Lecture 17: Policy Lab Does rural electrification work? I

PPHA 34600
Prof. Fiona Burlig

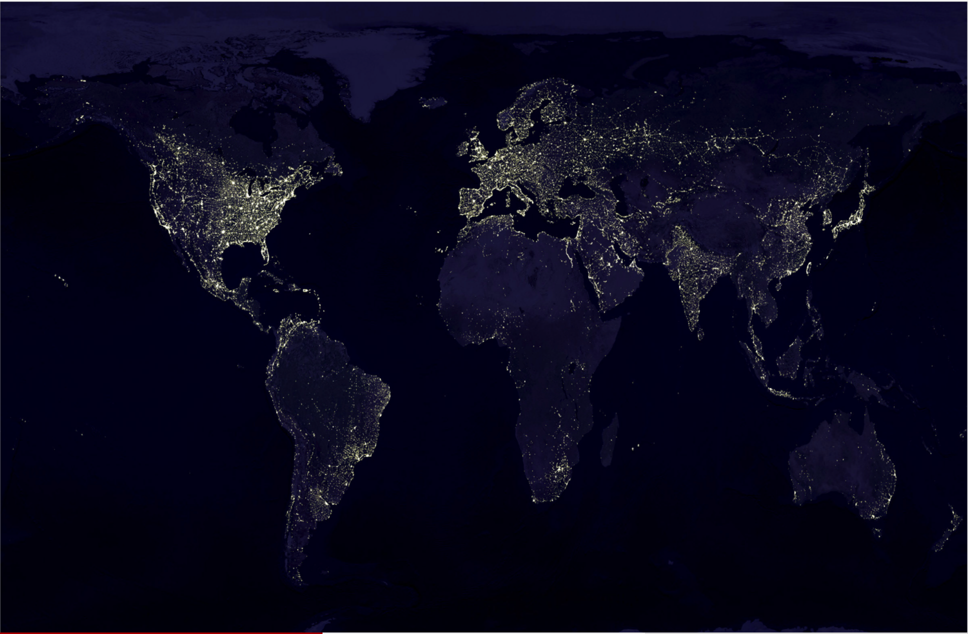
Harris School of Public Policy
University of Chicago

From last time: Machine learning for causal inference

We looked at several ways to incorporate ML into CI

1. Generating (big) data
2. Exploring heterogeneity
3. Improving research designs
   - ML works with SOO to handle functional form
   - And with SOU to aid in generating counterfactuals

The world at night



Electrification is a major policy goal

1.1 billion people still lack modern electricity access

Electrification is a major policy goal

1.1 billion people still lack modern electricity access

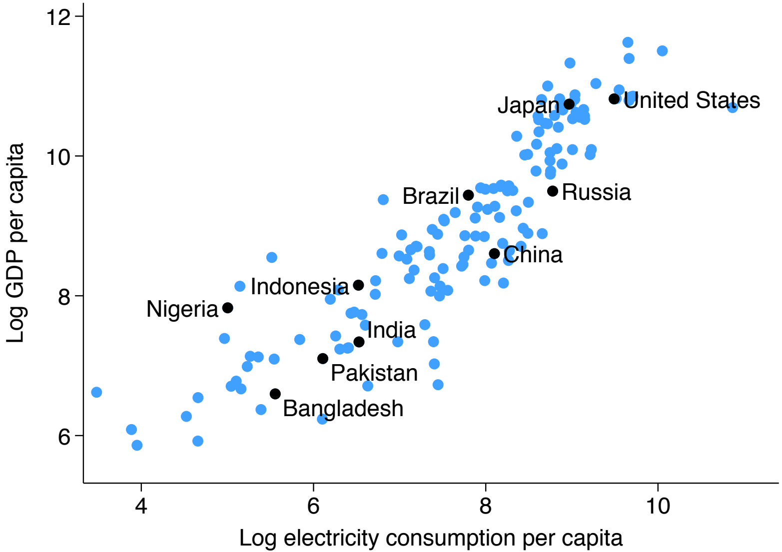
"Access to energy is essential to reduce poverty."

