14.750x: The Deep Determinants of Economic Development: Micro Evidence

Ben Olken

Introduction

- In the previous lecture, we introduced the idea that the effect of "institutions" may persist long after the institution itself is no longer there
  - E.g., the legacy of extractive institutions under colonialism
- But, how might this happen?
  - Something must happen in between to allow these things to persist. What is it?
  - And how do we know?

An example from Peru

Dell 2010: "The Persistent Effects of Peru's Mining Mita"

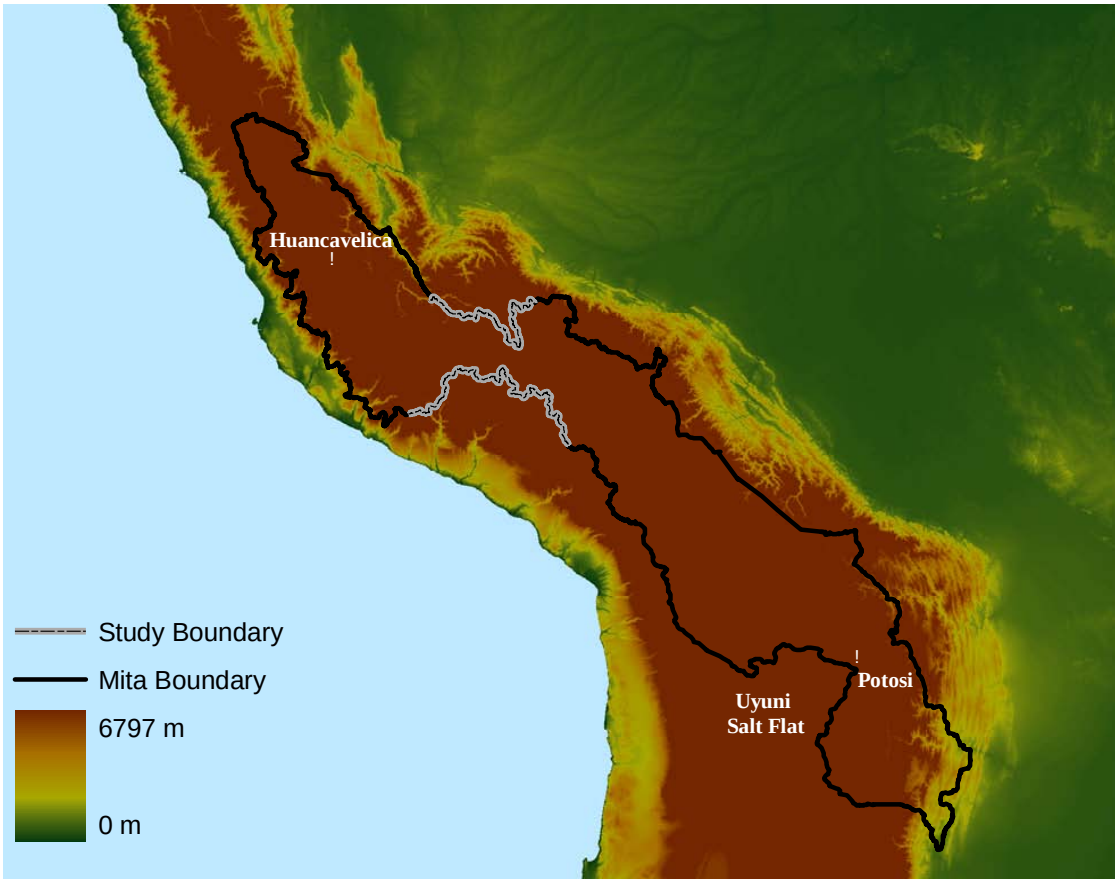
- What's the point of this paper?
- What was the Mita?
  - In colonial Peru and Bolivia, a major economic activity was mining, based at Potosí (silver) and Huancavelica (mercury)
  - To support the mines, from 1573 to 1812, indigenous communities were forced to send $\frac{1}{7}$ of their adult male population to work in the mines(!)
- Why might this matter at the time?
  - Local native elites were required to find the conscripts. What might this do?
  - Reduce trust, undermine institutions, encourage outmigration, make it hard to get good labor
- Why might it matter now?

Regression Discontinuity

- To identify the impact of the Mita, Dell takes advantage of the fact that the Mita had a well-defined border

The Mita

Figure 1



The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

Regression Discontinuity

- To identify the impact of the Mita, Dell takes advantage of the fact that the Mita had a well-defined border. Why might she do that?
- She focuses on the part of the border not coincident with mountains etc. Why?
- She then traces how communities on both sides of the border evolved from 1573 to present to tease out how the Mita may have had a long run impact
- This is an example of a general empirical design called "regression discontinuity."
- This will also come up a number of times this semester, so let's take a bit of a detour to explore what this is.

Regression discontinuity

- Consider a case where treatment is assigned based on a strict threshold.
- This is a sharp RD:

$$\begin{aligned} T_i &= 1 \text{ if } X_i \geq c \\ &= 0 \text{ if } X_i < c \end{aligned}$$

- Can you think of some examples of discontinuities?
  - Win an election if the most votes
  - Eligible for Medicare when you turn 65
  - Become a National Merit Semi-Finalist if your PSAT scores above a certain threshold
  - Legally allowed to buy alcohol when you turn 21
  - etc.

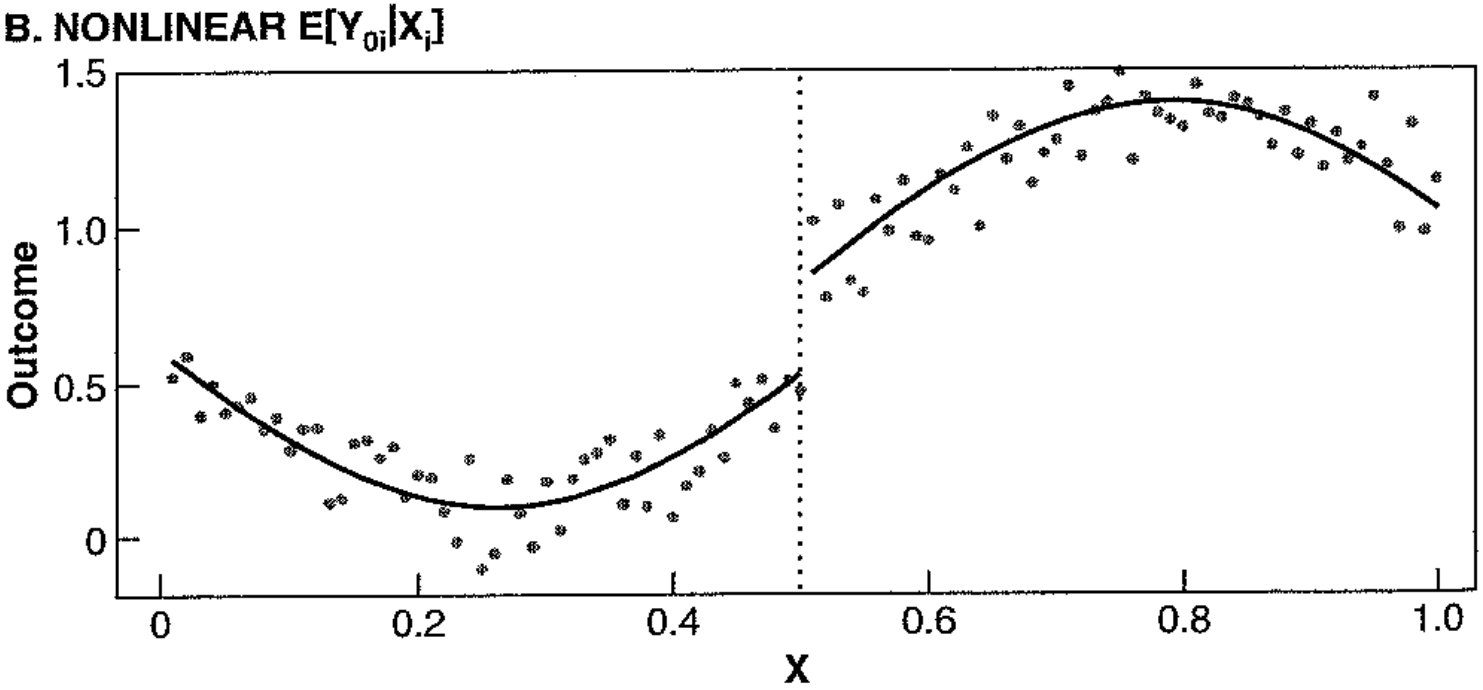
Regression discontinuity

- What does a discontinuity buy us?
- Lots of things in the world vary with X
  - e.g., you become more mature as you age
- The idea is that - except for the treatment, which varies discontinuously – the other covariates change continuously
  - e.g. you can legally buy alcohol when you turn 21, but your maturity grows continuously as you age
  - e.g., better popular politicians running for Governor get more votes, but there is a huge difference between getting 1 more vote than the competitor and getting 1 fewer vote, in that one of you becomes Governor and the other does not
- Technically, the assumption is that:

$$lim_{x \downarrow c} E[Y_i(0)|Xi = x] = lim_{x \uparrow c} E[Y_i(0)|Xi = x]$$

where $Y_i(0)$ is the counterfactual outcome variable if there had been no treatment

