Lecture 5:
Benefits and Costs of Energy Production
ECO 567A

Geoffrey Barrows[1]

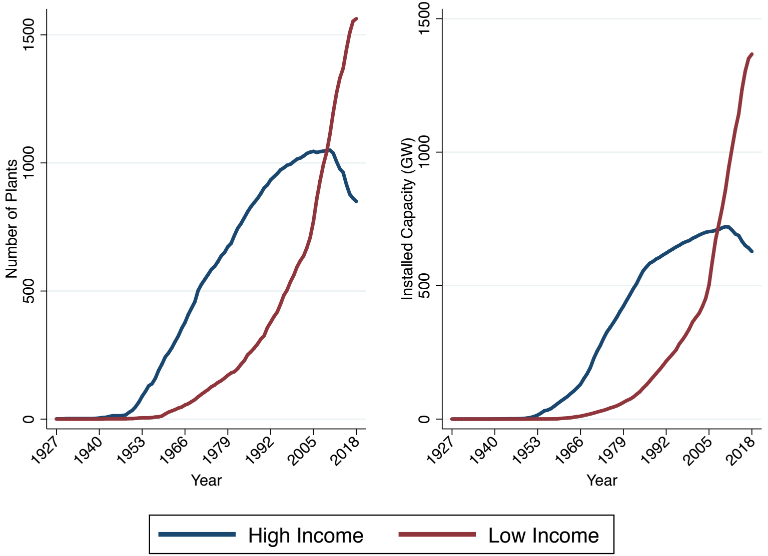
[1]CREST, CNRS, Ecole polytechnique

Feb 7, 2025

Syllabus

- Part I: Demand for Local Environmental Quality
  - Intro (Jan 10)
  - Demand I - Estimation (Jan 17)
  - Demand II - Sorting and Environmental Justice (Jan 24)
  - Amenities and Quant. Spatial Economic Models (Jan 31)
- Part II: Supply of Local Environmental Quality - Energy
  - Energy Production (Feb 7)
  - Energy Demand (Feb 14)
  - Energy Efficiency Innovation (Feb 21)
  - Trade and Pollution (March 7)
- Part III: Global Externalities
  - Climate Change (March 14)
- Final Exam March 19 9am - noon T5

Staggering Growth in Coal-fired Capacity in LDCs



India Continues to Build Coal-fired Capacity

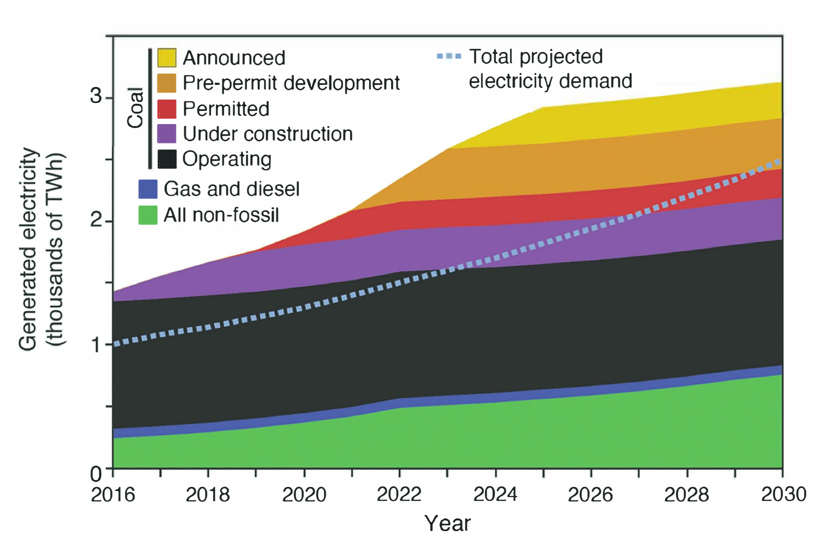


Figure: Shearer, Fofrich, and Davis (2017)

Conceptual Framework

- Utility

$$U = U(X, H) \tag{1}$$

- Health

$$H = H\left(M, \phi(Z(\theta))\right) \tag{2}$$

- Wages

$$w = w(H, \theta) \tag{3}$$

- Prices

$$p_x = p_x(\theta) \tag{4}$$

Conceptual Framework

$$\underset{X,M}{\text{Max}} \quad \mathcal{L} = U(X,H) - \lambda \left[p_x(\theta) * X + p_M * M - w(H,\theta)\right]$$

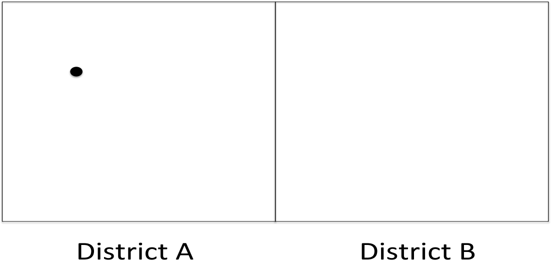
$$\begin{aligned} X^* &= X^*\left(p_x(\theta), p_M, \phi(\theta), \theta\right) \\ M^* &= M^*\left(p_x(\theta), p_M, \phi(\theta), \theta\right) \end{aligned}$$

