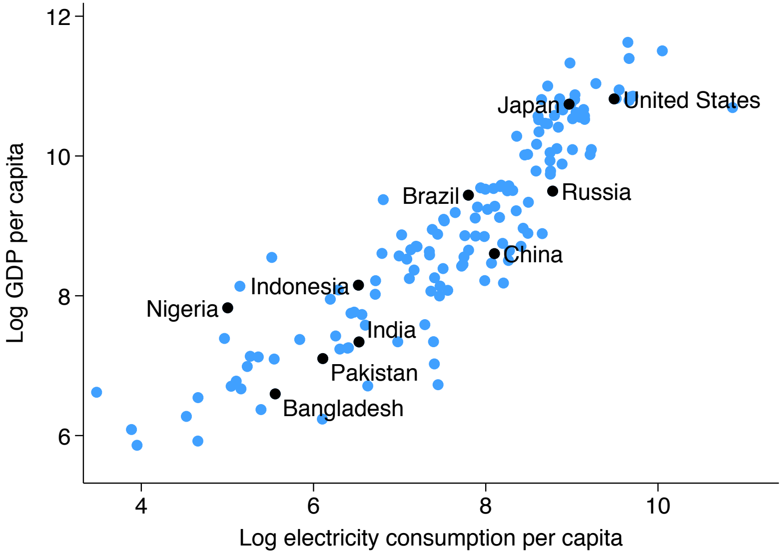
Lecture 19:
Policy Lab
Does rural electrification work? II

PPHA 34600
Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

From last time: sorry not sorry

From last time: rural electrification policy lab



Source: World Bank tables

From last time: rural electrification policy lab

What is the causal effect of rural electrification on economic development?

This is not totally straightforward to answer:

- Naive estimator: compare electrified to non-electrified places
- **Why is this problematic?**
- Electrified places might be...:
  - Growing faster (slower) than non-electrified places
  - More (less) politically connected
  - Have other infrastructure (roads, etc)
  - Be wealthier (less wealthy)
  - Etc

→ There are many forms of selection bias!

From last time: Does rural electrification work (in SA)?

Dinkelman (2011) is a seminal study of rural electrification:

- Estimates effects in post-Apartheid South Africa
- Deals with identification with an instrumental variables approach
- IV: Land gradient
- **Finding:** Rural electrification causes large changes in female employment

Alternative approaches to handling selection bias

We may find the geographic IV unsatisfying...

Today: two additional estimation approaches:

1. Lee, Miguel, Wolfram (2019): **RCT**
2. Burlig and Preonas (2016): **RD**

Second paper: Lee, Miguel, Wolfram (2019)

This is the first prominent econ paper to randomize RE:

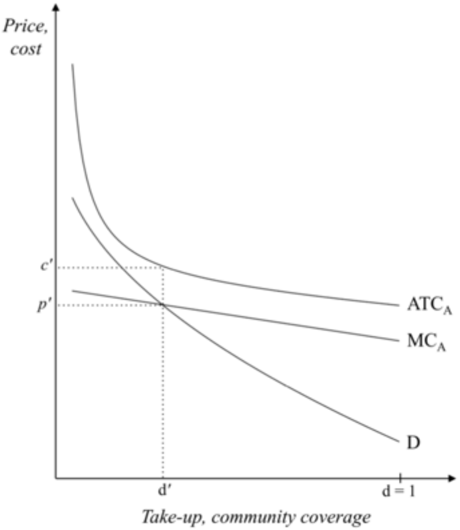
Research question: What are the "economics of rural electrification"?

→ AKA, what are the costs and benefits of rural electrification?

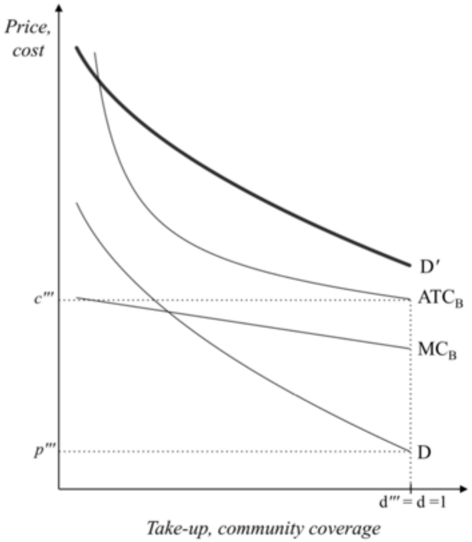
→ What are the effects of electrification on a variety of "benefits?"

→ And how much does it cost to achieve these benefits?

Theoretical framework: the utility as a natural monopoly



Theoretical framework: the utility as a natural monopoly



LMW (2019): Context

Lee, Miguel, and Wolfram study rural electrification in Kenya:

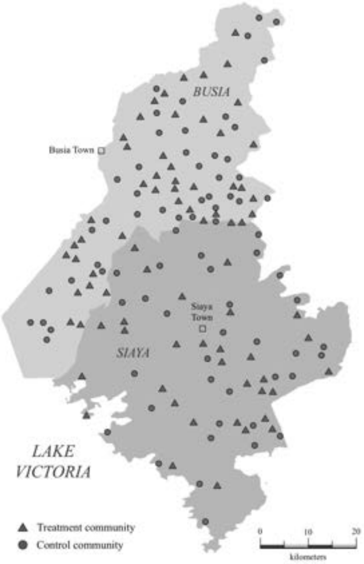
- Kenya's grid is green(ish): lots of hydro and geothermal
- Lots to do: Installed MW to increase 10 fold by 2031
- Huge increase in electricity access in recent years: REA pushed to electrify public places...
- ... but households remained at low levels (only 32% electrified in 2014)