— The World Bank

"By developing infrastructure that provides sustainable, reliable and affordable access to modern energy services, people, communities and countries can significantly improve their living standards and economic status."

— United Nations Development Programme

Wealthy countries use more electricity per capita



Source: World Bank tables

Research question

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

Research question

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

Research question

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?**

Research question

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

Research question

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!

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to electricity
- ... but even this is not straightforward

What do we mean by "rural electrification"?

What do we mean by "rural electrification"?

What do we mean by “rural electrification”?

What do we mean by "rural electrification"?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to electricity
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign 24x7 power?
    - Or should actual power supply reflect existing conditions?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to electricity
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign 24x7 power?
  - Or should actual power supply reflect existing conditions?
- Do we care about effects on households? Villages? Towns? Counties?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to electricity
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign 24x7 power?
  - Or should actual power supply reflect existing conditions?
- Do we care about effects on households? Villages? Towns? Counties?
- Do we think about electrification in one location at a time?
  - Or do we consider general equilibrium?

The ideal experiment

What would we do if we could do anything?

- Some kind of random assignment to electricity
- ... but even this is not straightforward

Important considerations:

- Do we want to randomly assign 24x7 power?
  - Or should actual power supply reflect existing conditions?
- Do we care about effects on households? Villages? Towns? Counties?
- Do we think about electrification in one location at a time?
  - Or do we consider general equilibrium?

→ Actually ideal experiment probably requires multiple Earths

→ Or at least a really large sample!

If you thought the *ideal* experiment was tricky...

An additional practical wrinkle:

- Randomizing electricity access is impractical!
  - Like roads, electricity works on a network
  - Putting a random segment in the middle of nowhere would not work
  - Simple randomization is therefore not really going to work
  - Not to mention that it's going to be extremely hard to randomize at any meaningful scale

If you thought the *ideal* experiment was tricky...

An additional practical wrinkle:

- Randomizing electricity access is impractical!
  - Like roads, electricity works on a network
  - Putting a random segment in the middle of nowhere would not work
  - Simple randomization is therefore not really going to work
  - Not to mention that it's going to be extremely hard to randomize at any meaningful scale

$\rightarrow$ A quasi-experimental approach may be useful here

First paper: Dinkelman (2011)

This is the most prominent early econ paper estimating effects of RE:

- **Research question:** What is the effect of electrification on "the ability of the poor to use their labor resources for market production?"

→ AKA, what is the effect of RE on employment?

What do we expect to happen?

There are lots of potential effects here:

- The nature of work at home may change
- So can work outside of the home

What do we expect to happen?

There are lots of potential effects here:

- The nature of work at home may change
- So can work outside of the home

The expected sign is ambiguous:

- New technology $\rightarrow$ higher productivity in home activity...
- ...but also more time in the day $\rightarrow$ more time spent on market

$\rightarrow$ Likely expect the market effects to dominate

What do we expect to happen?

There are lots of potential effects here:

- The nature of work at home may change
- So can work outside of the home

The expected sign is ambiguous:

- New technology $\rightarrow$ higher productivity in home activity...
- ...but also more time in the day $\rightarrow$ more time spent on market

$\rightarrow$ Likely expect the market effects to dominate

- Can also lead to firm growth
- Or more jobs in the household

What do we expect to happen?

There are lots of potential effects here:

- The nature of work at home may change
- So can work outside of the home

The expected sign is ambiguous:

- New technology $\rightarrow$ higher productivity in home activity...
- ...but also more time in the day $\rightarrow$ more time spent on market

$\rightarrow$ Likely expect the market effects to dominate

- Can also lead to firm growth
- Or more jobs in the household
- Both in- and out-migration effects

Dinkelman (2011): Context

Dinkelman studies immediately-post-Apartheid South Africa:

- **1993:** more than 2/3 of households didn't have electricity
  - 80% used wood burning for home production

Dinkelman (2011): Context

Dinkelman studies immediately-post-Apartheid South Africa:

- **1993:** more than 2/3 of households didn't have electricity
  - 80% used wood burning for home production
- **1994:** Elections $\rightarrow$ new Black-led government ends Apartheid
  - ANC commits to universal electrification