RD in pictures



Fuzzy regression discontinuity

- Note that in some of these examples the discontinuity may not be strict
  - e.g., some people ¡ 21 were able to buy alcohol
  - But, it discontinuously becomes easier to buy alcohol once you turn 21
  - This is called a "fuzzy" RD, and here what we require is

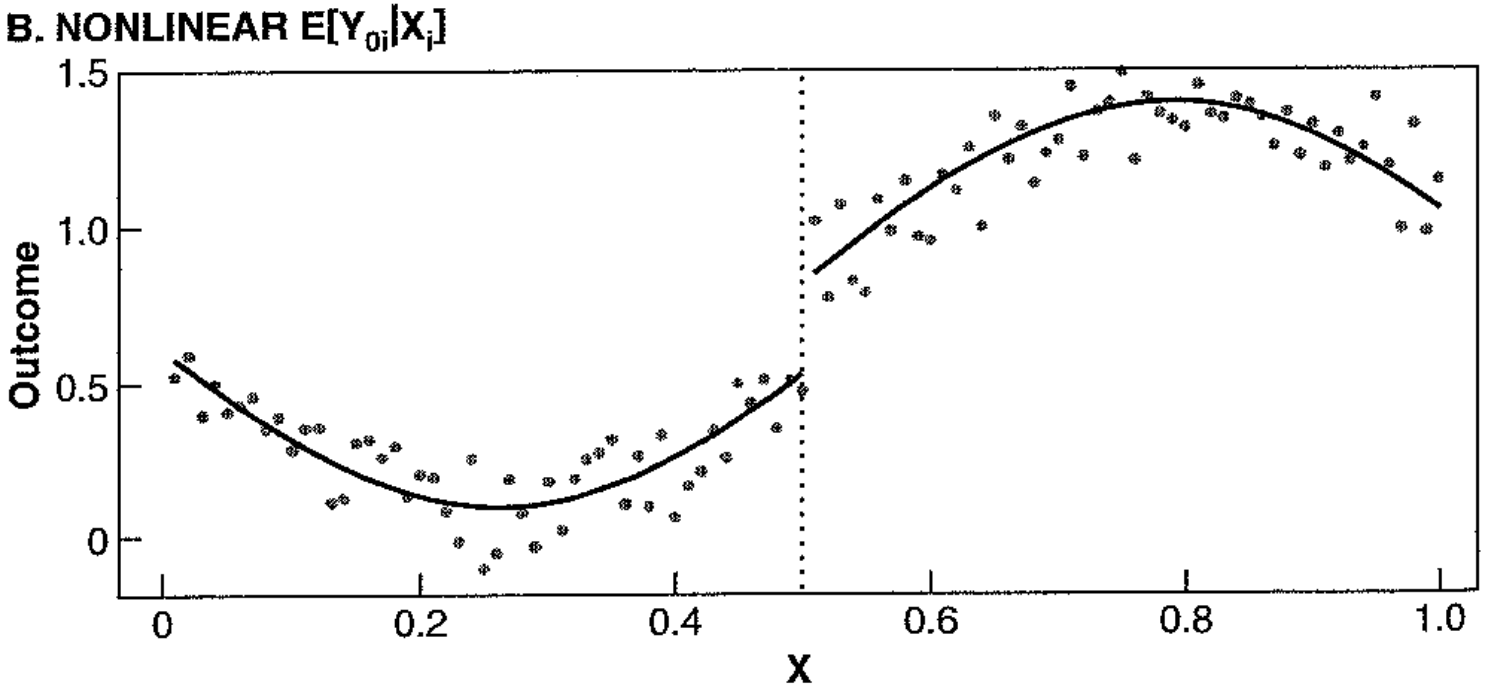
$$lim_{x \downarrow c} pr(T_i = 1 | X_i = x) \neq lim_{x \uparrow c} pr(T_i = 1 | X_i = x)$$

  - The same general ideas hold – we use the "jump" in T at the discontinuity to gain identification
- RD in practice
  - RD regressions tend to estimate equations like

$$y = \beta \mathbf{1}_{X_i > x} + f(X_i) + \varepsilon$$

    where $\mathbf{1}_{X_i > x}$ is a dummy variable for being above the discontinuity, and $f(X_i)$ is a very flexible function of the X_i (e.g. quadratic, cubic, etc)
  - β is the coefficient of interest

RD in pictures

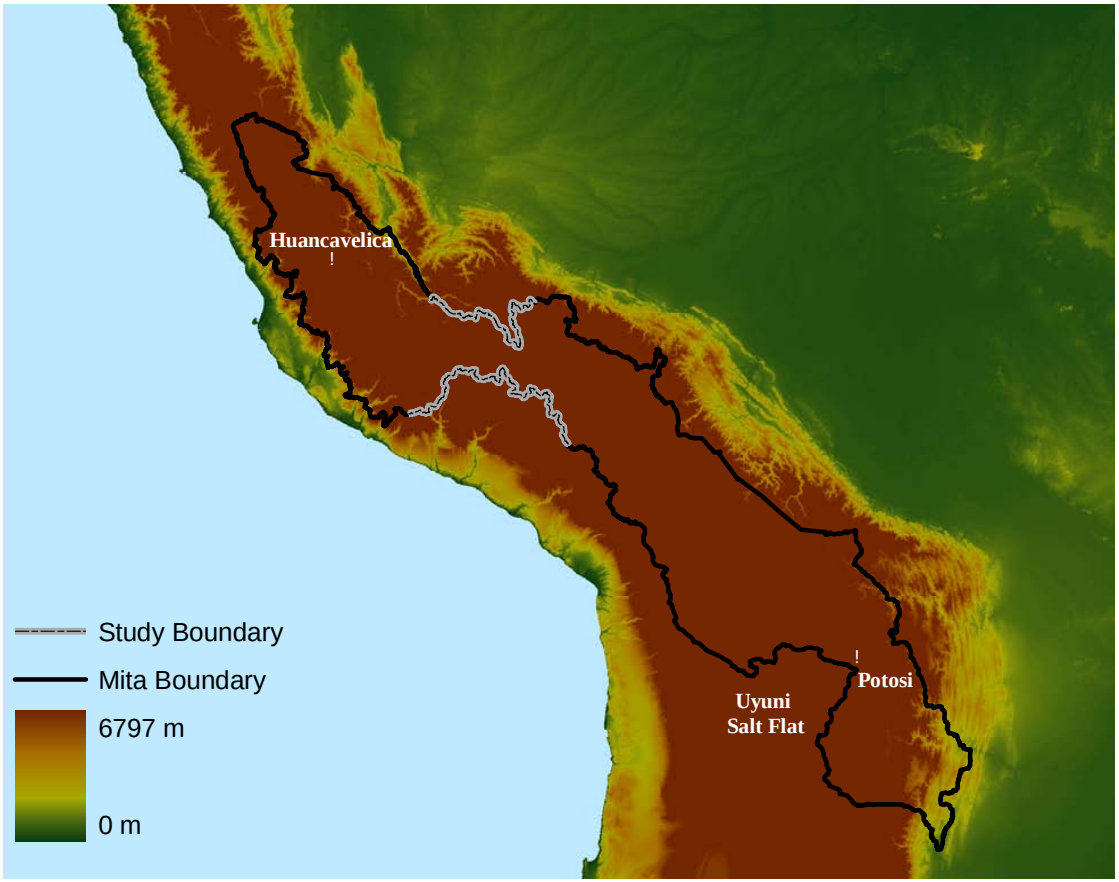


Back to the Mita

- In the Mita case, what is X? What is the discontinuity?

Back to the Mita

Figure 1



The *mita* boundary is in black and the study boundary in light gray. Districts falling inside the contiguous area formed by the *mita* boundary contributed to the *mita*. Elevation is shown in the background.