LMW (2019): Context

Lee, Miguel, and Wolfram study rural electrification in Kenya:

- Kenya's grid is green(ish): lots of hydro and geothermal
- Lots to do: Installed MW to increase 10 fold by 2031
- Huge increase in electricity access in recent years: REA pushed to electrify public places...
- ... but households remained at low levels (only 32% electrified in 2014)
- Households could pay for grid connections: **$398** within 600m of a transformer

LMW (2019): Context

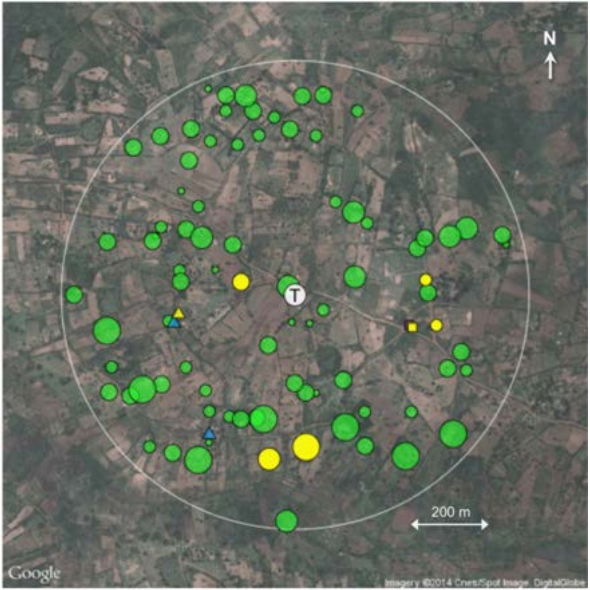


LMW (2019): Context

LMW (2019): Context

		Nationwide county percentiles		
	Study region	25th	50th	75th
Total population	793,125	528,054	724,186	958,791
per square kilometer	375.4	39.5	183.2	332.9
% rural	85.7	71.6	79.5	84.4
% at school	44.7	37.0	42.4	45.2
% in school with secondary education	10.3	9.7	11.0	13.4
Total households	176,630	103,114	154,073	202,291
per square kilometer	83.6	7.9	44.3	78.7
% with high quality roof	59.7	49.2	78.5	88.2
% with high quality floor	27.7	20.6	29.7	40.0
% with high quality walls	32.2	20.3	28.0	41.7
% with piped water	6.3	6.9	14.2	30.6
Total public facilities	644	356	521	813
per capita (000s)	0.81	0.59	0.75	0.98
Electrification rates				
Rural (%)	2.3	1.5	3.1	5.3
Urban (%)	21.8	20.2	27.2	43.2
Public facilities (%)	84.1	79.9	88.1	92.6

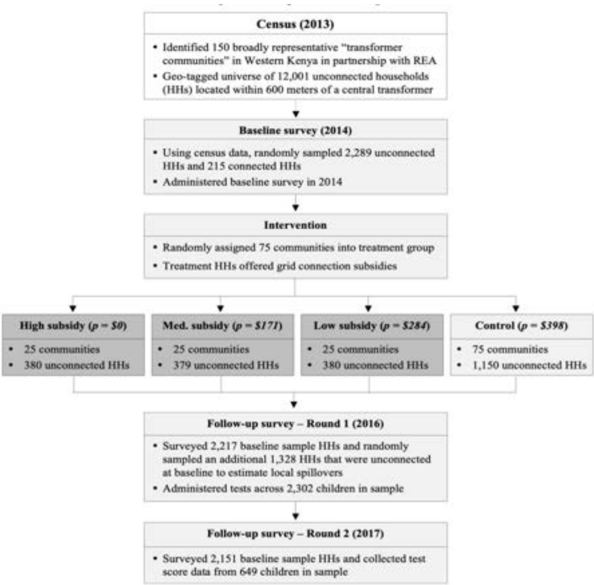
LMW (2019): Context



LMW (2019): Data

	Unconnected (1)	Connected (2)	p-value of diff. (3)
Panel A: Household head (respondent) characteristics			
Female (%)	62.9	58.6	0.22
Age (years)	52.3	55.8	< 0.01
Senior citizen (%)	27.5	32.6	0.11
Attended secondary schooling (%)	13.3	45.1	< 0.01
Married (%)	66.0	76.7	< 0.01
Not a farmer (%)	22.5	39.5	< 0.01
Employed (%)	36.1	47.0	< 0.01
Basic political awareness (%)	11.4	36.7	< 0.01
Has bank account (%)	18.3	60.9	< 0.01
Monthly earnings (USD)	16.9	50.6	< 0.01
Panel B: Household characteristics			
Number of members	5.2	5.3	0.76
Youth members (age ≤ 18)	3.0	2.6	0.01
High-quality walls (%)	16.0	80.0	< 0.01
Land (acres)	1.9	3.7	< 0.01
Distance to transformer (m)	356.5	350.9	0.58
Monthly (non-charcoal) energy (USD)	5.5	15.4	< 0.01
Panel C: Household assets			
Bednets	2.3	3.4	< 0.01
Sofa pieces	6.0	12.5	< 0.01
Chickens	7.0	14.3	< 0.01
Radios	0.35	0.62	< 0.01
Televisions	0.15	0.81	< 0.01
Sample size	2,289	215	