Dinkelman (2011): Context

Dinkelman studies immediately-post-Apartheid South Africa:

- **1993:** more than 2/3 of households didn't have electricity
  - 80% used wood burning for home production
- **1994:** Elections → new Black-led government ends Apartheid
  - ANC commits to universal electrification
- **2001:** 1/4 of households were newly connected to the grid
  - 2x as many households electrified as the first five years of the REA (US)

Dinkelman (2011): Context

Dinkelman studies immediately-post-Apartheid South Africa:

- **1993:** more than 2/3 of households didn't have electricity
  - 80% used wood burning for home production
- **1994:** Elections → new Black-led government ends Apartheid
  - ANC commits to universal electrification
- **2001:** 1/4 of households were newly connected to the grid
  - 2x as many households electrified as the first five years of the REA (US)

→ Massive push towards rural electrification

→ Opportunity for a natural experiment!

Dinkelman (2011): Context

What did the electrification program look like?

- National Electrification Programme (NEP): electrify 300k households
- Implemented through Eskom, the monopoly utility
- Cost $1.4bn; connected 470,000 hhs in KZN (1993–2003)
- Connections powerful enough to run a few (small) appliances
- Places that were more expensive to electrify got power later

Dinkelman (2011): Context

Post-Apartheid South Africa is an interesting place to study RE:

- A big switch from no power to power
- Very poor, marginalized groups getting access
- At the same time, many people getting access
- Important evidence from Sub-Saharan Africa

Dinkelman (2011): Context

Post-Apartheid South Africa is an interesting place to study RE:

- A big switch from no power to power
- Very poor, marginalized groups getting access
- At the same time, many people getting access
- Important evidence from Sub-Saharan Africa
- Also a unique setting
- Nothing has ever really looked just like this context

→ You should ask about external validity

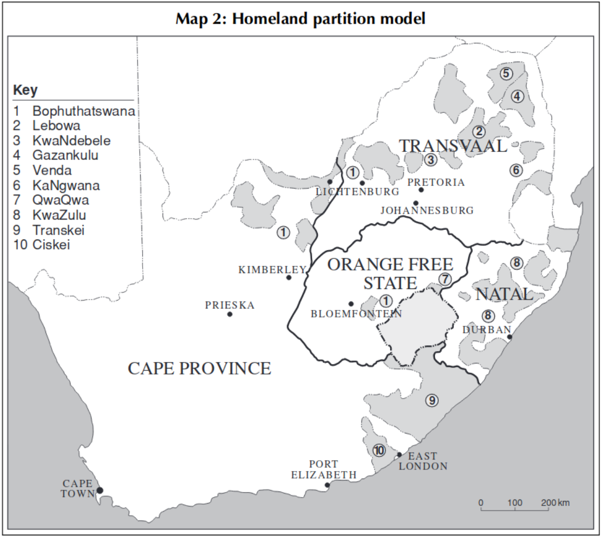
Dinkelman (2011): Data

	Means (standard deviation)			Differences in means (standard error)		
					By gradient	
Covariates in 1996	Full sample (1)	Eskom project (2)	No project (3)	Columns 2–3 (4)	No controls (5)	Controls (6)
Poverty rate	0.61	0.59	0.61	−0.024**	0.00	0.002
	(0.19)	(0.17)	(0.20)	(0.01)	(0.00)	(0.00)
Female-headed HHs	0.55	0.55	0.55	0.00	0.005***	0.001
	(0.13)	(0.12)	(0.13)	(0.01)	(0.00)	(0.00)
Adult sex ratio ($N_{females}/N_{males}$)	1.48	1.41	1.49	−0.080***	0.011***	0.004**
	(0.28)	(0.25)	(0.29)	(0.02)	(0.00)	(0.00)
Indian, white adults × 10	0.00	0.00	0.00	0.00	0.000	0.000
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Kilometers to road	37.95	35.62	38.54	−2.917**	−0.201	−0.156
	(24.57)	(24.18)	(24.64)	(1.44)	(0.41)	(0.18)
Kilometers to town	38.57	36.34	39.13	−2.790***	0.278	0.180
	(18.12)	(15.34)	(18.72)	(1.06)	(0.41)	(0.13)
Men with high school	0.06	0.08	0.06	0.016***	−0.002***	−0.003**
	(0.05)	(0.05)	(0.05)	(0.00)	(0.000)	(0.00)
Women with high school	0.07	0.08	0.06	0.020***	−0.002***	0.000
	(0.05)	(0.05)	(0.05)	(0.00)	(0.000)	(0.00)
Household density	22.05	32.56	19.41	13.152***	−0.523*	−0.944***
	(30.48)	(49.31)	(22.75)	(1.76)	(0.31)	(0.30)
Kilometers from grid	19.06	15.75	19.89	−4.139***	−0.235	0.029
	(13.32)	(10.20)	(13.88)	(0.77)	(0.36)	(0.12)
Land gradient	10.10	9.12	10.35	−1.232***		
	(4.89)	(4.21)	(5.02)	(0.29)		
N communities	1,816	365	1,451	1,816	1,816	1,816

