Just the difference in mean responses between treatment and control lists
- ▶ 24% = 2.31 2.06

- Do respondents really understand anonymity?
- Bias towards a 'reasonable'/central number?
- Floor/ceiling effects

When should we use a lab or survey experiment?