Environmental and Development Economics
Module 2 - How does development affect the environment?

Raahil Madhok
UMN Applied Economics

2024-09-12

Lecture 2

How does development affect the environment?

Housekeeping

- Replication project (Ryan announcement)
  - Due Oct 17th
- Research proposal
  - Start thinking about idea + come talk to me
  - First draft: Oct 3rd (pass/fail)
- Lecture slides will be posted before each class
- Reading list is a work in progress
- 5 min break mid-way through lectures

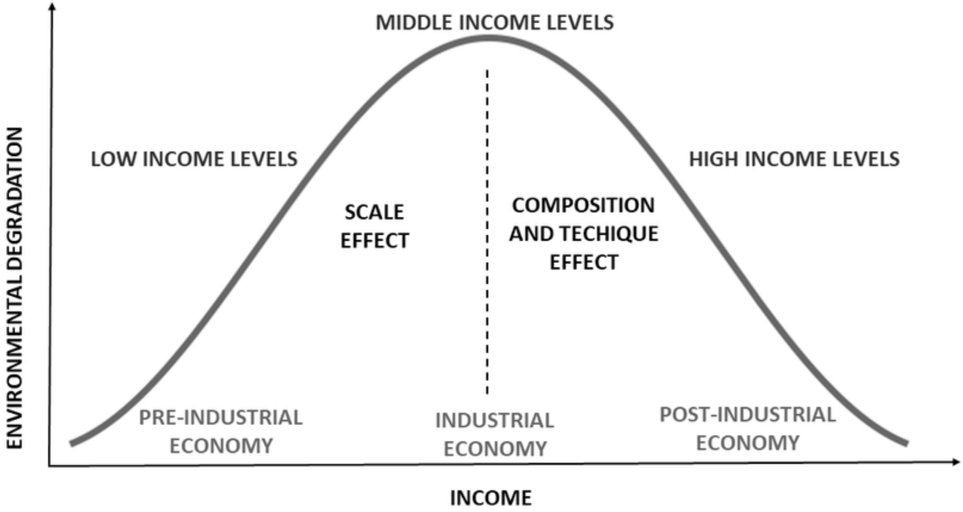
Today

- **Guiding question:** how does economic development affect the environment?
- Descriptive overview
- Channel I: Income effects (today)
  - **Changes in consumption**
  - **Energy**
  - Diet
- Channel II: Technology and Infrastructure (next week)
  - Agricultural productivity
  - Infrastructure
- Channel III: Institutions
  - Later in the course

How does economic development affect the environment?

- There is no one answer
- Choices that maximize economic growth often degrade environment
- But development expands our choice set (e.g. clean energy)
- Development also lowers $u'(c)$ (model from last week)
- Hence: the Environmental Kuznets Curve (EKC)

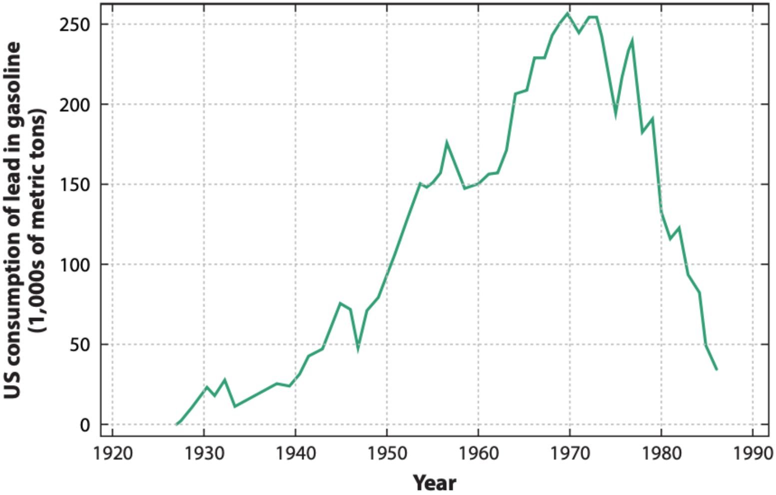
Environmental Kuznets Curve



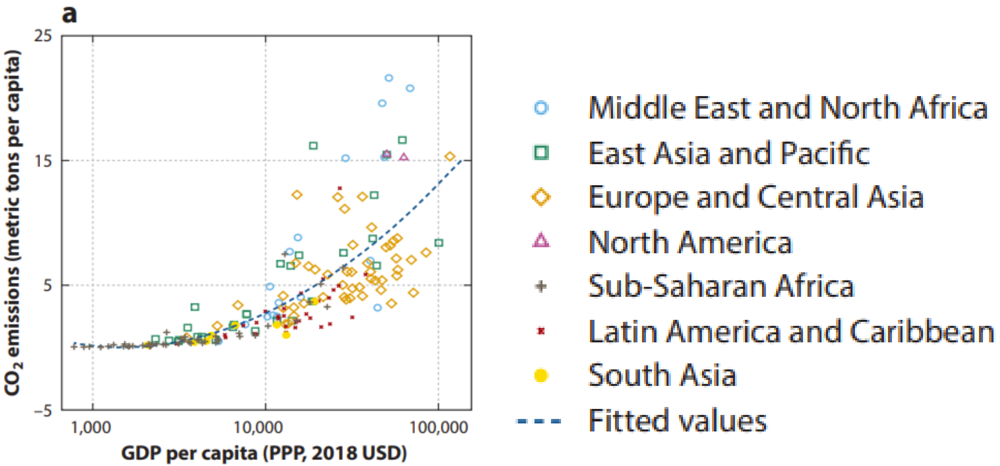
Discussion: Do you believe EKC?