LMW (2019): Study design



LMW (2019): Study design

LMW (2019): Demand – Estimation

LMW randomly assigned subsidies:

- This means they can estimate the price elasticity simply

All they need to do is estimate:

$$Y_{ic} = \alpha + \tau_L D_c^L + \tau_M D_c^M + \tau_H D_c^H + \beta X_{ic} + \varepsilon_{ic}$$

where:
Y_{ic} is take-up for household i in community c
D_c^L, D_c^M, and D_c^H are treatment for the Low, Medium, and High subsidies
X_{ic} are controls
ε_{ic} is an error term

LMW (2019): Demand – Estimation

LMW randomly assigned subsidies:

- This means they can estimate the price elasticity simply

All they need to do is estimate:

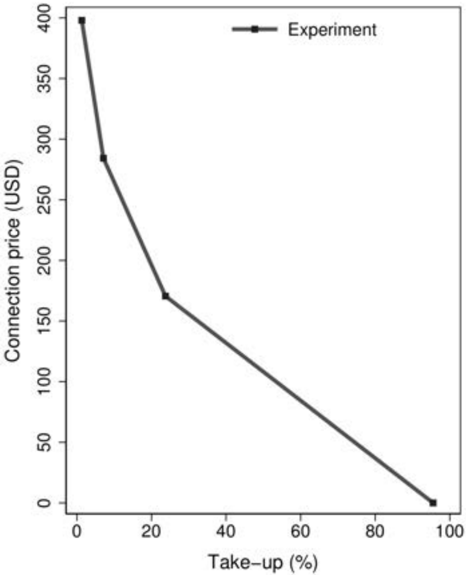
$$Y_{ic} = \alpha + \tau_L D_c^L + \tau_M D_c^M + \tau_H D_c^H + \beta X_{ic} + \varepsilon_{ic}$$

where:
Y_{ic} is take-up for household i in community c
D_c^L, D_c^M, and D_c^H are treatment for the Low, Medium, and High subsidies
X_{ic} are controls
ε_{ic} is an error term

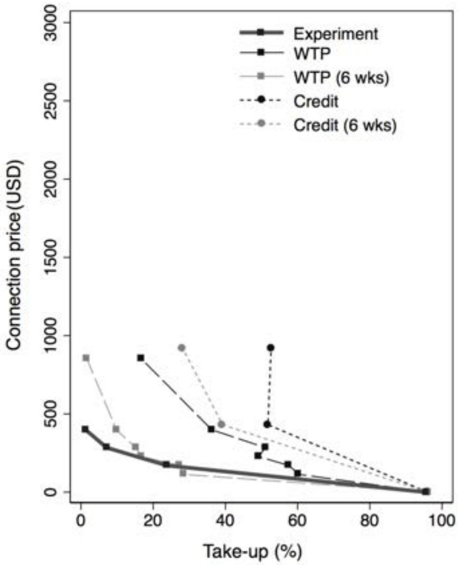
$\rightarrow$ τ_L, τ_M, τ_H are (treatment) effects on take-up at different subsidy levels

$\rightarrow$ α captures take-up in the control group

LMW (2019): Demand – Results



LMW (2019): Demand – Results



LMW (2019): Estimation – Supply

We need to know the cost of a connection:

- The average total cost (admin data) is $1,813 per connection
- Are there "economies of scale"?
- To estimate this:

$$ATC_c = \tau_0 + \tau_1 Connections_c + \tau_2 Connections_c^2 + \varepsilon_c$$

LMW (2019): Estimation – Supply

We need to know the cost of a connection:

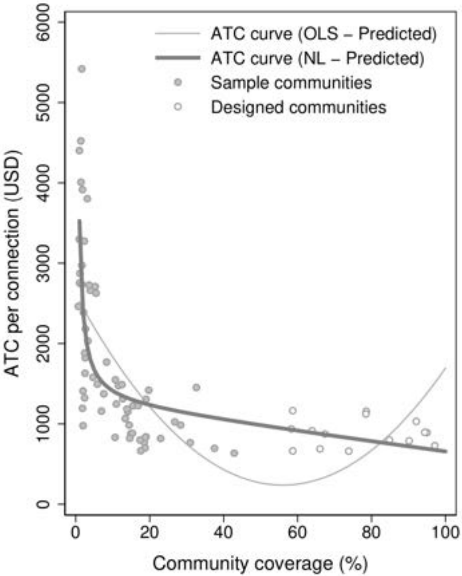
- The average total cost (admin data) is $1,813 per connection
- Are there "economies of scale"?
- To estimate this:

$$ATC_c = \tau_0 + \tau_1 Connections_c + \tau_2 Connections_c^2 + \varepsilon_c$$