Conceptual Framework

$$\frac{dH}{d\theta} = \underbrace{\left[\frac{\partial H}{\partial M}\frac{\partial M}{\partial \phi} + \frac{\partial H}{\partial \phi}\right]\frac{\partial \phi}{\partial Z}\frac{\partial Z}{\partial \theta}}_{\text{Net Pollution Effect}} + \underbrace{\frac{\partial H}{\partial M}\left[\frac{\partial M}{\partial w}\frac{dw}{d\theta} + \frac{\partial M}{\partial p_x}\frac{\partial p_x}{\partial \theta}\right]}_{\text{Supply Side Health Effect}} \tag{5}$$

$$\frac{dw}{d\theta} = \underbrace{\frac{\partial w}{\partial H}\frac{dH}{d\theta}}_{\text{Health Productivity Effect}} + \underbrace{\frac{\partial w}{\partial \theta}}_{\text{Supply Side Income Effect}} \tag{6}$$

Identification



$$
\begin{aligned}
E\left[\Delta Y_A\right] &= \Delta\text{Pollution} + \Delta\text{Local Income} + \Delta\text{Global Inc} + \Delta\text{Macro} \\
E\left[\Delta Y_B\right] &= \Delta\text{Global Income} + \Delta\text{Macro}
\end{aligned}
$$

$$
E\left[\Delta Y_A - \Delta Y_B\right] = \underbrace{\Delta\text{Pollution} + \Delta\text{Local Income}}_{\frac{dH}{d\theta}}
$$

Research Designs

- Fixed Effects/Difference-in-Difference
- Instrumental Variables

Fixed Effects

- Suppose the True Model is

$$y_{it} = \alpha + \beta * x_{it} + \underbrace{a_i + \epsilon_{it}}_{u_{it}}$$

$$E\left[\epsilon_{it}|x_{it}\right] = 0$$

$$E\left[u_{it}|x_{it}\right] \neq 0$$

- OLS is biased

$$y_{it} = \alpha + \beta * x_{it} + u_{it}$$

Within Transformation

Let

$$\begin{aligned}\bar{y}_i &= \frac{1}{T}\sum_{t=1}^{T}\left[\alpha + \beta * x_{it} + \underbrace{a_i + \epsilon_{it}}_{u_{it}}\right] \\ &= \alpha + \beta * \bar{x}_i + a_i + \bar{\epsilon}_i\end{aligned}$$

Then

$$y_{it} - \bar{y}_i = \beta * (x_{it} - \bar{x}_i) + (\epsilon_{it} - \bar{\epsilon}_i)$$

Or

$$\Delta y_{it} = \beta * \Delta x_{it} + \Delta \epsilon_{it}$$

And

$$cov(\Delta x_{it}, \Delta \epsilon_{it}) = 0$$

Difference Estimator/Fixed Effects

Difference estimator

$$\Delta y_{it} = \beta * \Delta x_{it} + \Delta \epsilon_{it}$$

Is functionally equivalent to Fixed Effect Model

$$\begin{aligned} y_{it} &= \alpha + \beta * x_{it} + \sum_{\ell} \delta_\ell * \mathbb{1}\left(i = \ell\right) + \epsilon_{it} \\ &= \alpha + \beta * x_{it} + \boldsymbol{\delta_i} + \epsilon_{it} \end{aligned}$$

And adding a time effect

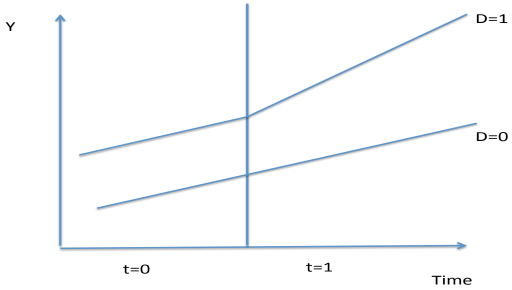
$$\begin{aligned} y_{it} &= \alpha + \beta * x_{it} + \sum_{\ell} \delta_\ell * \mathbb{1}\left(i = \ell\right) + \sum_{\tau} \delta_\tau * \mathbb{1}\left(t = \tau\right) + \epsilon_{it} \\ &= \alpha + \beta * x_{it} + \boldsymbol{\delta_i} + \boldsymbol{\delta_t} + \epsilon_{it} \end{aligned}$$

Diff-n-diff

- $y_{it} = \alpha_0 + \alpha_1 * \mathbb{1}(D_i = 1) + \alpha_2 * \mathbb{1}(t = 1) + \alpha_3 * \underbrace{\mathbb{1}(D_i = 1) * \mathbb{1}(t = 1)}_{=x_{it}} + \epsilon_{it}$

- Diff-in-diff

$$[E[y_{it}|D_i = 1, t = 1] - E[y_{it}|D_i = 0, t = 1]] -$$
$$[E[y_{it}|D_i = 1, t = 0] - E[y_{it}|D_i = 0, t = 0]] =$$
$$[\alpha_1 + \alpha_3] - [\alpha_1] = \alpha_3$$

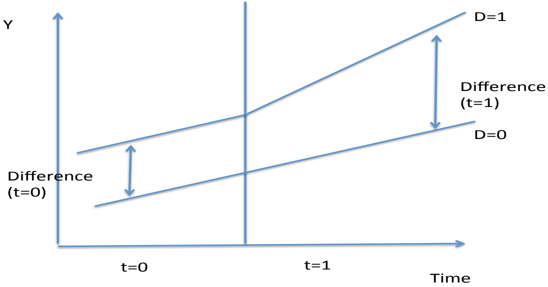


Diff-n-diff

- $y_{it} = \alpha_0 + \alpha_1 * \mathbb{1}(D_i = 1) + \alpha_2 * \mathbb{1}(t = 1) + \alpha_3 * \underbrace{\mathbb{1}(D_i = 1) * \mathbb{1}(t = 1)}_{=x_{it}} + \epsilon_{it}$