- Mechanisms?
- Causality?
- Robustness?

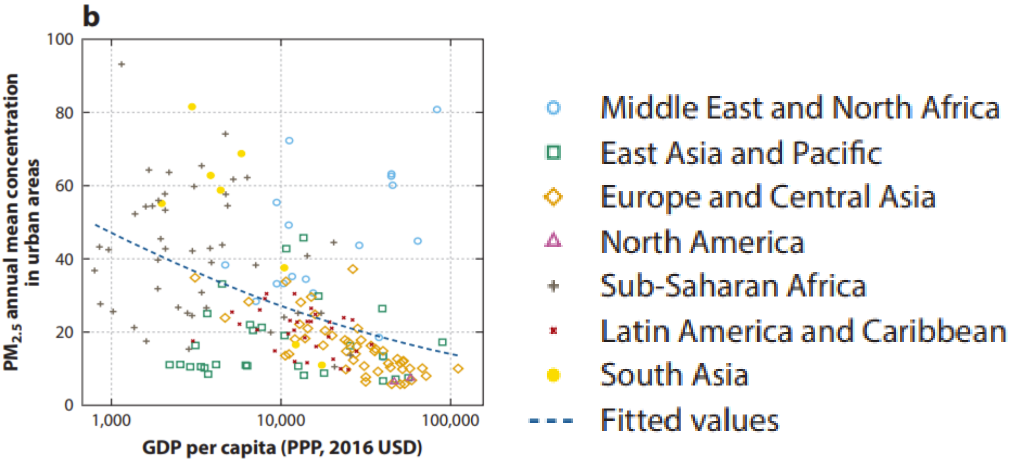
GDP and Lead in Gasoline



GDP and CO2

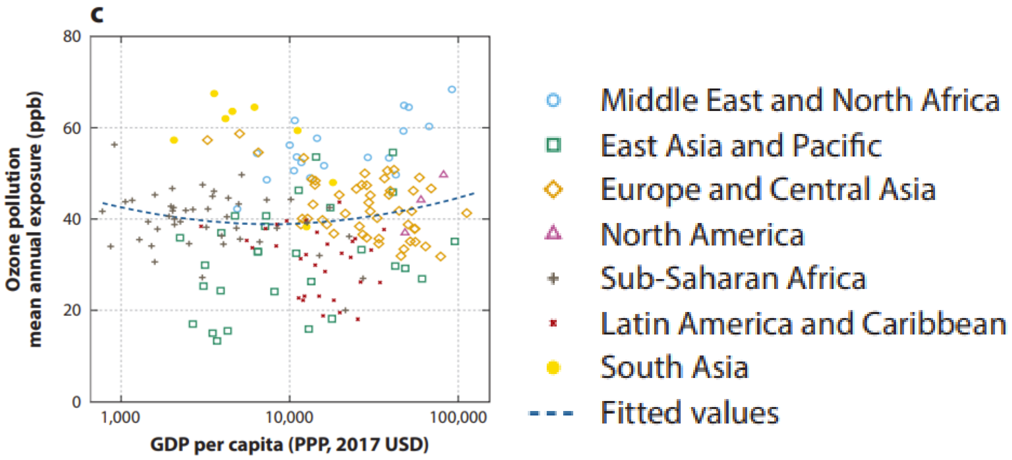

a
25
15
5
–5
1,000
10,000
100,000
GDP per capita (PPP, 2018 USD)
CO2 emissions (metric tons per capita)
Middle East and North Africa
East Asia and Pacific
Europe and Central Asia
North America
Sub-Saharan Africa
Latin America and Caribbean
South Asia
Fitted values

GDP and PM2.5



b
100
80
60
40
20
0
1,000
10,000
100,000
GDP per capita (PPP, 2016 USD)
PM2.5 annual mean concentration in urban areas
Middle East and North Africa
East Asia and Pacific
Europe and Central Asia
North America
Sub-Saharan Africa
Latin America and Caribbean
South Asia
Fitted values

GDP and Ozone



My view

- Macro correlations give limited answer
- Not the right question
  - What do we mean by development?
  - e.g. manufacturing $\uparrow$ pollution, but services may not
- Narrow, micro/empirical papers more helpful
  - even if they don't find consensus

Discussion: What is the ideal (quasi-) experiment?

- Can we randomize GDP?
- What variation in “development” can approximate the experiment?
  - what proxies development? (see ookla speed test)
  - what does the proxy miss?
  - what (non-random) variation isolates causal relationship?