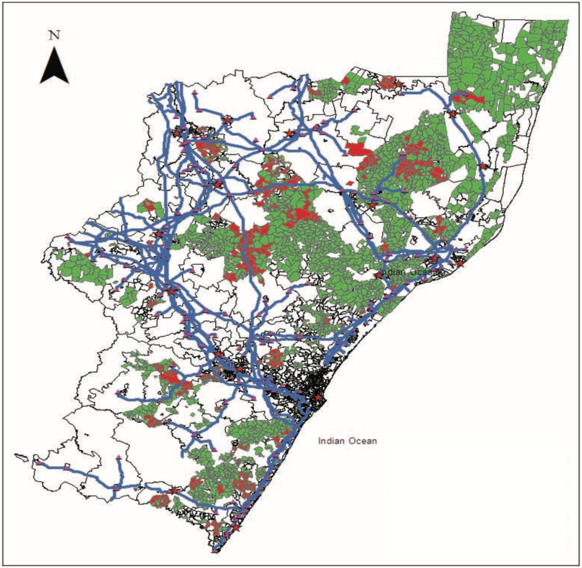
Dinkelman (2011): Data

	Year	Means (standard deviation)			Difference: Column 2–3
		Full sample (1)	Eskom project (2)	No project (3)	(4)
Female employment rate	1996	0.07	0.09	0.06	0.021***
		(0.08)	(0.07)	(0.08)	(0.00)
	2001	0.07	0.08	0.06	0.017***
		(0.07)	(0.07)	(0.07)	(0.00)
Difference	Δ_t	0.000	−0.003	0.001	−0.004
		(0.002)	(0.005)	(0.00)	(0.00)
Male employment rate	1996	0.14	0.16	0.13	0.031***
		(0.11)	(0.11)	(0.11)	(0.01)
	2001	0.10	0.11	0.10	0.014**
		(0.09)	(0.09)	(0.09)	(0.01)
Difference	Δ_t	−0.04***	−0.050***	−0.033***	−0.017***
		(0.00)	(0.01)	(0.00)	(0.01)
N		1,816	365	1,451	

Dinkelman (2011): Context



Dinkelman (2011): Data



Dinkelman (2011): Estimation approach

Electrification was not randomly assigned in South Africa:

- We need a research design to estimate the causal effect of interest

Ideally, we'd estimate:

$$Y_{id} = \alpha + \tau D_{id} + \varepsilon_{id}$$

where:

Y_{id} is the outcome (female employment rate) in community i in district d
D_{id} is an electrification indicator
ε_{id} is an error term

→ Without random assignment, we will get bias (why?)

Dinkelman (2011): Estimation approach

Without random assignment, we could leverage time:

$$Y_{idt} = \tau D_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

→ Note that Dinkelman writes this a bit weirdly

$$y_{jdt} = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \rho_d + \lambda_d t + \epsilon_{jdt},$$

Dinkelman (2011): Estimation approach

Without random assignment, we could leverage time:

$$Y_{idt} = \tau D_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

→ Note that Dinkelman writes this a bit weirdly

$$y_{jdt} = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \rho_d + \lambda_d t + \epsilon_{jdt},$$

- Why a time trend, not time FE?
- You can't have a district and a community FE?