- Diff-in-diff

$$[E[y_{it}|D_i = 1, t = 1] - E[y_{it}|D_i = 0, t = 1]] -$$
$$[E[y_{it}|D_i = 1, t = 0] - E[y_{it}|D_i = 0, t = 0]] =$$
$$[\alpha_1 + \alpha_3] - [\alpha_1] = \alpha_3$$

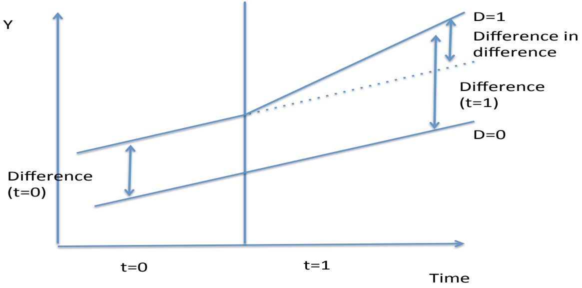


Diff-n-diff

- $y_{it} = \alpha_0 + \alpha_1 * \mathbb{1}(D_i = 1) + \alpha_2 * \mathbb{1}(t = 1) + \alpha_3 * \underbrace{\mathbb{1}(D_i = 1) * \mathbb{1}(t = 1)}_{=x_{it}} + \epsilon_{it}$

- Diff-in-diff

$$[E[y_{it}|D_i = 1, t = 1] - E[y_{it}|D_i = 0, t = 1]] -$$
$$[E[y_{it}|D_i = 1, t = 0] - E[y_{it}|D_i = 0, t = 0]] =$$
$$[\alpha_1 + \alpha_3] - [\alpha_1] = \alpha_3$$



Papers

- Benefits
  - Lipscomb et al (AEJ, 2013)
- Costs
  - Clay et al (NBER, 2016)
  - Davis (Restat, 2013)

Davis (Restat, 2011)

Research Question: What is the MWTP to avoid living near a power plant?

- Should be positive, so maybe we should site plants far from populations
- But MWTP to avoid power plants trades off against transmission costs

Davis (Restat, 2011)

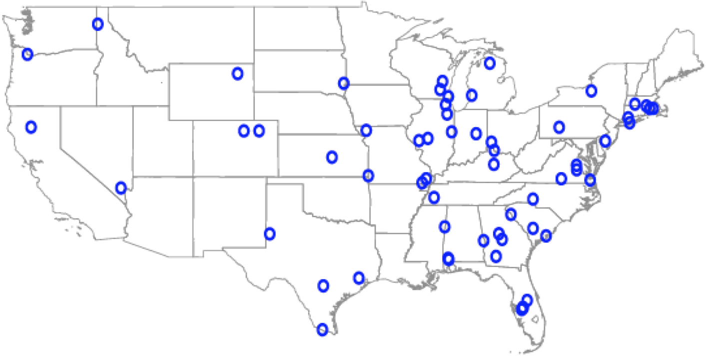
- Identification problem
  - placement of power plants is endogenous to local characteristics
- Empirical Strategy
  - Compare housing values before and after a plant opens within 2 miles vs between 2-5 miles
  - Should control for most omitted variables

Davis (Restat, 2011)

- Empirical setting: US
  - Opened 60 power plants between 1990 and 2000
  - Individual-level housing prices

Davis (2011)

Figure 1: Power Plants Opened During the 1990s



Comparing near vs far in 1990

	(1) 0-2 miles	(2) 2-5 miles	(3) Entire U.S.
Household Demographics			
Household Income (1000s)	30.2	34.2	32.5
Household Size (persons)	2.57	2.49	2.35
Number of Individuals Under 18 Per Household	.71	.63	.60
Number of Individuals Over 65 Per Household	.27	.31	.29
Proportion Household Head Completed High School	.71	.75	.77
Proportion Household Head Completed College	.25	.24	.31
Proportion Household Head Black	.11	.07	.08
Proportion Household Head Hispanic	.21	.12	.18
Housing Characteristics			
House Value (1000s)	100.1	104.3	98.7
Monthly Rent	478.9	461.5	470.0
Proportion Occupied	.88	.92	.87
Proportion Owner Occupied	.52	.60	.59
Proportion 0-2 Bedrooms	.50	.45	.44
Proportion 3-4 Bedrooms	.47	.53	.52
Proportion Built Last 5 Years	.09	.09	.11
Proportion Built Last 10 Years	.16	.16	.20
Proportion Complete Plumbing	.97	.99	.98
Proportion One or More Acres	.08	.08	.14
Proportion Ten or More Acres	.03	.02	.07
Proportion Multi-Unit	.41	.30	.24

Difference-in-difference Estimation

$$\begin{aligned}
\ln\ price_{ij_pt} &= \alpha_1 * x_{ij_pt} + \alpha_2 * \mathbb{1}(j \text{ within 2 miles of } p) * \mathbb{1}(\text{t=2000}) \\
&+ \boldsymbol{\alpha}_3 * \mathbb{1}(\text{t=2000}) * \mathbb{1}(\text{plant=}p) + CF_{j_p} + \epsilon_{ij_pt}
\end{aligned}$$

- p indexes plant
- j indexes census block
- i indexes house
- $\boldsymbol{\alpha}_3$ is a vector of plant-specific time-trend parameters to be estimated