Back to the Mita

- In the Mita case, what is X? What is the discontinuity?
- This case is a little tricky since the discontinuity is a border, and X is really in two dimensions
- So in this case, Dell does it several ways
  - polynomial in latitude and longitude
  - distance to the mine at Potosi
  - distance to the border
- She also "zooms in" to get closer to the border – this is also common practice

Checking the discontinuity

- How might you check that the discontinuity – the border – is a good empirical design?
- You should check that nothing else systematically varies at the border
- In this case, what could you check?
  - Geography (elevation, slope)
  - Pre-period characteristics (log 1572 tribute rate)
- (Aside: why do we log so many variables? If a variable is in logs, how do we interpret coefficients?)

Checking the discontinuity

TABLE I

SUMMARY STATISTICS[a]

	Sample Falls Within											
	<100 km of *Mita* Boundary			<75 km of *Mita* Boundary			<50 km of *Mita* Boundary			<25 km of *Mita* Boundary		
	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.
GIS Measures												
Elevation	4042	4018	[188.77] (85.54)	4085	4103	[166.92] (82.75)	4117	4096	[169.45] (89.61)	4135	4060	[146.16] (115.15)
Slope	5.54	7.21	[0.88]* (0.49)***	5.75	7.02	[0.86] (0.52)**	5.87	6.95	[0.95] (0.58)*	5.77	7.21	[0.90] (0.79)*
Observations	177	95		144	86		104	73		48	52	
% Indigenous	63.59	58.84	[11.19] (9.76)	71.00	64.55	[8.04] (8.14)	71.01	64.54	[8.42] (8.43)	74.47	63.35	[10.87] (10.52)
Observations	1112	366		831	330		683	330		329	251	
Log 1572 tribute rate	1.57	1.60	[0.04] (0.03)	1.57	1.60	[0.04] (0.03)	1.58	1.61	[0.05] (0.04)	1.65	1.61	[0.02]* (0.03)

(*Continues*)

Checking the discontinuity

TABLE I—*Continued*

	Sample Falls Within											
	<100 km of *Mita* Boundary			<75 km of *Mita* Boundary			<50 km of *Mita* Boundary			<25 km of *Mita* Boundary		
	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.	Inside	Outside	s.e.
% 1572 tribute to Spanish Nobility	59.80	63.82	[1.39]*** (1.36)***	59.98	63.69	[1.56]** (1.53)**	62.01	63.07	[1.12] (1.34)	61.01	63.17	[1.58] (2.21)
Spanish Priests	21.05	19.10	[0.90]** (0.94)**	21.90	19.45	[1.02]** (1.02)**	20.59	19.93	[0.76] (0.92)	21.45	19.98	[1.01] (1.33)
Spanish Justices	13.36	12.58	[0.53] (0.48)*	13.31	12.46	[0.65] (0.60)	12.81	12.48	[0.43] (0.55)	13.06	12.37	[0.56] (0.79)
Indigenous Mayors	5.67	4.40	[0.78] (0.85)	4.55	4.29	[0.26] (0.29)	4.42	4.47	[0.34] (0.33)	4.48	4.42	[0.29] (0.39)
Observations	63	41		47	37		35	30		18	24	

Results

Does the Mita matter today?

TABLE II

LIVING STANDARDS[a]

	Dependent Variable						
	Log Equiv. Household Consumption (2001)			Stunted Growth, Children 6–9 (2005)			
Sample Within:	<100 km of Bound. (1)	<75 km of Bound. (2)	<50 km of Bound. (3)	<100 km of Bound. (4)	<75 km of Bound. (5)	<50 km of Bound. (6)	Border District (7)
			Panel A. Cubic Polynomial in Latitude and Longitude				
Mita	−0.284	−0.216	−0.331	0.070	0.084*	0.087*	0.114**
	(0.198)	(0.207)	(0.219)	(0.043)	(0.046)	(0.048)	(0.049)
R^2	0.060	0.060	0.069	0.051	0.020	0.017	0.050
			Panel B. Cubic Polynomial in Distance to Potosí				
Mita	−0.337***	−0.307***	−0.329***	0.080***	0.078***	0.078***	0.063*
	(0.087)	(0.101)	(0.096)	(0.021)	(0.022)	(0.024)	(0.032)
R^2	0.046	0.036	0.047	0.049	0.017	0.013	0.047
			Panel C. Cubic Polynomial in Distance to *Mita* Boundary				
Mita	−0.277***	−0.230**	−0.224**	0.073***	0.061***	0.064***	0.055*
	(0.078)	(0.089)	(0.092)	(0.023)	(0.022)	(0.023)	(0.030)
R^2	0.044	0.042	0.040	0.040	0.015	0.013	0.043
Geo. controls	yes	yes	yes	yes	yes	yes	yes
Boundary F.E.s	yes	yes	yes	yes	yes	yes	yes
Clusters	71	60	52	289	239	185	63
Observations	1478	1161	1013	158,848	115,761	100,446	37,421

TABLE VI

LAND TENURE AND LABOR SYSTEMS[a]

	Dependent Variable				
	Haciendas per District in 1689 (1)	*Haciendas* per 1000 District Residents in 1689 (2)	Percent of Rural Tributary Population in *Haciendas* in ca. 1845 (3)	Percent of Rural Population in *Haciendas* in 1940 (4)	Land Gini in 1994 (5)
	Panel A. Cubic Polynomial in Latitude and Longitude				
Mita	−12.683***	−6.453**	−0.127*	−0.066	0.078
	(3.221)	(2.490)	(0.067)	(0.086)	(0.053)
R^2	0.538	0.582	0.410	0.421	0.245
	Panel B. Cubic Polynomial in Distance to Potosí				
Mita	−10.316***	−7.570***	−0.204**	−0.143***	0.107***
	(2.057)	(1.478)	(0.082)	(0.051)	(0.036)
R^2	0.494	0.514	0.308	0.346	0.194
	Panel C. Cubic Polynomial in Distance to *Mita* Boundary				
Mita	−11.336***	−8.516***	−0.212***	−0.120***	0.124***
	(2.074)	(1.665)	(0.060)	(0.045)	(0.033)
R^2	0.494	0.497	0.316	0.336	0.226
Geo. controls	yes	yes	yes	yes	yes
Boundary F.E.s	yes	yes	yes	yes	yes
Mean dep. var.	6.500	5.336	0.135	0.263	0.783
Observations	74	74	81	119	181

Channels of Persistence

Education

TABLE VII

EDUCATION[a]

	Dependent Variable		
	Literacy 1876 (1)	Mean Years of Schooling 1940 (2)	Mean Years of Schooling 2001 (3)
	Panel A. Cubic Polynomial in Latitude and Longitude		
Mita	−0.015	−0.265	−1.479*
	(0.012)	(0.177)	(0.872)
R^2	0.401	0.280	0.020
	Panel B. Cubic Polynomial in Distance to Potosí		
Mita	−0.020***	−0.181**	−0.341
	(0.007)	(0.078)	(0.451)
R^2	0.345	0.187	0.007
	Panel C. Cubic Polynomial in Distance to *Mita* Boundary		
Mita	−0.022***	−0.209***	−0.111
	(0.006)	(0.076)	(0.429)
R^2	0.301	0.234	0.004
Geo. controls	yes	yes	yes
Boundary F.E.s	yes	yes	yes
Mean dep. var.	0.036	0.470	4.457
Clusters	95	118	52
Observations	95	118	4038

Channels of Persistence

Roads

TABLE VIII

ROADS[a]

	Dependent Variable		
	Density of Local Road Networks (1)	Density of Regional Road Networks (2)	Density of Paved/Gravel Regional Roads (3)
	Panel A. Cubic Polynomial in Latitude and Longitude		
Mita	0.464	−29.276*	−22.426*
	(18.575)	(16.038)	(12.178)
R^2	0.232	0.293	0.271
	Panel B. Cubic Polynomial in Distance to Potosí		
Mita	−1.522	−32.644***	−30.698***
	(12.101)	(8.988)	(8.155)
R^2	0.217	0.271	0.256
	Panel C. Cubic Polynomial in Distance to *Mita* Boundary		
Mita	0.535	−35.831***	−32.458***
	(12.227)	(9.386)	(8.638)
R^2	0.213	0.226	0.208
Geo. controls	yes	yes	yes
Boundary F.E.s	yes	yes	yes
Mean dep. var.	85.34	33.55	22.51
Observations	185	185	185

Bottom line

- What has the paper shown?
- Looking over time, areas inside the Mita area had:
  - Fewer haciendas
  - Fewer public goods (roads)
  - Less education
  - Less income today
  - And more likely to have Shining Path violence
- Suggests channels of institutional persistence

Generalizing?