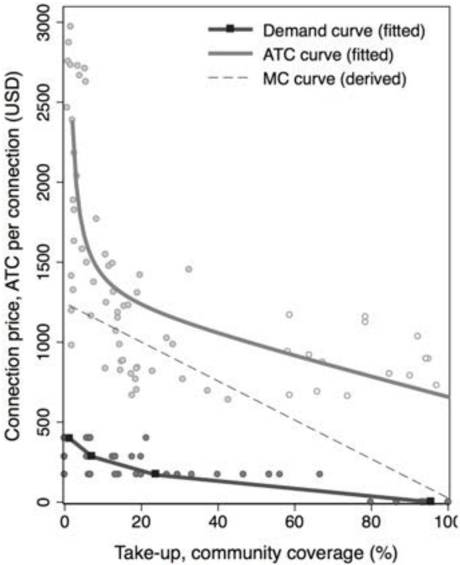
or

$$ATC_c = \underbrace{\frac{\tau_0}{Connections_c} + \tau_1}_{\text{fixed cost}} + \underbrace{\tau_2 Connections_c}_{\text{marginal cost}} + \varepsilon_c$$

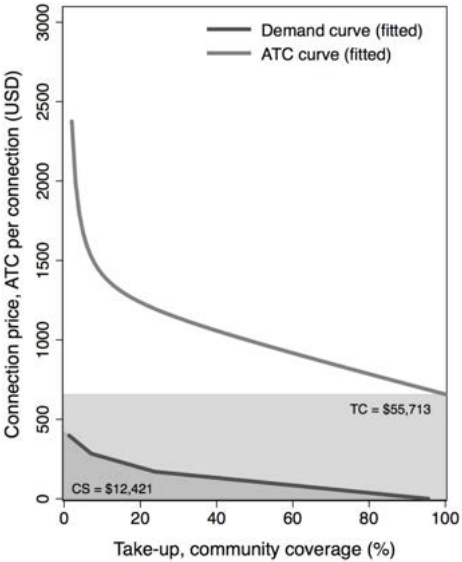
LMW (2019): Supply – Results



LMW (2019): Supply and demand – Results



LMW (2019): Supply and demand – Results



LMW (2019): Outcomes – Estimation

Following from the randomly-assigned subsidies:

- To get the ITT, they just estimate:

$$Y_{ic} = \alpha + \tau_H D_c^H + \beta X_{ic} + \varepsilon_{ic}$$

(no D_c^L or D_c^M because take-up was so low)

LMW (2019): Outcomes – Estimation

Following from the randomly-assigned subsidies:

- To get the ITT, they just estimate:

$$Y_{ic} = \alpha + \tau_H D_c^H + \beta X_{ic} + \varepsilon_{ic}$$

(no D_c^L or D_c^M because take-up was so low)

- To get the ATT, they estimate:

$$E_{ic} = \alpha + \gamma_L D_c^L + \gamma_M D_c^M + \gamma_H D_c^H + \beta X_{ic} + \varepsilon_{ic}$$

and

$$Y_{ic} = \alpha + \tau \hat{E}_{ic} + \beta X_{ic} + \varepsilon_{ic}$$

→ They instrument to get from ITT to ATT

LMW (2019): Outcomes – Results

	Control	ITT	TOT	FDR q-val
	(1)	(2)	(3)	(4)
Panel A: Treatment effects on pre-specified outcomes				
P1. Grid connected (%)	5.6	89.7***	–	–
	[23.0]	(1.4)		
P2. Monthly electricity spending (USD)	0.16	2.00***	2.20***	.001
	[1.29]	(0.18)	(0.20)	
P3. Household employed or own business (%)	36.8	5.1	4.5	.416
	[38.8]	(3.1)	(3.4)	
P4. Total hours worked last week	50.9	-2.8*	-3.6**	.167
	[32.8]	(1.5)	(1.7)	
P5. Total asset value (USD)	888	109	110	.540
	[851]	(108)	(120)	
P6. Ann. consumption of major food items (USD)	117	-3	-5	.548
	[92]	(6)	(7)	
P7. Recent health symptoms index	0	-0.03	-0.05	.548
	[1]	(0.06)	(0.07)	
P8. Normalized life satisfaction	0	0.12**	0.13*	.179
	[1]	(0.06)	(0.07)	
P9. Political and social awareness index	0	-0.03	-0.02	.731
	[1]	(0.05)	(0.05)	
P10. Average student test Z-score	0	-0.08	-0.10	.540
	[0.99]	(0.10)	(0.10)	
Panel B: Mean treatment effects on grouped outcomes				
G1. Economic Index (P3 to P6 outcomes)	0	0.06	0.03	–
	[1]	(0.08)	(0.08)	
G2. Non-Economic Index (P7 to P10 outcomes)	0	-0.01	-0.02	–
	[1]	(0.06)	(0.07)	

LMW (2019): Summary

	C	Experimental approach CS	Experimental approach NW	Alternative approach CS	Alternative approach NW	Key assumption(s)
Main estimates	658	147	-511	147	-511	
a) Income growth (*experimental approach*);	–	+139		–		Income growth of 3 percent per annum over 30 years (based on demand curves in figure 2, panel B);
Electricity consumption growth (*alternative approach*)	–	–		+182		Electricity consumption growth of 10 percent per annum over 30 years (see table 4, column 2, row 3).
b) No credit constraints for grid connections	–	+301		–		Stated WTP without time constraints (see figure 5)
c) No transformer breakdowns	–	+33		+19		Reduce likelihood of transformer breakdowns from 5.4 to 0 percent (see appendix table B10).
d) No grid connection delays	–	+46		+26		Reduce waiting period from 188 to 0 days (see appendix figure A1).
e) No construction cost leakage	-140	–		–		Decrease total construction costs by 21.3 percent (see appendix table B8).
Ideal scenario	518	665	148	374	-144	

Third paper: Burlig and Preonas (2016)

What is the impact of rural electrification on economic development?