Davis (2011)

Table 2. The Effect of Power Plants on Housing Values and Rents
Difference-in-Differences Estimates

	Housing Values			Rents		
	(1)	(2)	(3)	(4)	(5)	(6)
*1(within two miles) * 1(year 2000)*	-.030	-.023	-.030	-.044	-.055	-.049
	(.040)	(.033)	(.027)	(.013)	(.016)	(.016)
Housing Characteristics	yes	yes	yes	yes	yes	yes
Power Plant Indicators x *1(within two miles)*	yes	-	-	yes	-	-
Power Plant Indicators x *1(year 2000)*	yes	yes	yes	yes	yes	yes
Census Block Fixed Effects	no	yes	yes	no	yes	yes
Propensity Score Weighting	no	no	yes	no	no	yes
Number of Observations	170,821	170,821	170,821	87,690	87,690	87,690
Number of Census Blocks	27,848	27,848	27,848	17,765	17,765	17,765
R^2	.62	.35	.34	.31	.18	.16

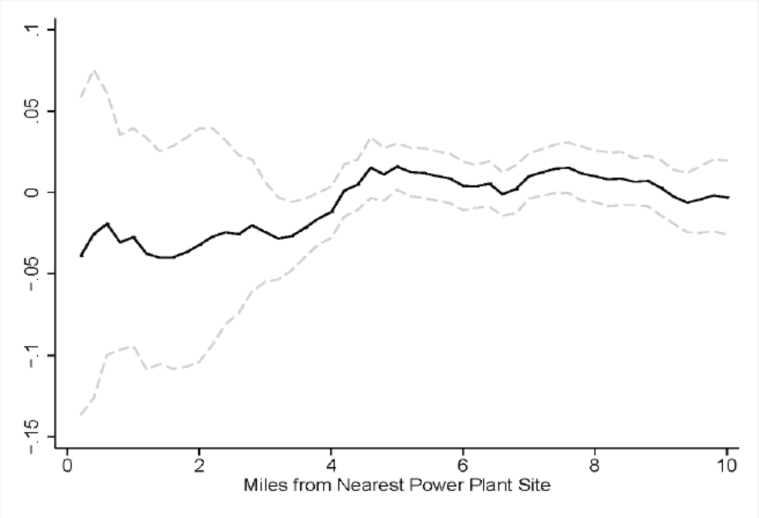
Note: The dependent variable in columns (1)-(3) is housing value in logs and the dependent variable in columns (4)-(6) is monthly rent in logs. The sample includes homes located within five miles of one of sixty 100 megawatt power plants opened in the United States during the 1990s. The variable *1(within two miles)* indicates homes located within two miles of the nearest power plant site. The variable *1(year 2000)* is an indicator for observations from the 2000 census. Standard errors clustered by power plant site are shown in parentheses. In columns (1),(2),(4) and (5) observations are weighted using sampling weights. In columns (3) and (6) observations in the comparison group are weighted using propensity scores.

Key Concerns

- Contamination of the controls (outward migration increases housing prices in control areas)
- Treated vs control areas were trending differentially prior to 1990
- Siting based on expectations of future trends

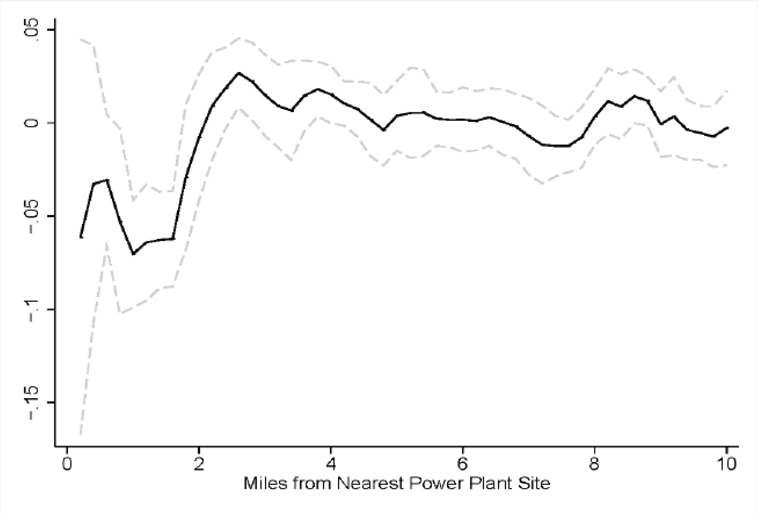
Davis (2011)

Figure 2: Distance Gradient for Housing Values



Davis (2011)

Figure 3: Distance Gradient for Rents



Heterogeneous Effects

	Housing Values		Rents	
By Primary Fuel:				
Natural Gas Plants	-.033	(.029)	-.050	(.016)
Coal Plants	-.027	(.050)	-.069	(.050)
By Year Plant Opened:				
1991-1995	-.010	(.019)	-.039	(.020)
1996-1999	-.053	(.060)	-.069	(.016)
By Plant Capacity:				
Large Capacity Plants	-.055	(.063)	-.053	(.014)
Small Capacity Plants	-.012	(.015)	-.038	(.045)
By Plant Utilization:				
Base Load Plants	-.035	(.053)	-.045	(.013)
Peaker Plants	-.024	(.016)	-.005	(.038)
By Prevailing Wind Direction:				
Homes Downwind of Plant	-.033	(.034)	-.043	(.034)
Homes Upwind of Plant	-.036	(.028)	-.052	(.016)

Summary of Results

- Results
  - power plants drop housing prices by 3-5%.
  - Average number of houses within 2 miles is 2900, with an average value of 166,000 $\implies$ .03*(2900*166,00) = 14.5 million USD in market capitalization
  - Cost of transmission lines $800,000 USD per mile
  - If density is not too high (you don't have to move the plant to far), moving a plant 1 or 2 miles away might be worth it.