Dinkelman (2011): Estimation approach

Without random assignment, we could leverage time:

$$Y_{idt} = \tau D_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

→ Note that Dinkelman writes this a bit weirdly

$$y_{jdt} = \alpha_0 + \alpha_1 t + \alpha_2 T_{jdt} + \mu_j + \delta_j t + \rho_d + \lambda_d t + \epsilon_{jdt},$$

- Why a time trend, not time FE?
- You can't have a district and a community FE?

→ Even with time, we still have identification concerns (why?)

Dinkelman (2011): Estimation approach

Dinkelman uses an IV approach to overcome the selection problem:

- We want to isolate the effect of electrification from everything else

For the instrument to be valid, we need:

1. **First stage:** Our IV needs to be correlated with electrification
2. **Exclusion restriction:** Our IV needs to only move Y through electrification

Dinkelman (2011): Estimation approach

Dinkelman uses an IV approach to overcome the selection problem:

- We want to isolate the effect of electrification from everything else

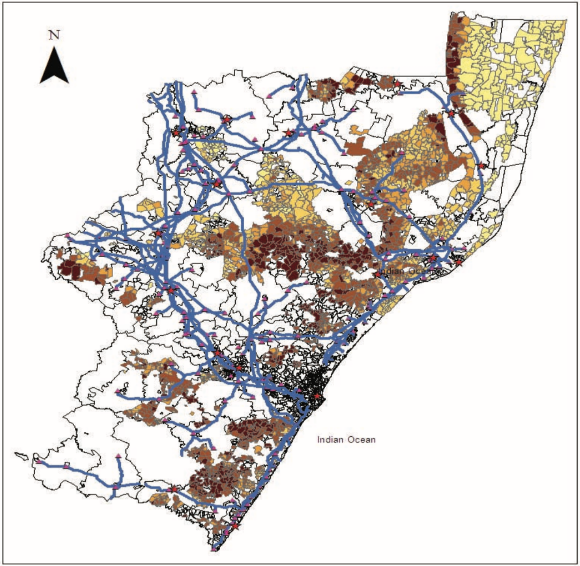
For the instrument to be valid, we need:

1. **First stage:** Our IV needs to be correlated with electrification
2. **Exclusion restriction:** Our IV needs to only move Y through electrification

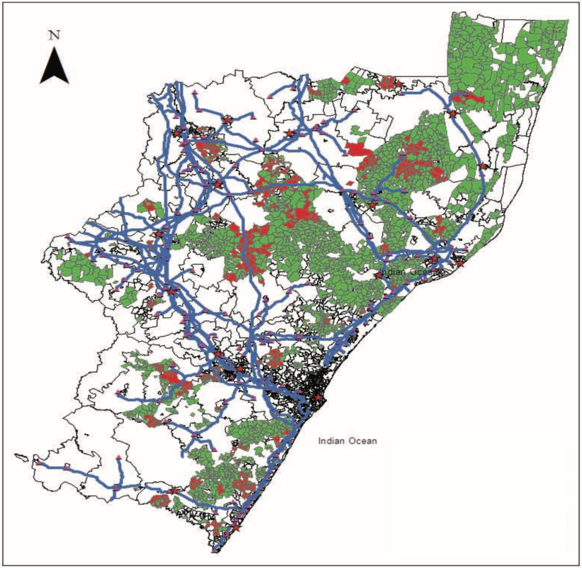
Instrument of choice: land gradient

- Steeper land is more expensive to electrify
- The first stage should be negative