Context: Massive rural electrification program in India

- Home to world's largest unelectrified population
- Program targeted $> 400{,}000$ villages ($\approx 2/3$)

Research design: Regression discontinuity

- Population-based eligibility cutoff

Outcomes: Rich administrative data on development indicators

- Results from 3-5 years into the program

BP (2016): Context

"Rajiv Gandhi Grameen Vidyutikaran Yojana" electrification program

- Enacted in 2005; Goal: bring electricity access to all rural villages
- Approx. $17.2 billion in federal funds budgeted for the program
- Covered over 400,000 villages in 27 states
  - > 100,000 unelectrified villages
  - > 300,000 "under-electrified" villages

Empirical strategy: regression discontinuity

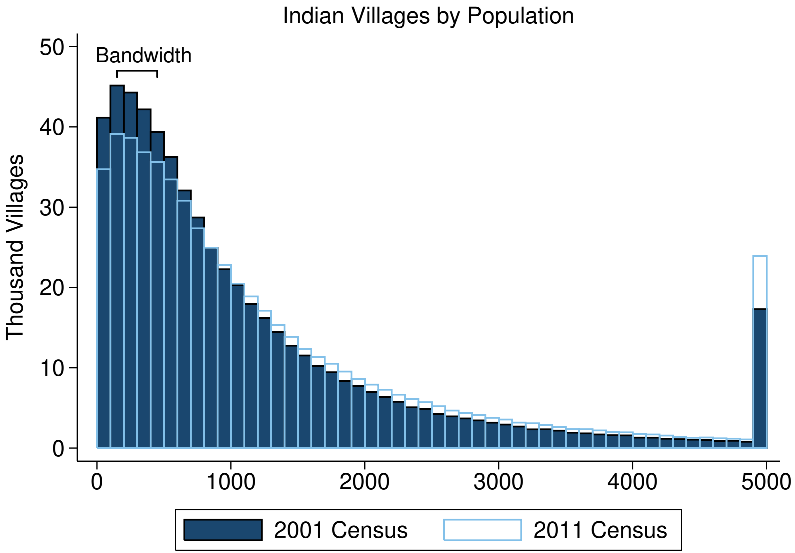
We use RGGVY's first wave only

- 225 districts across 25 states
- Earliest wave of program $\rightarrow$ more years of data
- Village eligibility cutoff: **300 people**

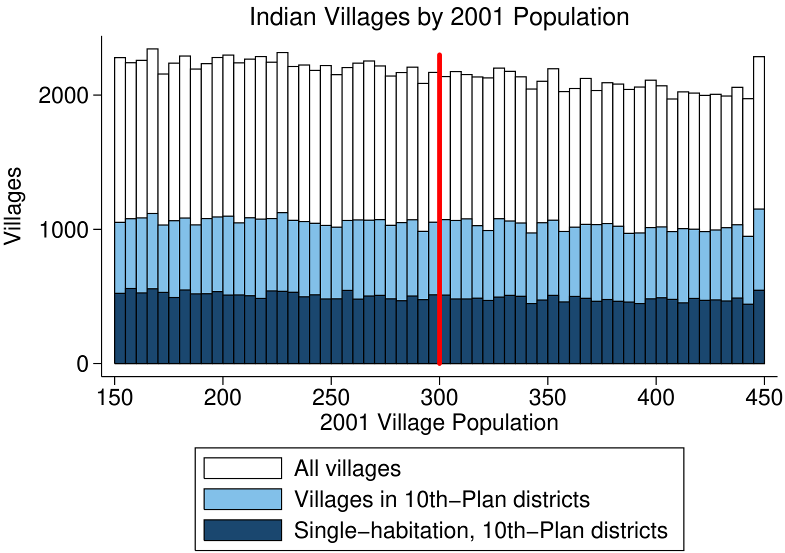
Regression discontinuity design

- Population-based eligibility cutoff (running variable set in 2001)
- We estimate an intent-to-treat effect
- Identifying assumptions:
  - continuity across threshold
  - population not manipulable

Population is smooth across the threshold



Population is smooth across the threshold



BP (2016): Estimation

$$Y_{vs}^{2011} = \alpha + \tau D_{vs} + \beta_1(P_{vs} - 300) + \beta_2(P_{vs} - 300) \cdot D_{vs} + \ Y_{vs}^{2001} + \eta_s + \varepsilon_{vs}$$