Lipscomb et al (AEJ, 2013)

Research Question: does large-scale electrification infrastructure growth increase development?

- Increase electricity supply should lower cost of electricity, which should have lots of benefits (absent negative externalities)
- But how big are the gains?
- Also, if observed rollout depends on political calculations, then benefits might be small

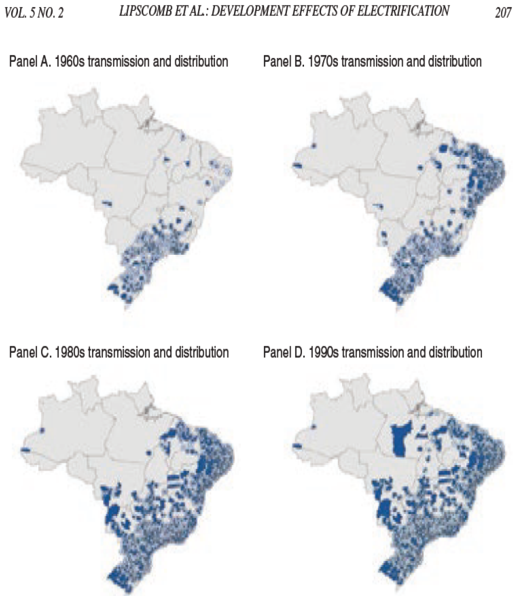
Lipscomb et al (AEJ, 2013)

- Identification problem
  - placement of power plants/transmission lines is endogenous to local characteristics - wealthy regions (cross-section), growing regions (over time)
- Empirical Strategy
  - Estimate IV impact on county-level development

Lipscomb et al (AEJ, 2013)

- Empirical setting: Brazil
  - Electricity access expanded from 32% to 75% between 1960 and 2000
  - 85% Hydro power

Observed Electrification Rollout



Hydro Instrument

1. Predict suitability of grid points for building dams based on geographical characteristics

   whether there is a river, the average and maximum gradient of the river, maximum water flow accumulation

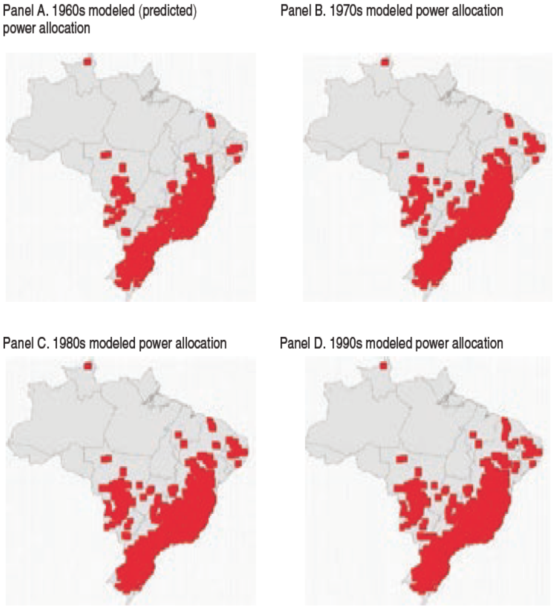
2. Rank grid points in terms of suitability

3. "build dams" starting from the top of the list until the national budget for the decade is exhausted. Then continue with the list in the next decade, etc.

4. Given predicted dam placement, build counterfactual transmission lines using lowest cost algorithm

5. Identify any grid point within 50km of a substation of the counterfactual transmission lines is counterfactually electrified

Predicted Electrification Rollout



Estimation

- First Stage

$$E_{c,t-1} = \alpha_c^1 + \gamma_{t-1}^1 + \theta Z_{c,t-1} + \eta_{c,t-1} \tag{7}$$

$E_{c,t-1}$ is the proportion of grid points electrified in county c in decade $t-1$ and $Z_{c,t-1}$ is the model-predicted proportion

- Second Stage

$$Y_{ct} = \alpha_c^2 + \gamma_t^2 + \beta \widehat{E_{c,t-1}} + \epsilon_{ct} \tag{8}$$

$\widehat{E_{c,t-1}}$ is predicted electrification from the first stage

Identification Assumption

The assumption is $E\left[Z_{c,t-1} * \epsilon_{ct}\right] = 0$

Identification Assumption

The assumption is $E\left[Z_{c,t-1} * \epsilon_{ct}\right] = 0$

- How could the assumption fail?