- Do you think the patterns from the Mita example are likely to be general?
- That is – what does this tell us about the long-run impacts of colonialism more generally?
- Reasons things might go in the other direction?

Another example

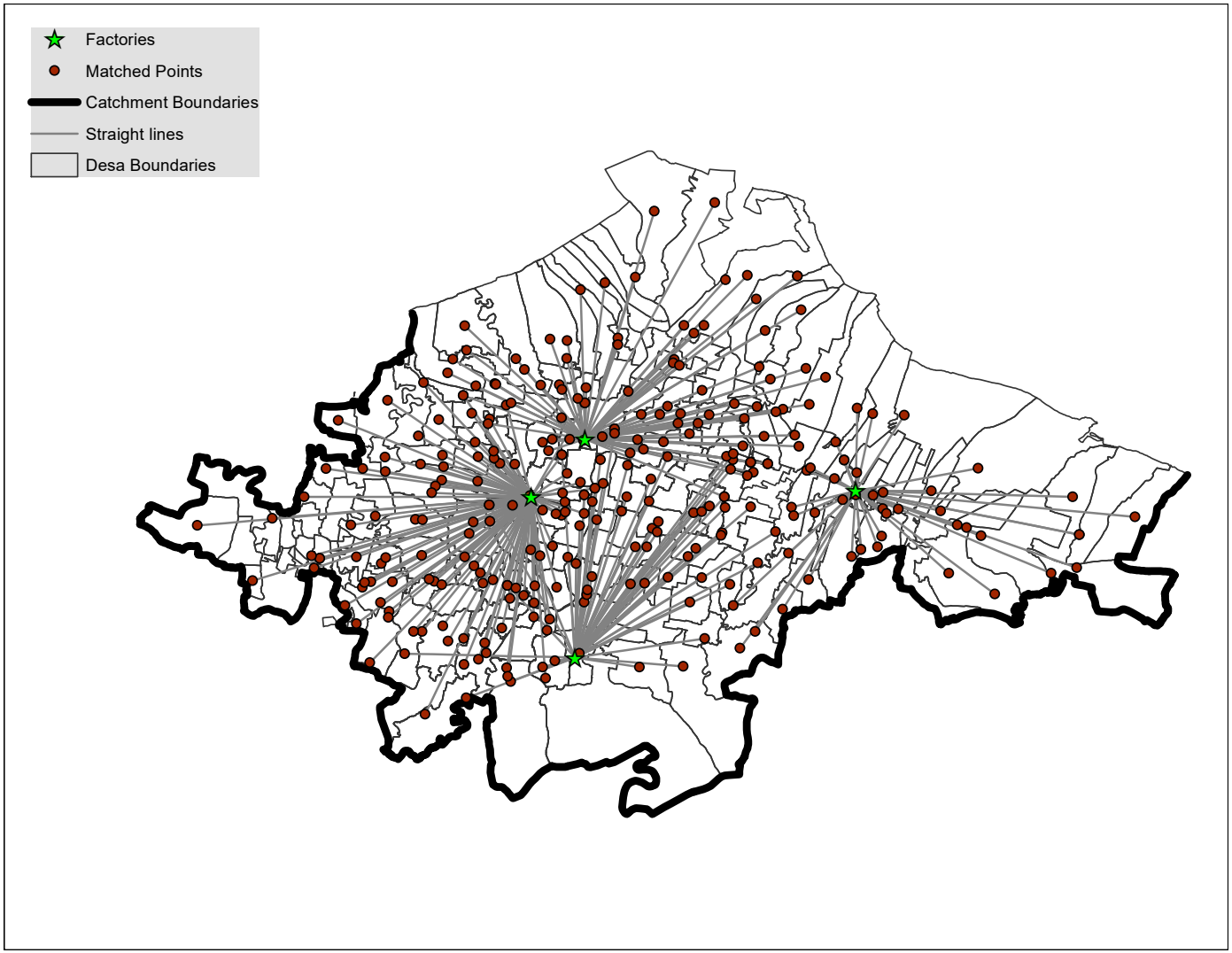
Dell and Olken (2018): The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java

- Dell and I look at an important historical example: Dutch colonial experience in Java (now Indonesia)
- Key facts:
  - From the early 1830s through the 1870s, the Dutch colonial state forced peasants along Java's northern coast to cultivate sugar, which was then processed in nearby colonial sugar factories for export to Europe.
  - The Dutch established 94 water-powered sugar processing factories. The cane had to be grown nearby.
  - The Dutch constructed a catchment area with a radius of approximately five kilometers around each factory, and forced all villages within the catchment area to grow cane.
  - Reports from the 1860s show that over 2.5 million forced workers labored in the sugar factories or related services.
- Based on the Mita paper, what would you expect? Other views?

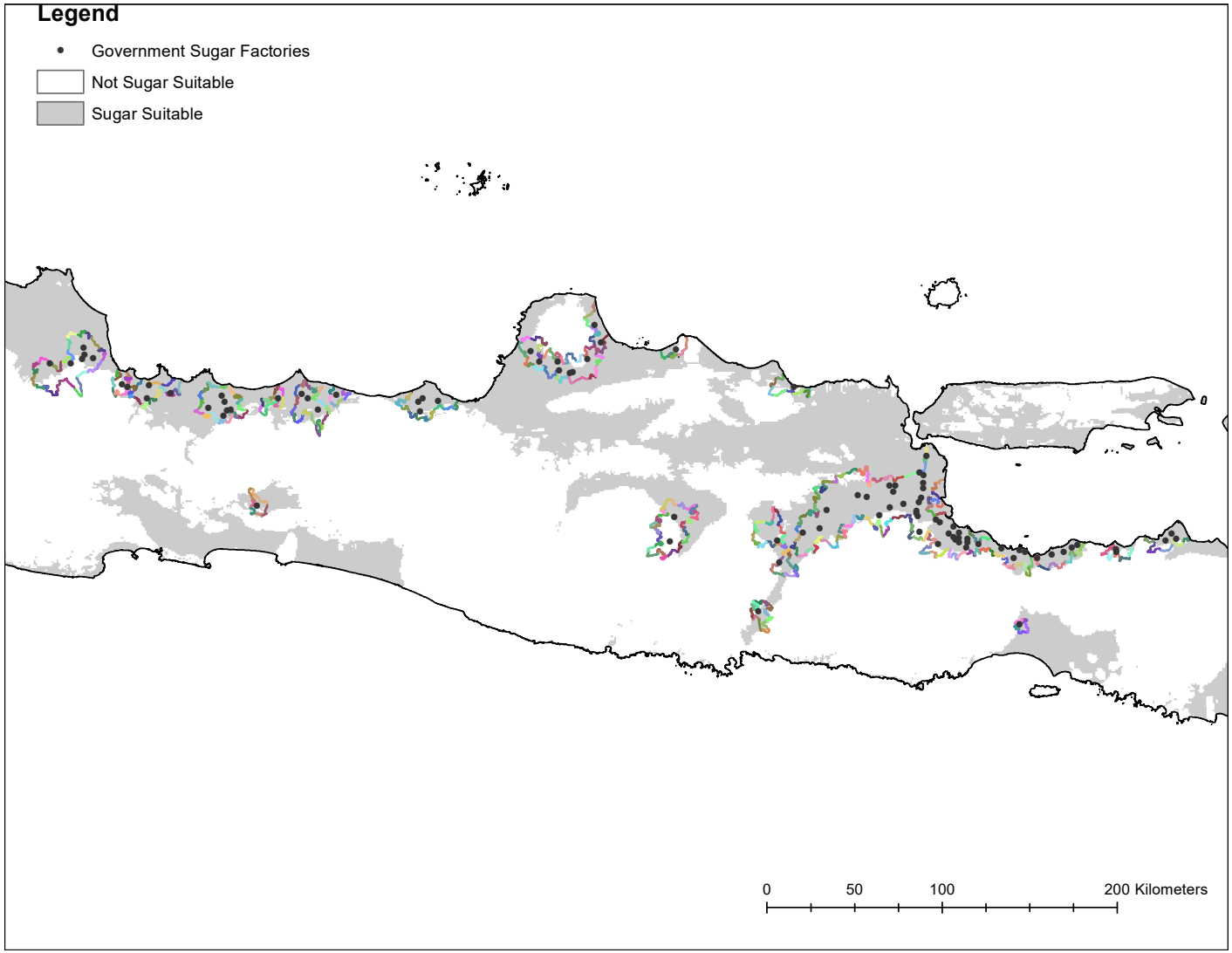
Methodologies

- We use two different strategies in this paper
- How to figure out the impact of being forced to grow cane?
- This is an RD, just like the Mita paper