for $300 - h \leq P_{vs} \leq 300 + h\,,$

where $w_{vs} \equiv \mathbf{1}[P_{vs} \geq 300]$

Y_{vs}^t : outcome variable for village v in year t

P_{vs} : population of village v in 2001

D_{vs} : RD eligibility indicator for village v

η_s : state fixed effects

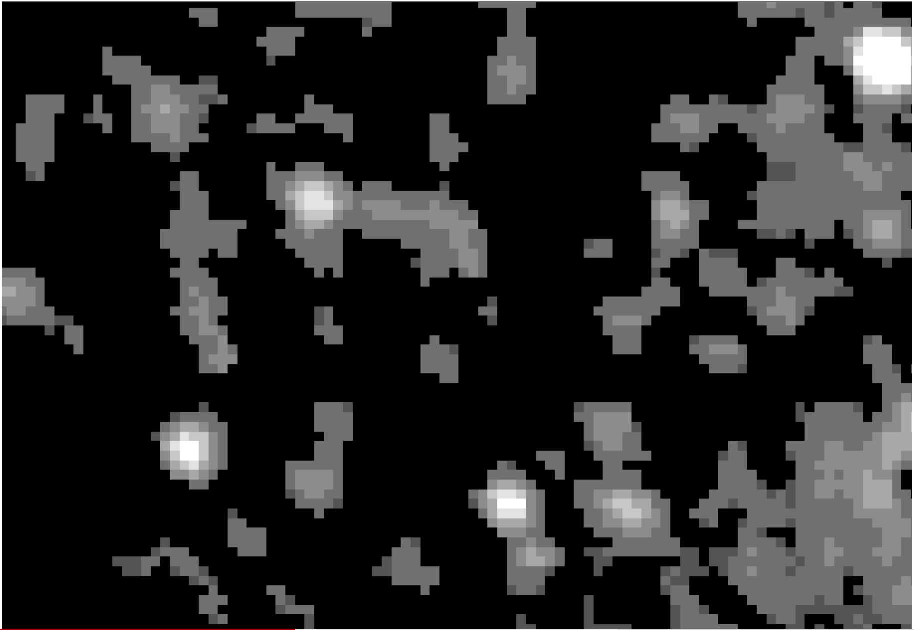
ε_{vs} : error term (clustered by district)

h : RD bandwidth ($h = 150$)

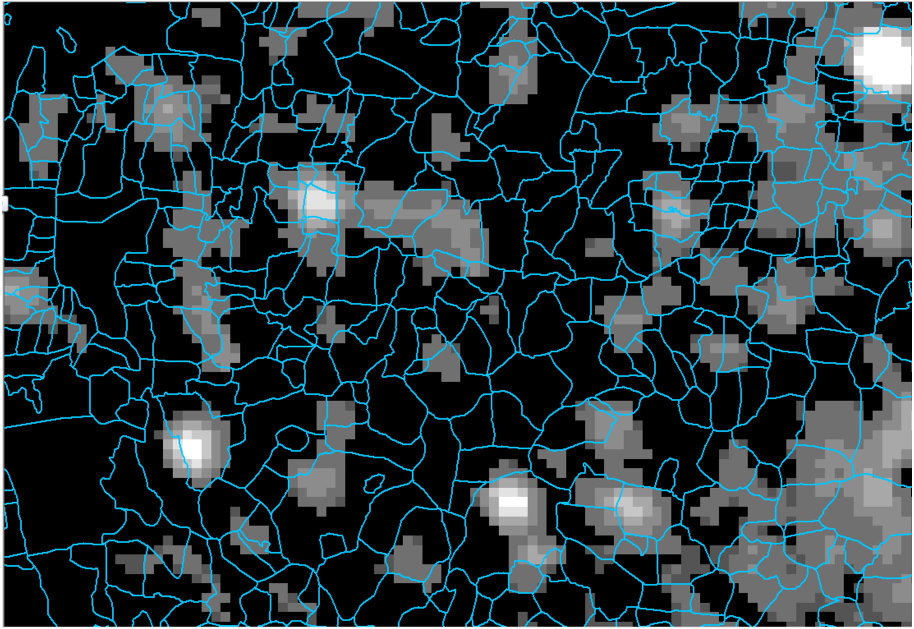
Nighttime lights (2001)

Nighttime lights (2011)