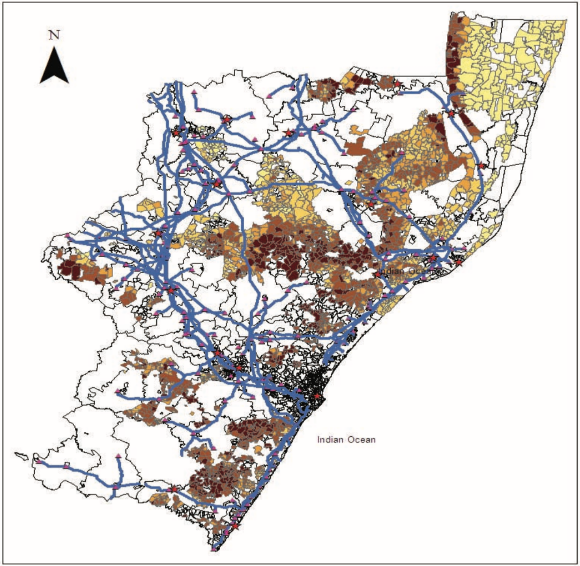
Dinkelman (2011): Estimation approach



Dinkelman (2011): Estimation approach



Dinkelman (2011): Estimation approach



Dinkelman (2011): Estimation approach

With the instrument, we simply estimate:

$$D_{idt} = \theta Z_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

$$Y_{idt} = \tau \hat{D}_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

Dinkelman (2011): Estimation approach

With the instrument, we simply estimate:

$$D_{idt} = \theta Z_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

$$Y_{idt} = \tau \hat{D}_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

Identifying assumption: Conditional on fixed effects, land gradient does not affect employment growth other than through electricity

- Is this reasonable?

Dinkelman (2011): Estimation approach

With the instrument, we simply estimate:

$$D_{idt} = \theta Z_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

$$Y_{idt} = \tau \hat{D}_{idt} + \alpha_{id} + \delta_t + \varepsilon_{idt}$$

Identifying assumption: Conditional on fixed effects, land gradient does not affect employment growth other than through electricity

- Is this reasonable?

LATEs: Estimates are the LATE for relatively flatter places

- How should this compare to ATE?

Dinkelman (2011): First stage

Dependent variable: Eskom project = [1 or 0]	(1)	(2)	(3)	(4)
Gradient × 10	−0.083**	−0.075**	−0.078***	−0.077***
	(0.040)	(0.034)	(0.027)	(0.027)
Kilometers to grid × 10		−0.040*	−0.012	−0.011
		(0.021)	(0.023)	(0.023)
Household density × 10		0.017***	0.012**	0.013**
		(0.004)	(0.006)	(0.006)
Poverty rate		0.023	0.019	0.017
		(0.069)	(0.070)	(0.069)
Female-headed HHs		0.393***	0.165	0.155
		(0.120)	(0.107)	(0.107)
Adult sex ratio		−0.173***	−0.130***	−0.121***
		(0.052)	(0.042)	(0.042)
Indian, white adults × 10		−1.236***	−1.116**	−1.105**
		(0.401)	(0.459)	(0.452)
Kilometers to road × 10		0.003	−0.010	−0.010
		(0.009)	(0.010)	(0.010)
Kilometers to town × 10		0.016	0.008	0.008
		(0.015)	(0.015)	(0.016)
Men with high school		−0.269	−0.185	−0.152
		(0.500)	(0.411)	(0.417)
Women with high school		1.046**	0.965**	0.984**
		(0.475)	(0.413)	(0.409)
Δ_t water access				0.012
				(0.048)
Δ_t toilet access				0.155
				(0.104)
District fixed effects	N	N	Y	Y
Mean of outcome variable	0.20	0.20	0.20	0.20
N communities	1,816	1,816	1,816	1,816
R^2	0.01	0.07	0.18	0.18
F-statistic on gradient	4.20	4.87	8.34	8.26
Pr > F	0.04	0.03	0.00	0.00

Dinkelman (2011): Household Behavior

Outcome is Δ_t in:	OLS No controls (1)	OLS Controls (2)	IV No controls (3)	IV Controls (4)
(1) Lighting with electricity	0.251***	0.221***	0.577***	0.635***
Mean: 0.08	(0.032)	(0.031)	(0.188)	(0.227)
(2) Cooking with wood	−0.045***	−0.039***	−0.266	−0.275*
Mean: −0.035	(0.012)	(0.012)	(0.179)	(0.147)
(3) Cooking with electricity	0.068***	0.056***	0.250**	0.228**
Mean: 0.037	(0.009)	(0.009)	(0.107)	(0.101)
(4) Water nearby	−0.029	0.005	−0.483*	−0.372
Mean: 0.007	(0.029)	(0.024)	(0.249)	(0.248)
(5) Flush toilet	0.003	0.008	0.018	0.067
Mean: 0.03	(0.006)	(0.005)	(0.069)	(0.068)