RD example


Factories
Matched Points
Catchment Boundaries
Straight lines
Desa Boundaries

RD example

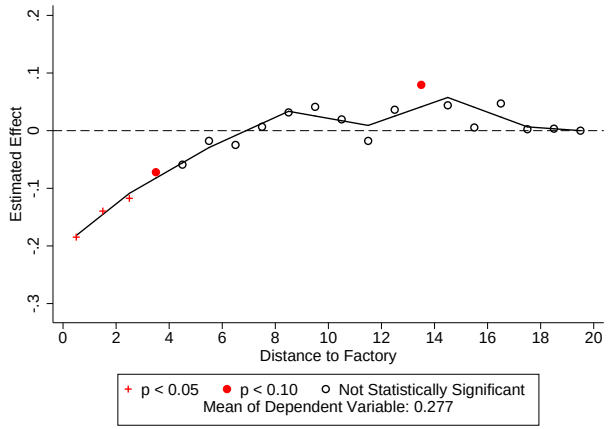

Legend
Government Sugar Factories
Not Sugar Suitable
Sugar Suitable
0
50
100
200 Kilometers

What else?

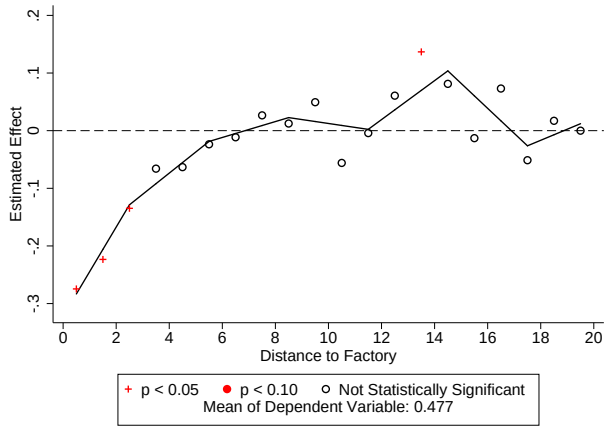
- But this is only half the story. Why?
- The factories themselves may have had impacts
- To explore this, we compare areas near the actual factories to areas near where the Dutch *could* have build factories

Impact of factories on long-run development

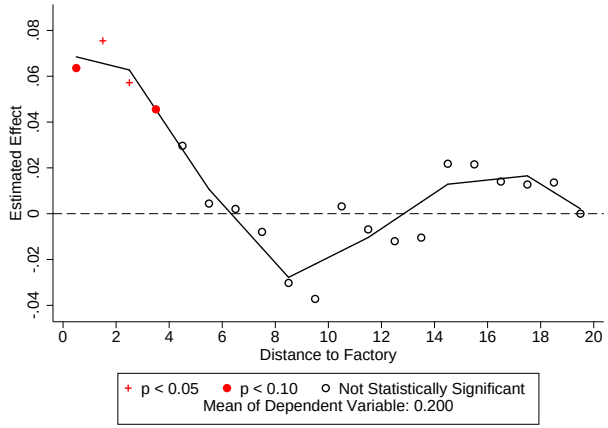
(a) Agriculture (Susenas 2001-11)



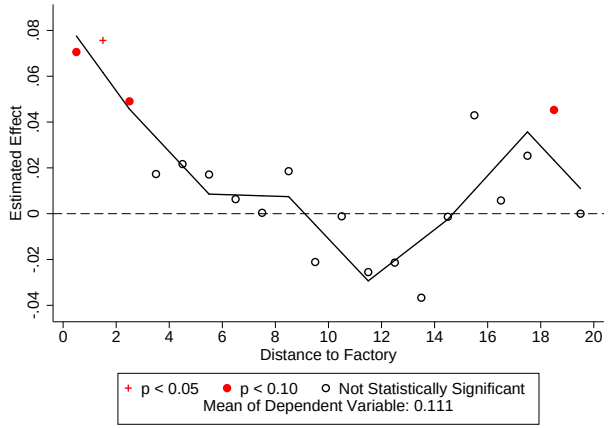
(b) Agriculture (Census 1980)



(c) Manufacturing (Susenas 2001-11)

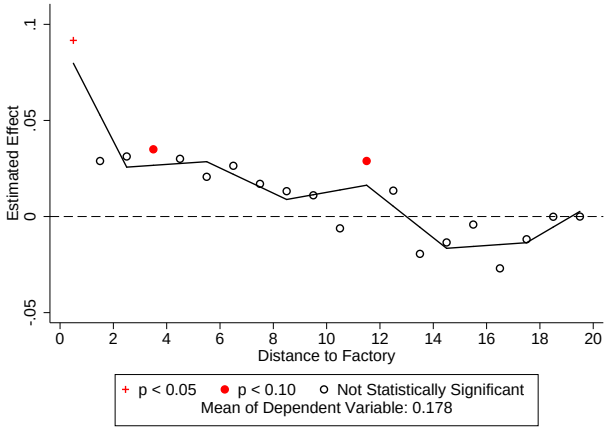


(d) Manufacturing (Census 1980)

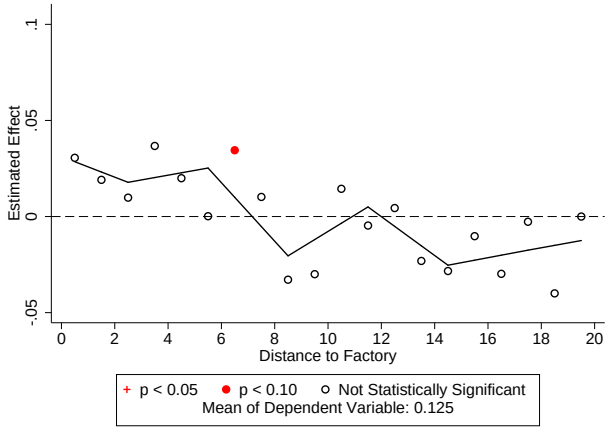


Impact of factories on long-run development

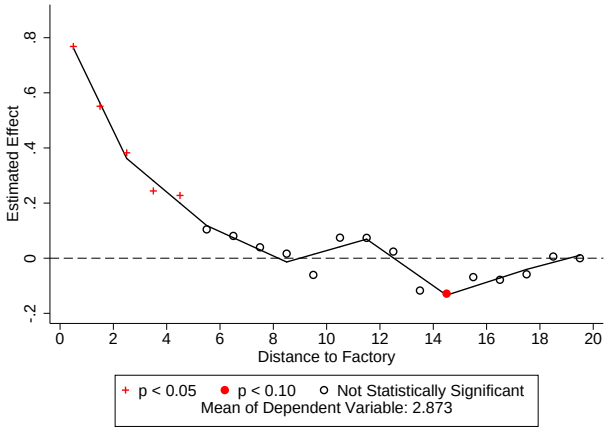
(e) Retail (Susenas 2001-11)



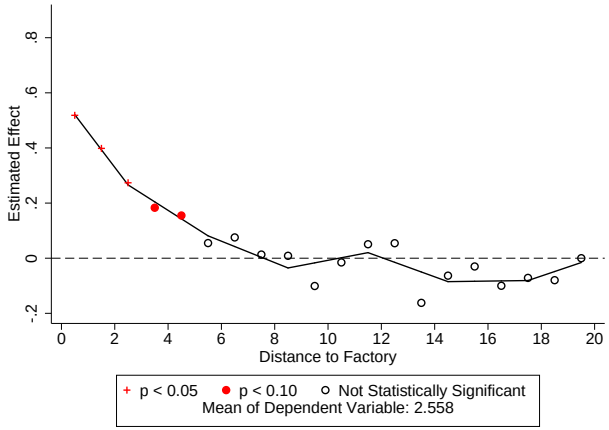
(f) Retail (Census 1980)



(g) Log Pop. Density (PODES 2003)