Zoomed in



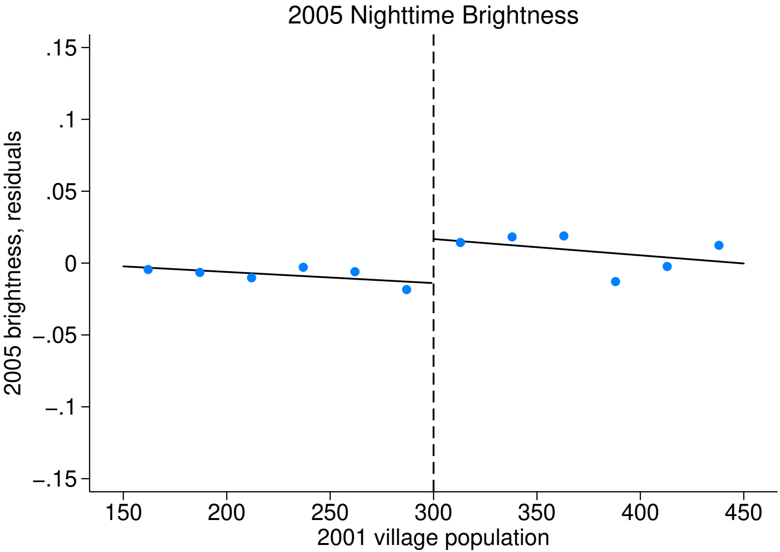
Village-level brightness



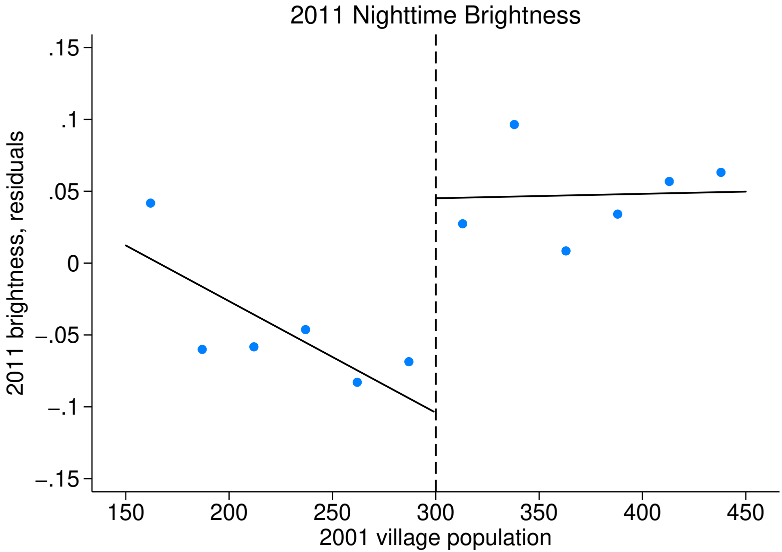
Outcomes come from several large administrative datasets

Dataset	Information
Primary Census Abstract (2001, 2011)	- population (running variable) - # workers, by gender/type
Houselisting Primary Census Abstract (2011)	- asset ownership - housing characteristics
Village Directory (2001, 2011)	- village-level amenities
Socioeconomic and Caste Census (2001, 2011)	- poverty - household wealth
District Info. System on Education (2005-06 – 2014-15)	- school enrollment

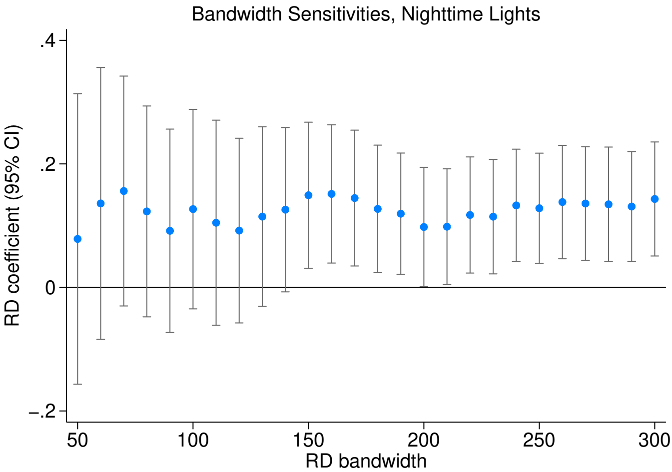
RD on nighttime brightness – pre-program



RD on nighttime brightness – post-program



Nighttime brightness bandwidth sensitivity



Imbens and Kalyanaraman (2012) optimal bandwidth: 137–174

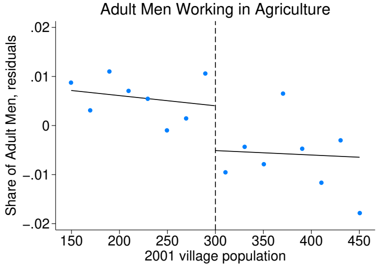
Results withstand a variety of robustness checks

Our results are robust to:

- A placebo test
- A randomization inference check
- Alternative bandwidths
- Alternative functional forms
- Alternative measures of brightness
- Alternative standard errors
- Inclusion of controls
- Three falsification tests

RD results: agricultural employment

$\tau = -0.0071^{**}$ CI: [-0.013, -0.002]

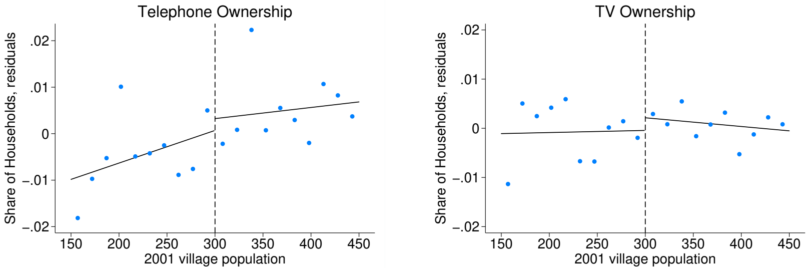


$\tau = -0.0091^{*}$ CI: [-0.019, 0.001]

$\tau = -0.0049$ CI: [-0.013, 0.003]

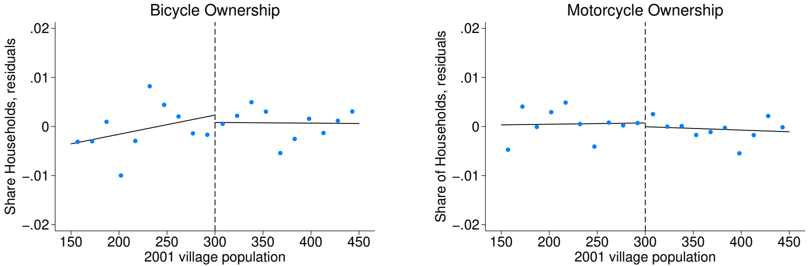
$\tau = -0.0039$ CI: [-0.013, 0.006]

RD results: asset ownership



$\tau = 0.0025$ CI: [-0.008, 0.013]