Dinkelman (2011): Female Employment

	Δ_t female employment rate							
	OLS regression coefficients				IV regression coefficients			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eskom project	−0.004	−0.001	0.000	−0.001	0.025	0.074	0.090*	0.095*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.045)	(0.060)	(0.055)	(0.055)
A. R. 95 percent C.I.							*[0.05; 0.3]*	*[0.05; 0.3]*
Poverty rate		0.029***	0.033***	0.031***		0.027**	0.032**	0.031**
		(0.011)	(0.010)	(0.010)		(0.012)	(0.013)	(0.013)
Female-headed HHs		0.042**	0.051***	0.047**		0.014	0.036	0.033
		(0.019)	(0.019)	(0.020)		(0.031)	(0.026)	(0.026)
Adult sex ratio		0.019**	0.017**	0.020***		0.033**	0.029**	0.032***
		(0.009)	(0.008)	(0.007)		(0.014)	(0.012)	(0.012)
Baseline controls?	N	Y	Y	Y	N	Y	Y	Y
District fixed effects?	N	N	Y	Y	N	N	Y	Y
Δ_t other services?	N	N	N	Y	N	N	N	Y
N communities	1,816	1,816	1,816	1,816	1,816	1,816	1,816	1,816

Dinkelman (2011): Male Employment

	Δ_t male employment rate							
	OLS regression coefficients				IV regression coefficients			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eskom project	−0.017**	−0.015***	−0.009	−0.010*	−0.063	0.069	0.033	0.035
	(0.007)	(0.006)	(0.006)	(0.006)	(0.073)	(0.082)	(0.064)	(0.066)
A. R. 95 percent C.I.							*[−0.05; 0.25]*	*[−0.05; 0.25]*
Poverty rate		0.062***	0.064***	0.063***		0.059***	0.064***	0.062***
		(0.020)	(0.018)	(0.018)		(0.022)	(0.019)	(0.019)
Female-headed HHs		0.217***	0.233***	0.227***		0.187***	0.227***	0.220***
		(0.029)	(0.030)	(0.030)		(0.042)	(0.034)	(0.034)
Adult sex ratio		0.018*	0.012	0.017		0.034*	0.018	0.023
		(0.011)	(0.011)	(0.011)		(0.019)	(0.015)	(0.015)
Baseline controls?	N	Y	Y	Y	N	Y	Y	Y
District fixed effects?	N	N	Y	Y	N	N	Y	Y
Δ_t other services?	N	N	N	Y	N	N	N	Y
N communities	1,816	1,816	1,816	1,816	1,816	1,816	1,816	1,816

Dinkelman (2011): Panel Results

	Females		Males		Females		Males	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A. Employment [1/0]				*Panel B. Usual weekly hours of work*			
MD electrification rate	0.126**	0.128	0.090	0.134	6.646***	8.920	5.671**	13.090
	(0.058)	(0.149)	(0.077)	(0.164)	(1.771)	(6.634)	(2.597)	(12.947)
Trend (1995–2001)	−0.010	0.046**	−0.051***	−0.075***	−0.407	−0.588	−0.322	−1.424
	(0.012)	(0.020)	(0.012)	(0.022)	(0.491)	(0.872)	(0.620)	(1.701)
N	152	152	152	152	151	151	151	151
Mean of outcome	0.25	0.25	0.42	0.42	42.82	42.82	46.94	46.94
R^2	0.06	0.63	0.09	0.76	0.06	0.42	0.03	0.45
	Panel C. Log hourly wage				*Panel D. Log monthly earnings*			
MD electrification rate	−0.148	−1.380	0.101	0.171	−0.070	−0.616	0.414**	1.107**
	(0.253)	(1.046)	(0.211)	(0.483)	(0.225)	(0.995)	(0.191)	(0.477)
Trend (1995–2001)	−0.079***	0.132	−0.027	0.077	−0.091**	−0.065	−0.047	−0.085
	(0.030)	(0.137)	(0.032)	(0.063)	(0.037)	(0.131)	(0.033)	(0.063)
N	146	146	148	148	146	146	148	148
Mean of outcome	1.17	1.17	1.49	1.49	6.42	6.42	6.80	6.80
R^2	0.03	0.52	0.00	0.51	0.03	0.52	0.05	0.57