Identification Assumption

The assumption is $E\left[Z_{c,t-1} * \epsilon_{ct}\right] = 0$

- How could the assumption fail?
  - If people move over time towards the locations predicted to receive electricity for other reasons

    from the lowest cost locations (robust water flow with a steep river gradient) in the early decades to slightly more expensive (flatter and less water-rich) locations in later years

First Stage Results

TABLE 3—FIRST-STAGE REGRESSIONS
DEPENDENT VARIABLE: ACTUAL ELECTRICITY AVAILABILITY FROM INFRASTUCTURE INVENTORIES

Modeled electricity availability	0.563***	0.323***	0.222***
	(0.03)	(0.05)	(0.05)
Year FE	Yes	Yes	Yes
County FE	No	Yes	Yes
Amazon × year dummies	No	No	Yes
R^2	0.369	0.840	0.866
Observations	8,730	8,730	8,730
F-Stat	336.3	34.71	24.6
p-value	0.00	0.00	0.00

Notes: The dependent variable is prevalence of electricity infrastructure in the county. Regressions are weighted by county area. Standard errors clustered by county in parentheses. Measures of electrification are lagged by a decade in all our second-stage regressions, and the data used for the first-stage regression therefore covers 1960–1990. The Amazon and Pantanal are referred to jointly as the Amazon.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

OLS and Second Stage IV - Housing Prices

TABLE 6—HOUSING VALUES
DEPENDENT VARIABLE: AVERAGE VALUE OF HOUSING

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged electricity infrastructure	5.023***	1.326***	0.801***	8.468***	7.792***	8.811***
	(0.90)	(0.35)	(0.27)	(1.52)	(1.72)	(3.03)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
County FE?	No	Yes	Yes	No	Yes	Yes
Amazon × year dummies[a]	No	No	Yes	No	No	Yes
R^2	0.153	0.922	0.925	0.106	0.191	0.151
Observations	8,730	8,730	8,730	8,730	8,730	8,730

Notes: Standard errors clustered by county in parentheses. The dependent variable is average housing value in thousands of reais. All regressions have county size weights and year dummies. The average housing value in the sample is 13.048.

[a]Topographic factor is interacted with a full set of decade fixed effects in order to flexibly control for differential trends by that Topographic factor.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

OLS and Second Stage IV - Development Index

Table 7—Human Development Index

	OLS			IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Lagged electricity infrastructure	0.036**	0.009	0.006	0.091***	0.091***	0.109**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.04)
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes
County FE?	No	Yes	Yes	No	Yes	Yes
Jungle × year dummies[a]	No	No	Yes	No	No	Yes
R^2	0.657	0.960	0.960	0.640	0.931	0.930
Observations	8,730	8,730	8,730	8,730	8,730	8,730

Notes: Standard errors clustered by county in parentheses. The dependent variable is the human development index. Year dummies are included in all regressions. All regressions have county size weights. The average HDI value in the sample is 0.557.

[a]Topographic factor is interacted with a full set of decade fixed effects in order to flexibly control for differential trends by that Topographic factor.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

Measurement Error

- Let the true model be $y = \beta x + \epsilon$
- Let x be measured with error $\tilde{x} = x + u$, $E(u) = 0$
- Substituting in we have $y = \beta \tilde{x} + \underbrace{\gamma}_{(\epsilon - \beta u)}$
- $\tilde{x}$ and u are correlated (from $\tilde{x} = x + u$), so the OLS is biased if we take $\tilde{x}$ as the input variable instead of x.

OLS and Second Stage IV - Development Index

TABLE 11—HUMAN DEVELOPMENT INDEX COMPONENTS AND OTHER POVERTY MEASURES

	HDI: Longevity		HDI: Income		HDI: Education	
	OLS	IV	OLS	IV	OLS	IV
Panel A						
Lagged electricity infrastructure	0	−0.01	−0.03*	0.45***	0.03***	0.19***
	(0.01)	(0.05)	(0.02)	(0.15)	(0.01)	(0.06)
R^2	0.84	0.8	0.89	0.5	0.91	0.65
Observations	8,730	8,730	8,730	8,730	8,730	8,730
Mean of dep var	0.569		0.472		0.515	
	Infant mortality		Gross income PC		Poverty	
	OLS	IV	OLS	IV	OLS	IV
Panel B						
Lagged electricity infrastructure	−7.99***	−11.97	−0.01	0.11**	−0.76	−42.17***
	(2.42)	(18.08)	(0.01)	(0.05)	(1.39)	(13.84)
R^2	0.9	0.86	0.84	0.58	0.85	0.53
Observations	8,730	8,730	8,730	8,730	8,730	8,730
Mean of dep var	71.96		0.114		60.469	

Notes: Standard errors clustered by county in parentheses. Dependent variables are the component indices of the HDI. All regressions have county size weights, year dummies, and jungle × year dummies.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Interpreting the Results

- Why are the IV estimates larger than the OLS?
  - Measurement error in $E_{c,t-1}$
  - Heterogeneous treatment effects (LATE)
  - Government targeted poorer areas in actual infrastructure rollout

Summary of Results

- Cost-Benefit Analysis
  - It costs \$33 million USD to electrify a county $\Longrightarrow$ \$40 billion USD for observed electrification of Brazil between 1960 - 2000
  - Electrification (moving from 0 to 100% electrification in the county) increases housing values about 9000 reis (point estimate 8.811) , which is equivalent to \$4,900 USD per house. With an average of 12,400 HH per county $\Longrightarrow$ \$61 million USD per county.
  - Internal rate of return $= 61/33 = 184\%$

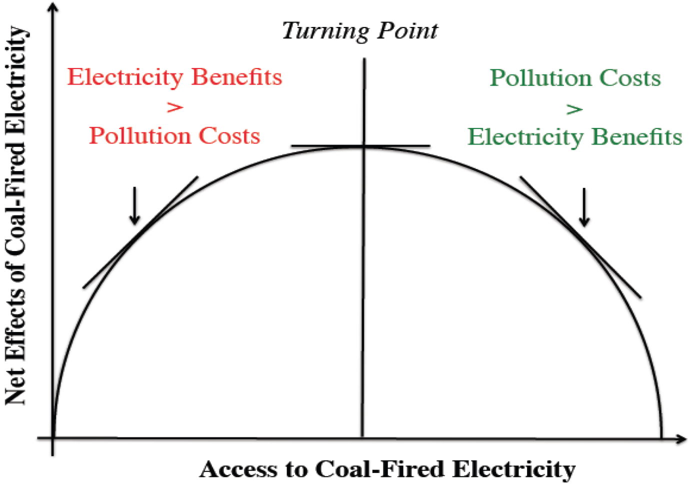
Clay et al (2016)

Research Question: How do coal-fired power plants impact housing prices and infant health?