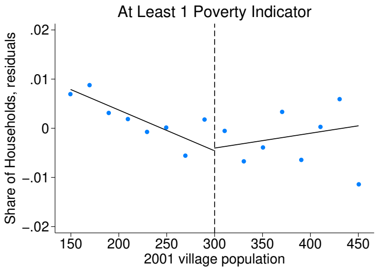
$\tau = 0.0026$ CI: [-0.005, 0.010]



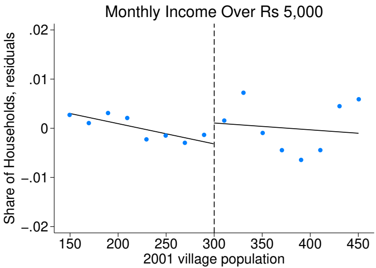
$\tau = -0.0015$ CI: [-0.010, 0.007]

$\tau = -0.008$ CI: [-0.006, 0.004]

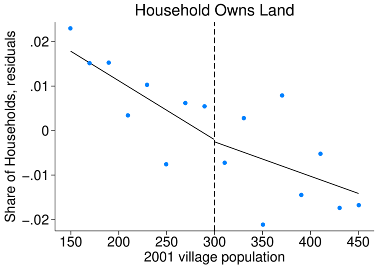
RD results: household wealth



$\tau = 0.0006$ CI: [-0.011, 0.012]



$\tau = 0.0043$ CI: [-0.004, 0.013]



$\tau = -0.0005$ CI: [-0.017, 0.016]

$\tau = 0.0030$ CI: [-0.002, 0.008]

Scaling our results

- Convert from ITT to LATE

  $\approx$ 56–82% of RGGVY-eligible villages received treatment

  $\rightarrow$ Scaling factor $\approx$ **1.5**

Scaling our results

- Convert from ITT to LATE

  $\approx$ 56–82% of RGGVY-eligible villages received treatment

  $\rightarrow$ Scaling factor $\approx$ **1.5**

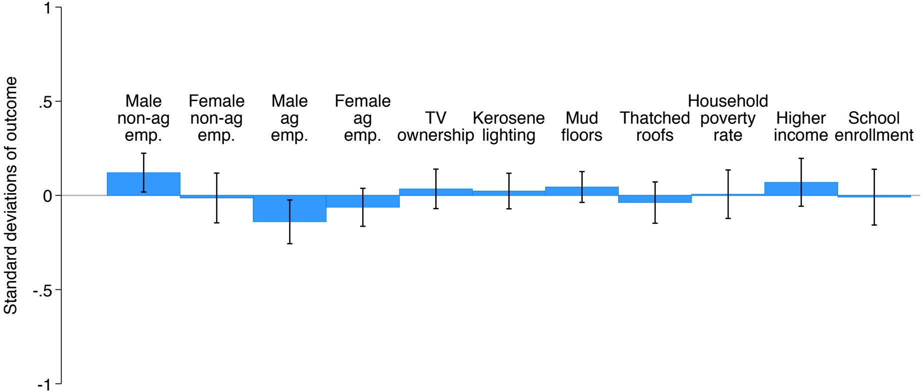
- Calibrate $\tau = 0.15$ to remote sensing estimates

  – village-wide electrification $\approx$ 0.4-unit increase

  – per-household conversion $\approx$ 0.2-unit increase

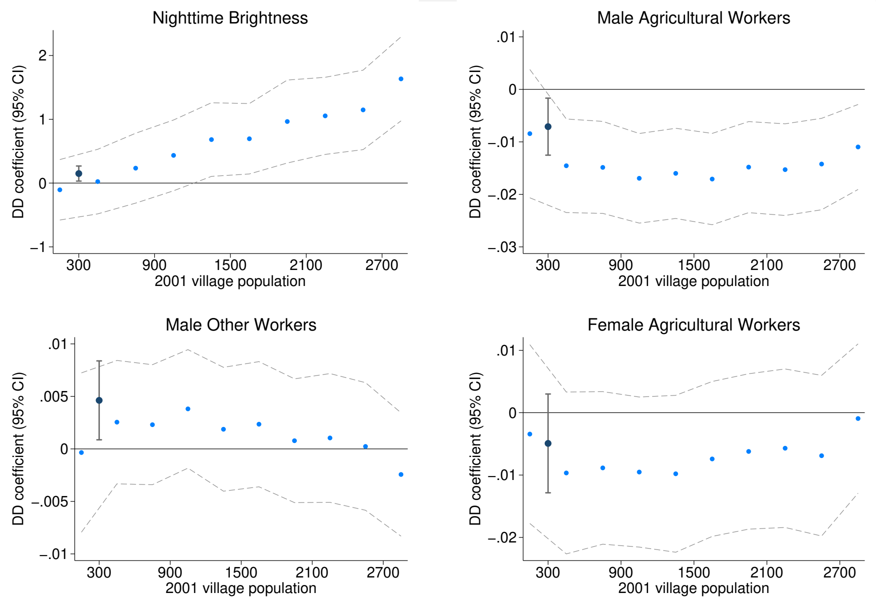
  $\rightarrow$ Scaling factor $\approx$ **1.3 to 3**

Applying a scaling factor of 3



We can reject changes $\geq$ 0.26 of one standard deviation, for all development outcomes

Results: going beyond LATE



Welfare

Recap

TL;DR:

1. Lee, Miguel, Wolfram (2019) randomizes subsidies for household electricity connections
2. Burlig and Preonas (2016) use an RD to study electrification
3. Both find extremely limited benefits