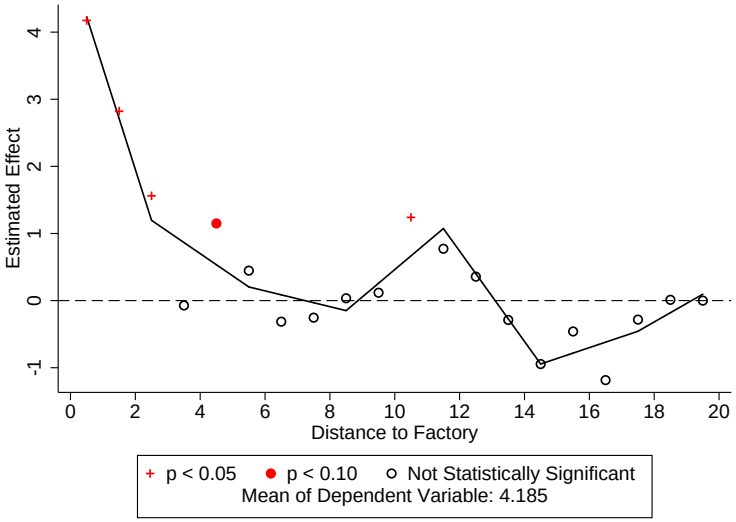


(h) Log Pop. Density (PODES 1980)

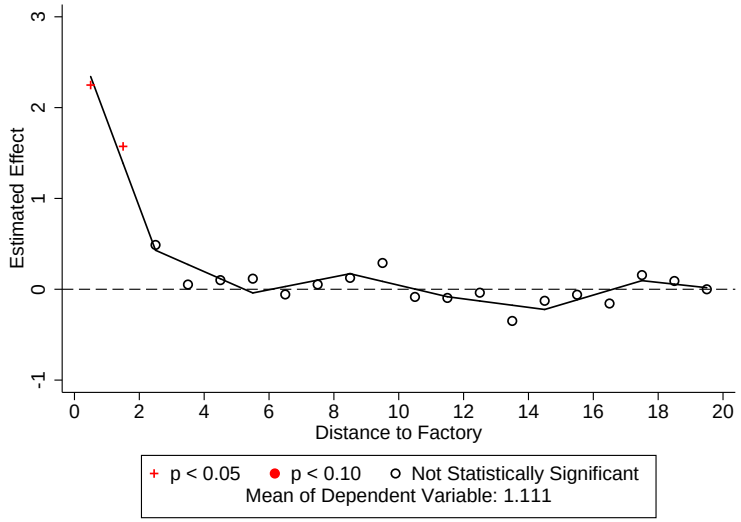


Impact of factories on long-run development

(a) Colonial Road Density (1900)

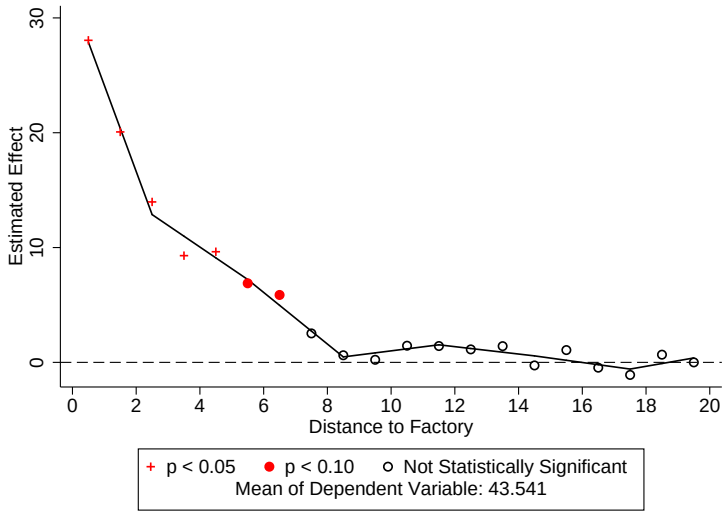


(b) Colonial Railroad Density (1900)

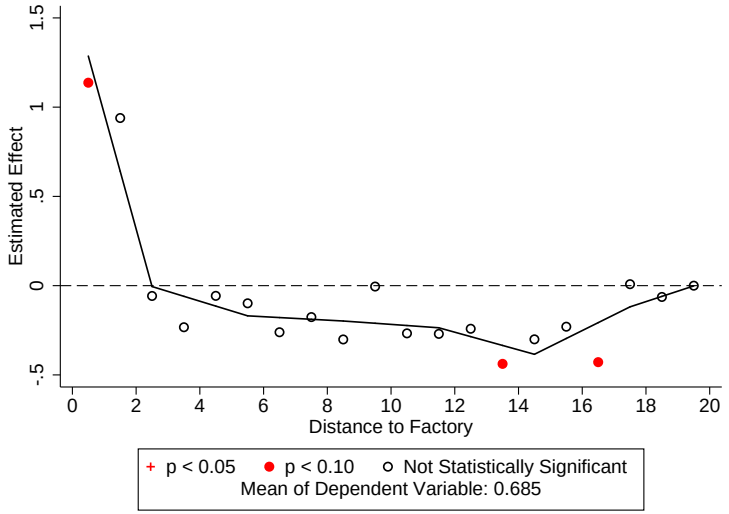


Impact of factories on long-run development

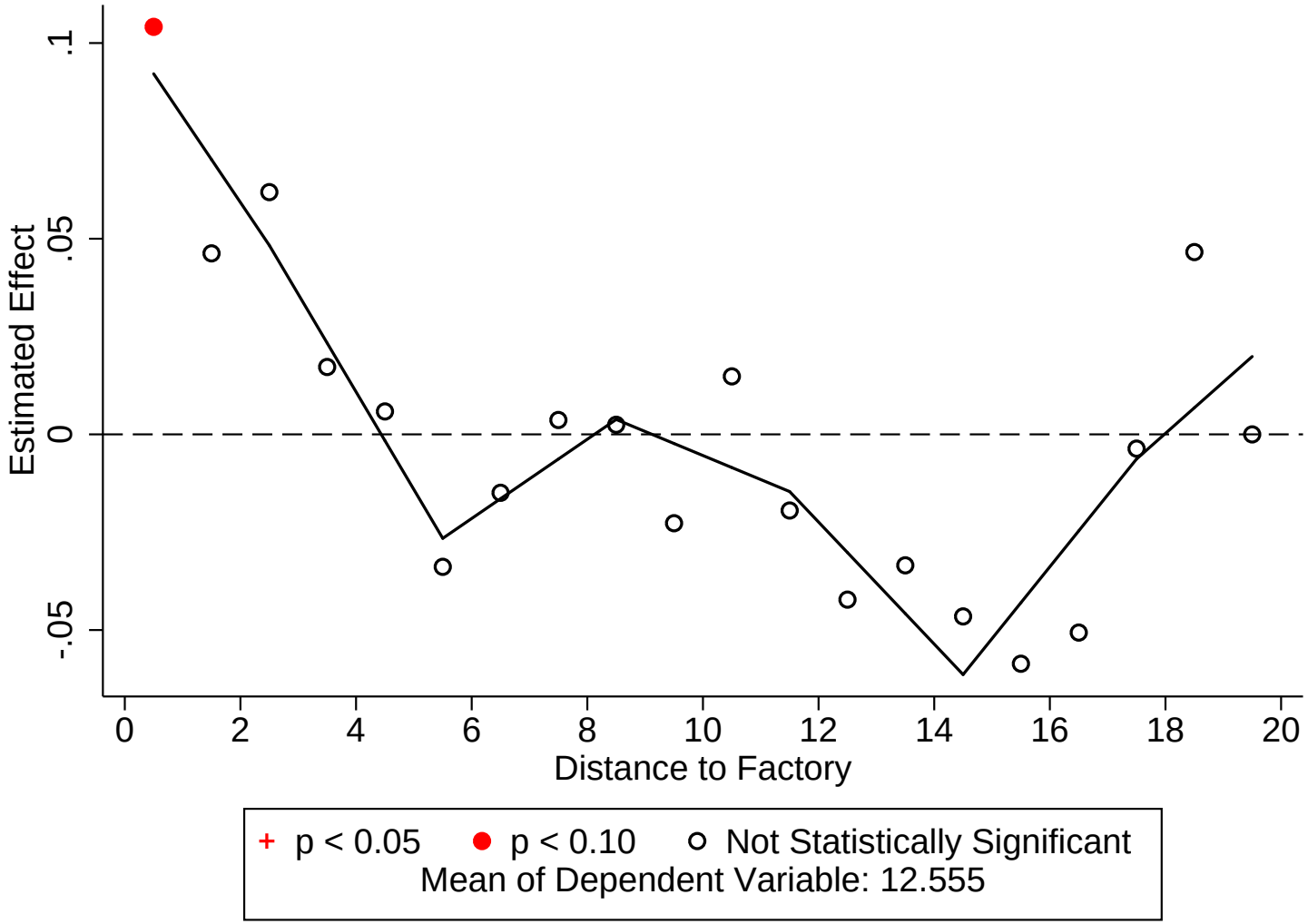
(e) Local Road Density (2017)



(f) Railroad Density (2017)



Impact of factories on long-run development



Impact of being in a subjected village

	Num. Manuf. Firms (1)	Manuf. Emp. Share (2)	Log Population Density (2003) (3)	Log Population Density (1980) (4)	Log Equiv. Consumption (5)
Cultivation	21.734	0.025	0.065	0.062	0.006
	(8.639)	(0.018)	(0.034)	(0.032)	(0.011)
Obs	4,549	4,549	4,550	4,107	144,046
Clusters	383	383	383	380	381
Mean	71.72	0.16	2.87	2.54	12.55

Discussion

- So? What do you make of all this?