- Pollution certainly lowers housing prices and increases IMR
- But electrification also increases wages, productivity, lowers prices, etc

1. Do effects differ by initial electrification rates
2. Do agents weight costs vs benefits

Clay et al (2016)

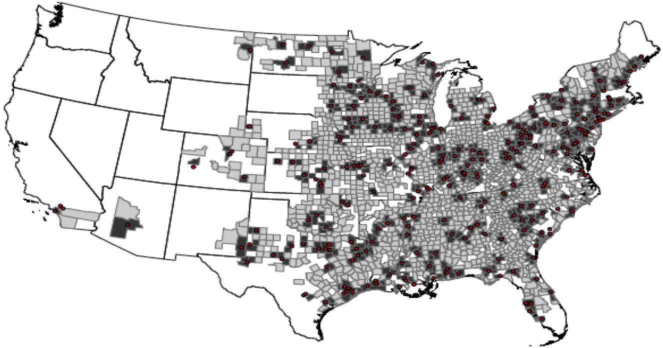


Clay et al (2016)

- Identification problem
  - Coal plant placement is endogenous to determinants of infant health and housing prices
- Empirical Strategy
  - Looking at historical US, high transmission costs means that people near the plants are affected by both channels
  - Compare county-level housing prices and IMR for counties within 30 miles of a plant and those between 30-90 miles from the plant

Treated vs Control Counties

Figure A.3: Sample Counties for Power Plants Openings Analysis



Notes: This figure reports the sample of 1,924 counties used in the difference-in-differences. Red dots identify the 272 power plants that opened between 1938 and 1962. Black identifies 'treatment' counties ($\leq$ 30 miles of a power plant), grey identifies 'control' counties (30-90 from a power plant).

Empirical Model

- Plant opening Model

$$\begin{aligned} Y_{cpt} &= \alpha + \beta * \mathbb{1}(\text{PPOpen})_{pt} + \delta * [\mathbb{1}(\text{PPOpen})_{pt} * \mathbb{1}(\text{Near})_{cp}] \\ &+ \psi X_{cpt} + \eta_{cp} + \lambda_{st} + \epsilon_{cpt} \end{aligned}$$

- Event Study

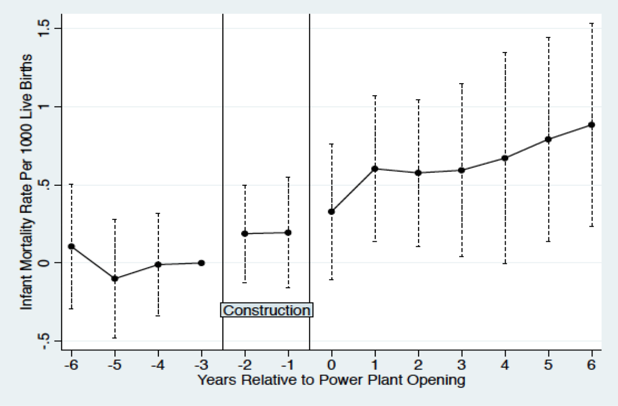
$$\begin{aligned} Y_{cpt} &= \alpha + \sum_{\tau=-6}^{6} \beta^{\tau} * 1(t - e_p = \tau) * \mathbb{1}(\text{PPOpen})_{pt} \\ &+ \sum_{\tau=-6}^{6} \delta^{\tau} * 1(t - e_p = \tau) * [\mathbb{1}(\text{PPOpen})_{pt} * \mathbb{1}(\text{Near})_{cp}] \\ &+ \psi X_{cpt} + \eta_{cp} + \lambda_{st} + \epsilon_{cpt} \end{aligned}$$

- Coal Consumption Model

$$Y_{ct} = \alpha + \beta * Coal_{ct} + \psi X_{ct} + \theta_t Z_c + \eta_c + \lambda_{st} + \epsilon_{ct}$$

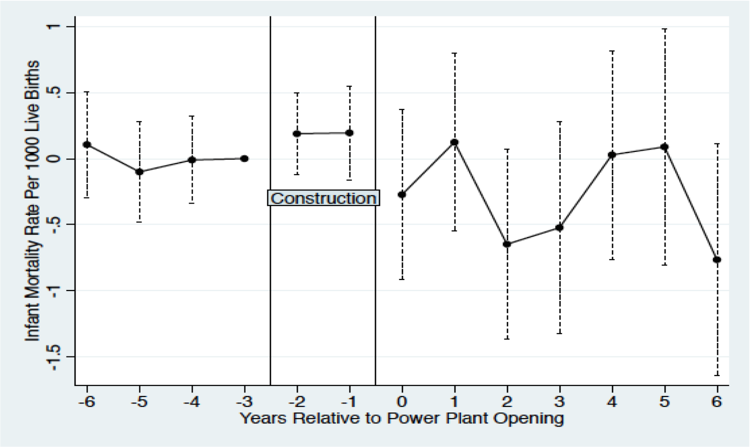
Event Study - Large Plants

Figure 5: Event Study: The Effect of Coal-Fired Power Plant Openings on Infant Mortality