A slight digression: multiple testing

The “reproducibility crisis” is becoming a Thing:

- We have lots of results that don’t replicate
- When we try to re-do experiments, we don’t find the same results
- Famous example: power poses!

A slight digression: multiple testing

The "reproducibility crisis" is becoming a Thing:

- We have lots of results that don't replicate
- When we try to re-do experiments, we don't find the same results
- Famous example: power poses!

→ A central culprit: **p-hacking**

  - Researchers get jobs based on statistically significant effects
  - This generates incentives to find them
  - We get a lot of results with $0.051 < p < 0.043$
  - ... a lot more than 20%!
  - Adding controls, etc to get "stars" is common

A slight digression: multiple testing

One way to generate "stars" is to run a lot of regressions:

- With a 95% confidence threshold, you will find stars 5% of the time, even if your null hypothesis is true
- So if you run a lot of regressions, you are bound to find some with stars
- If you then only report the ones with stars, we have a problem
- (If you report all of them, we're fine!)

A slight digression: multiple testing

One way to generate "stars" is to run a lot of regressions:

- With a 95% confidence threshold, you will find stars 5% of the time, even if your null hypothesis is true
- So if you run a lot of regressions, you are bound to find some with stars
- If you then only report the ones with stars, we have a problem
- (If you report all of them, we're fine!)

→ If we do this, we're killing the usefulness of stars

→ And generating results that won't replicate

Addressing multiple testing

There are a few fixes to this issue:

1. Pre-specification
    - Before you look at data, write down exactly what you're going to run
    - And make this public
    - So I can tell how many tests you actually did!

    → To make this credible, need to prove you couldn't see data first

Addressing multiple testing

There are a few fixes to this issue:

1. Pre-specification
   - Before you look at data, write down exactly what you're going to run
   - And make this public
   - So I can tell how many tests you actually did!
   - → To make this credible, need to prove you couldn't see data first
2. Multiple correction adjustment
   - Adjust for the fact that you ran many tests
   - Essentially involves inflating your p-values
   - Many ways to do this: FWER, FDR, Bonferroni
   - → This is best paired with prespecification

Dinkelman (2011): Spillovers?

Outcome: Δ_t female employment	OLS (1)	IV (2)	*N* communities (3)
Panel A.			
Full sample	−0.001	0.095*	1,816
	(0.005)	(0.055)	
Panel B.			
Excluding nonproject areas $<$ 1 km from project site	−0.004	0.076	1,205
	(0.006)	(0.057)	
Panel C.			
Excluding nonproject areas $<$ 5 km from project site	−0.003	0.069	840
	(0.008)	(0.077)	

Dinkelman (2011): Migration

	Δ_t log population		Δ_t females with high school		Δ_t males with high school	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)
Panel A.						
Eskom project	0.171***	3.897***	0.001	0.129*	0.001	0.076
	(0.045)	(1.427)	(0.005)	(0.058)	(0.003)	(0.050)
N	1,816	1,816	1,816	1,816	1,816	1,816
	Δ_t log non–in-migrant population		Δ_t female employment excluding in-migrants		Δ_t male employment excluding in-migrants	
Panel B.						
Eskom project	0.181***	4.349***	0.000	0.116*	−0.008	0.086
	(0.048)	(1.586)	(0.005)	(0.069)	(0.005)	(0.069)
N	1,816	1,816	1,816	1,816	1,816	1,816

Recap

TL;DR:

1. Dinkelman (2011) is a seminal study of the effects of rural electrification
2. Finds that electrification dramatically increases female employment
3. Uses an IV strategy based on land gradient (credible?)