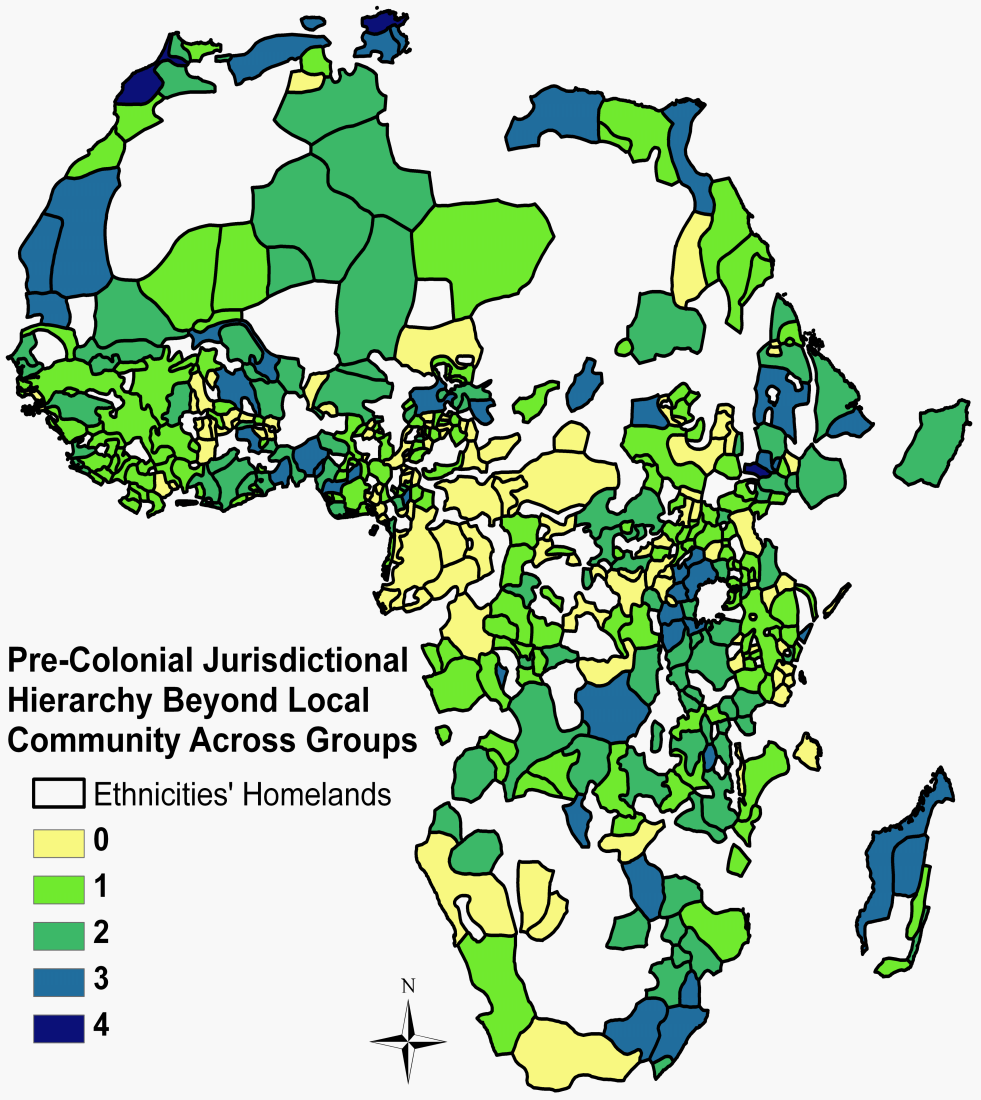
Ethnic institutions or national institutions

Michalopoulos and Papaioannou: "Divide and Rule or Rule of the Divided? Evidence from Africa"

- Acemoglu, Johnson, and Robinson argued it was national institutions (influenced by colonialization) that affect modern development
- However, others have argued that in fact, pre-colonial ethnic institutions may also be important – how pre-colonial ethnic institutions were organized may also affect contemporary development.
- Why? One reason is that in many African countries, national governments have little power outside of national capitals, so old, pre-colonial institutions may largely be in place
- MP seek to answer this question, using a similar RD approach
  - Obtain data on pre-colonial ethnic boundaries and ethnic instituions from Murdock's *Ethnolinguistic Atlas of Africa*
  - Use the fact that national borders in Africa were drawn artificially
  - Use data on light intensity at night (remember the Korea picture?) to see how economic activity changes across the national border within