(a) Large Power Plants (>75 MW)

Event Study - Small Plants



(b) Small Power Plants (<75 MW)

Results - Plant Openings

Table 3: Effects of Large and Small Power Plant Openings on Infant Mortality

Dependent Variable	Infant Mortality Rate			
	(1)	(2)	(3)	(4)
Large Power Plants (>75 MW) vs. Small Power Plants ($\leq$75 MW)				
1(Large Plant Operating) $\times$ 1(County within 30 miles)	2.1700**	1.7577***	1.0093***	0.7862***
	(1.0025)	(0.4321)	(0.3848)	(0.2643)
1(Small Plant Operating) $\times$ 1(County within 30 miles)	1.4292**	0.2846	-0.0972	-0.2101
	(0.6249)	(0.3902)	(0.3625)	(0.3084)
R-squared	0.6441	0.6946	0.6961	0.7045
F-statistic: $\delta^{Large} = \delta^{Small}$	1.42	15.13	9.04	9.19
P-value	0.2332	0.0001	0.0027	0.0025
Observations	130,025	130,025	130,025	130,025
County-Plant Pair FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
State-by-Year FE	N	Y	Y	Y
1940 Manufacturing Employment $\times$ Year FE	N	N	Y	Y
Additional Covariates	N	N	N	Y

Results - Coal Consumption

Table 6: Effects of Coal Consumption and Coal-fired Capacity by Baseline Electricity Access

Dependent Variable	**Infant Mortality**		
	(1)	(2)	(3)
Panel A. Coal Consumption within 30 Miles			
Coal Consumption × 1(L-Electricity)	-0.1414***	-0.0629**	-0.0503**
	(0.0303)	(0.0246)	(0.0240)
Coal Consumption × 1(H-Electricity)	0.1767***	0.1483***	0.1359***
	(0.0196)	(0.0180)	(0.0198)
R-squared	0.6496	0.6956	0.7000
Panel B. Coal-Fired Capacity within 30 Miles			
Coal Capacity × 1(L-Electricity)	-0.3350***	-0.1260*	-0.0962
	(0.0885)	(0.0720)	(0.0674)
Coal Capacity × 1(H-Electricity)	0.2309***	0.2094***	0.2014***
	(0.0416)	(0.0272)	(0.0343)
R-squared	0.6488	0.6957	0.7001
Observations	49,575	49,575	49,575
County & Year FE	Y	Y	Y
State-by-Year FE	N	Y	Y
All Controls	N	N	Y

Results - Coal Consumption

Table 8: Effects of Coal Consumption and Coal-fired Capacity on Property Values, by Baseline Electricity Access

Dependent Variable	Infant Mortality		Log Median Rent		Log Median Housing Value	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Coal Consumption within 30 Miles						
Coal Consumption × 1(L-Electricity)	-0.2169***	-0.1760***	0.0032***	0.0028***	0.0007	0.0018
	(0.0678)	(0.0610)	(0.0009)	(0.0008)	(0.0013)	(0.0011)
Coal Consumption × 1(H-Electricity)	0.2411***	0.2487***	-0.0204***	-0.0061**	-0.0147***	-0.0032**
	(0.0709)	(0.0546)	(0.0042)	(0.0024)	(0.0027)	(0.0013)
R-squared	0.7411	0.7613	0.9305	0.9551	0.9248	0.9422
Panel B. Coal-Fired Capacity within 30 Miles						
Coal Capacity × 1(L-Electricity)	-0.5817***	-0.4108**	0.0085***	0.0064***	0.0023	0.0039
	(0.1893)	(0.1634)	(0.0024)	(0.0022)	(0.0029)	(0.0027)
Coal Capacity × 1(H-Electricity)	0.2976***	0.3701***	-0.0247***	-0.0086***	-0.0184***	-0.0019
	(0.0652)	(0.0768)	(0.0043)	(0.0028)	(0.0036)	(0.0024)
R-squared	0.7437	0.7640	0.9302	0.9551	0.9247	0.9421
Observations	3,708	3,708	3,708	3,708	3,707	3,707
County & Year FE	Y	Y	Y	Y	Y	Y
All Controls	N	Y	N	Y	N	Y

Summary of Results

- Opening a large coal-fired power plant generated 0.8 additional infant deaths per 1000 live births per year (3 percent increase in IMR), or 22,015 deaths during the sample.

- But coal-fired plants only increase IMR in place that already had high electrification rates. In places with low electrification rates, IMR declined.

- Housing values echo the IMR results: at low levels of electrification, coal plants increase property values, but at high levels of electrification, coal plants lower housing values.

References

- Davis, Lucas W. "The effect of power plants on local housing values and rents." Review of Economics and Statistics 93.4 (2011): 1391-1402.
- Clay, Karen, Joshua Lewis, and Edson Severnini. Canary in a Coal Mine: Infant Mortality, Property Values, and Tradeoffs Associated with Mid-20th Century Air Pollution. No. w22155. National Bureau of Economic Research, 2016.
- Lipscomb, Molly, Mushfiq A. Mobarak, and Tania Barham. "Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil." American Economic Journal: Applied Economics 5.2 (2013): 200-231.