Ethnic Homelands before Colonization

Pre-Colonial Institutions before Colonization

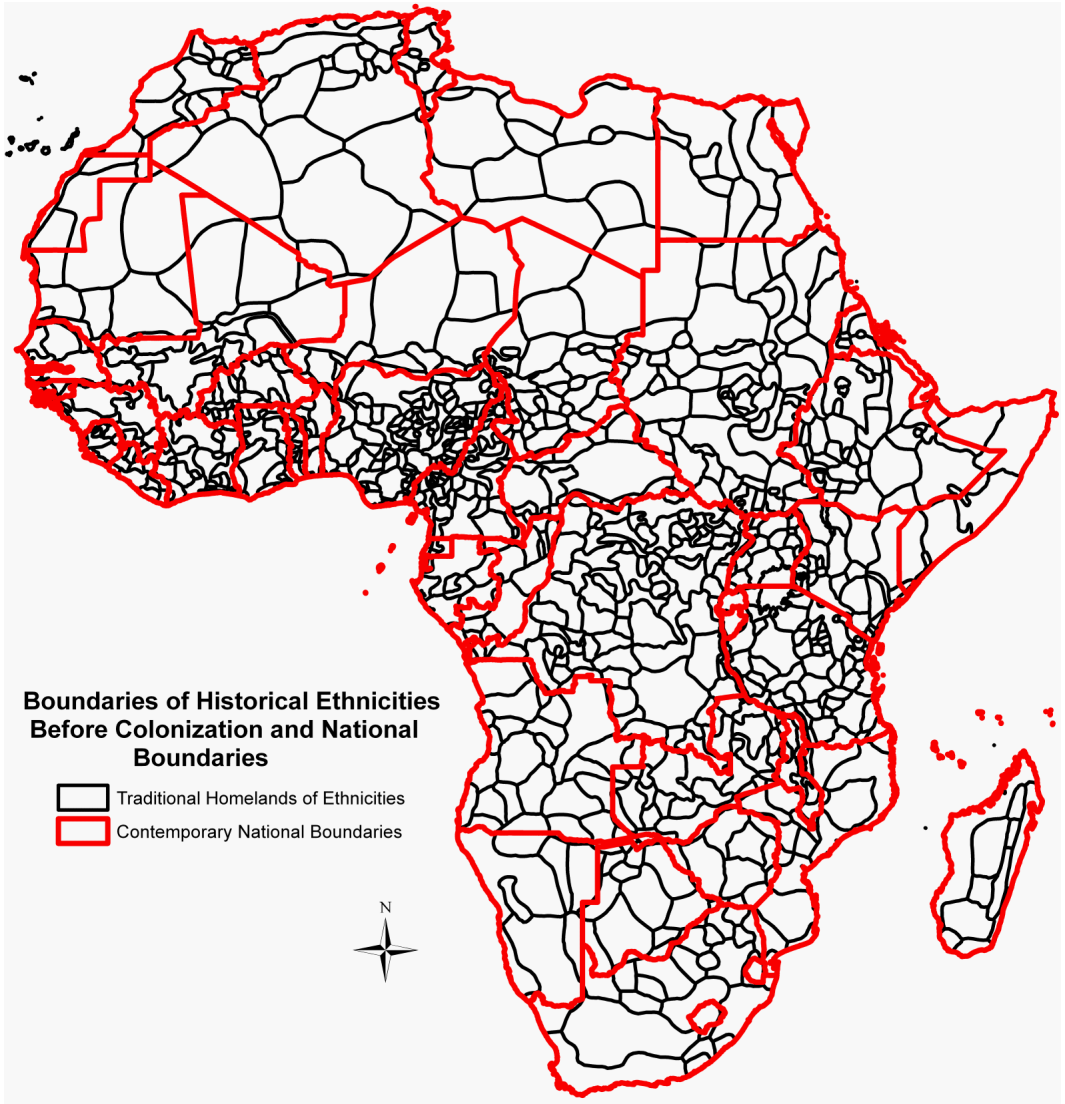


How African Borders Were Created

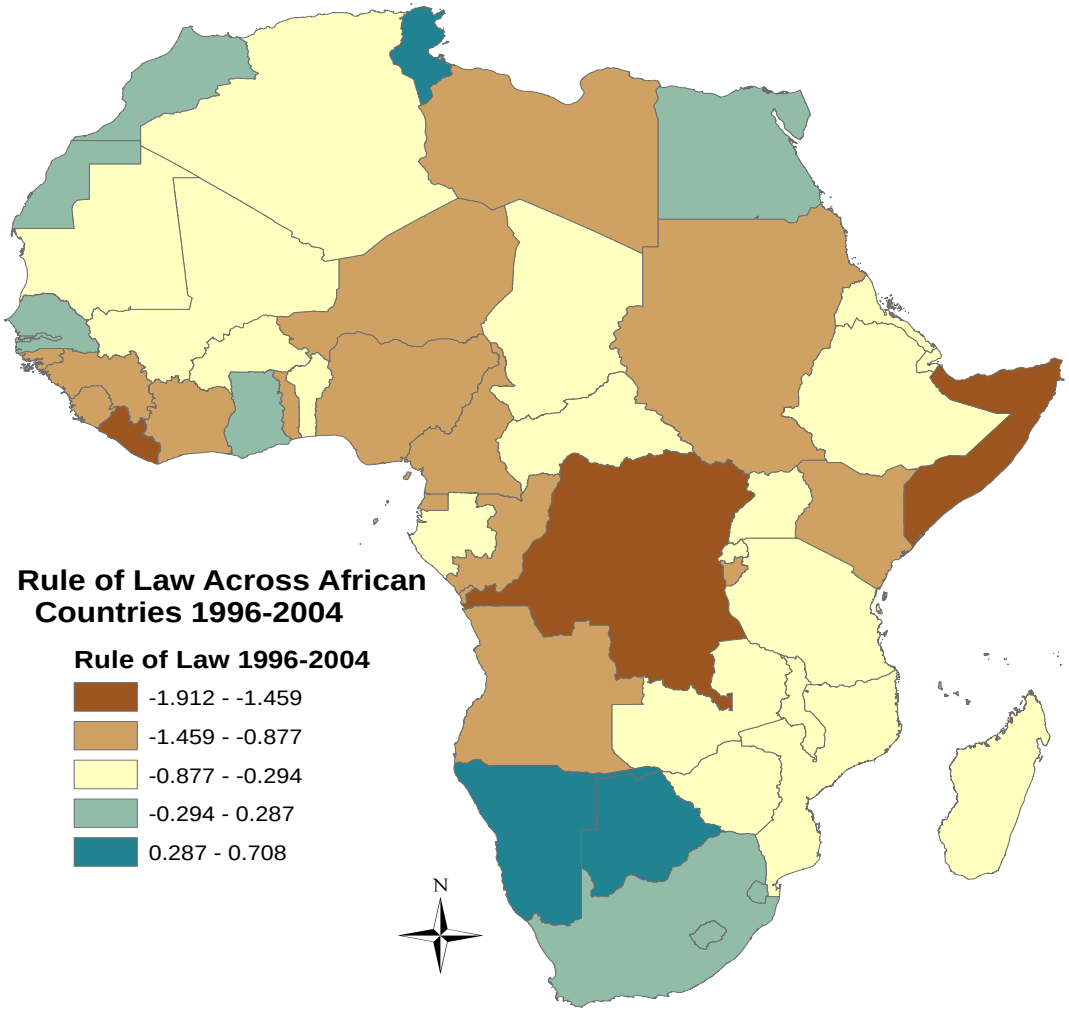
How African Borders Were Created

- African borders were decided in Berlin in 1884/1885 and 1890s
  - Colonizers had not even explored most of Africa when borders were agreed.
  - No ethnicity-specific measure predicts which ethnicities were partitioned
  - Drawing involved protectorates, large (free-trade) areas, and spheres of in‡uence rather than potential states.

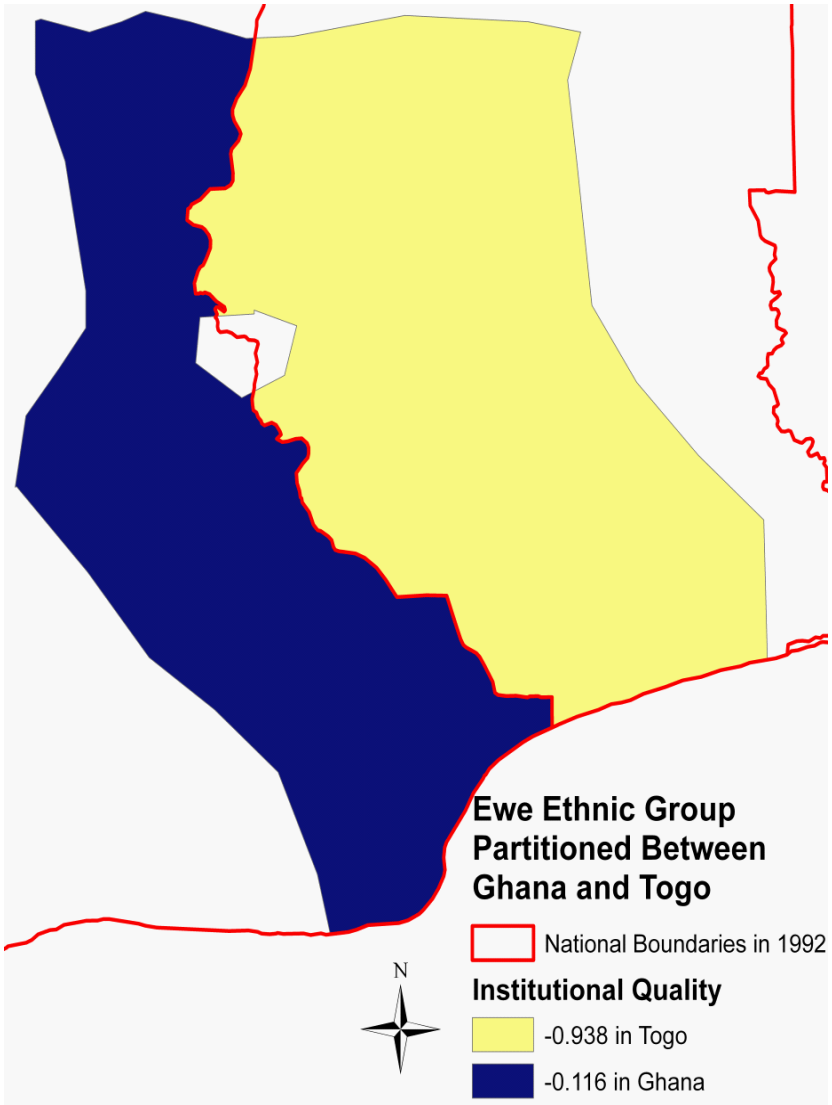
African Borders vs. Ethnic Borders



African Borders and National Institutions



They examine partitioned ethnic groups



Results show that national institutions matter much less within ethnicities

Panel A: Global Polynomial Control Function Method

	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	0.8153***	0.0644	0.6432**	0.0349		
	(0.2645)	(0.2795)	(0.2591)	(0.3294)		
Control of Corruption					0.9522***	-0.1235
					(0.2990)	(0.3214)
Adjusted R-squared	0.301	0.843	0.403	0.846	0.298	0.841
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No
Geographic Controls	No	No	Yes	Yes	No	No
Observations	454	454	454	454	454	454

Conclusions

- What have we learned from this?
  - Long-lasting effects of institutions – colonial and ethnic
  - Visible within countries, not just between countries