Discussion: What about mechanisms?

- **Why and how** does economic development affect the environment?
- Theory is your friend: especially total derivatives. . .
  - Sometimes enough to rationalize mechanisms
  - Otherwise generate testable predictions
- Heterogeneity is your friend
  - But have a reason **why** you choose one channel over another

Alix-Garcia et al. (2013): Ecological footprint of poverty alleviation

Question: What is the effect of income on environmental quality?

- Poverty alleviation may $\uparrow$ demand for resource-intensive good
- Or, it may $\uparrow$ demand for conservation
- **This paper:** Exploit RDD eligibility to study impact of CCT on forests

Why can't we regress forests on income?

- Income correlated with other things that affect environment
- Like what?
- If we control for OVB, will panel fixed effects work?

Mexico's Oportunidades Program

- **Goal:** Ambitious program to build human capital
- **Scope:** \$2.6 billion/year (half of anti-poverty budget)
  - Transfers are 1/3 of total income
  - Represents shift to new income path
- **Size:** 4 million households treated
- **Design:** Randomized rollout from 1998-2000
  - Afterwards, enrollment is non-random and based on **marginality index**
  - Two step: 1) village selected on index; 2) households selected in eligible villages

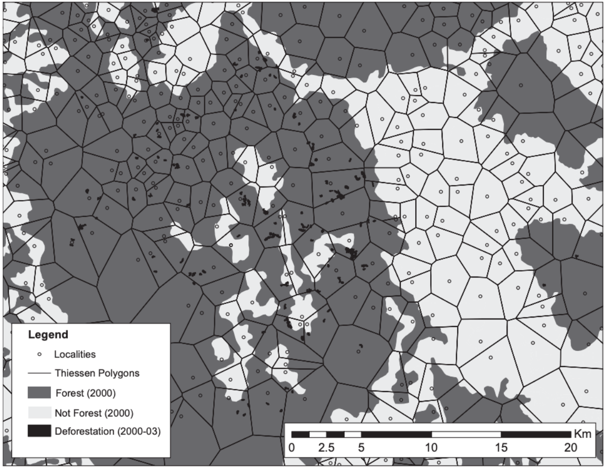
Results Preview: conditional cash transfer increases deforestation

Data

- **Main variables:** village lat/lon, eligibility index, forest cover (satellite)
  - Enrollment by village until 2003
- **Sample size:** 105,749 villages
- **Unit of analysis:** only village centroids provided?
  - How do we estimate village-level impacts? Thiessen polygons
  - Area around a point where every location closer to the point than to all others
  - Problems?

Thiessen/Voronoi Polygons

Figure 2.—Illustration of Locality Boundaries Defined Using Thiessen Polygons



Data Visualization

Eligibility cutoff: -1.2

FIGURE 3.—ENTIRE SAMPLE MINUS OBSERVATIONS WITH INDEX > 3

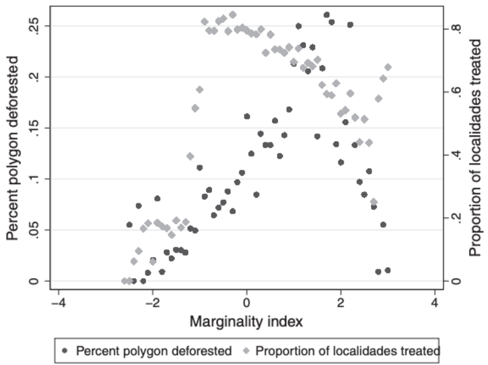
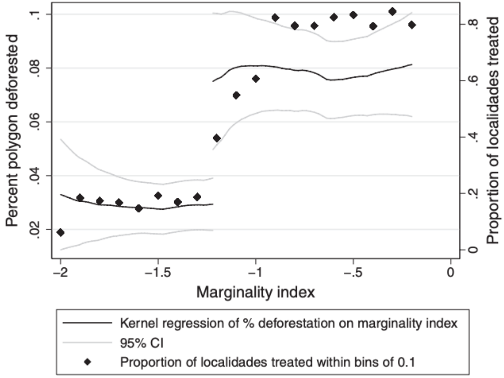


FIGURE 4.—KERNEL ESTIMATION OF DEFORESTATION ON MARGINALITY INDEX—RESTRICTED SAMPLE



Empirical Strategy: OLS Equation

$$\Delta f_i = \alpha + \delta T_i + \beta' X_i + \epsilon_i$$

- where Δf_i is % deforestation in polygon i over 2000-2003
- $T_i = 1$ if the locality enrolled by 2003
- X_i includes poverty
- But T_i is not random!

- If discontinuity is sharp, replace T_i with $1[E_i > -1.2]$
- RDD: Sample window $-2 < I < -0.2$

RDD Equation

- Discontinuity is fuzzy, not sharp
  - How do we know? Two reasons
- Solution: Use cutoff to instrument treatment probability
- Fuzzy RDD/IV:

$$\Delta f_i = \alpha + \delta T_i + \gamma I_i + \beta' X_i + \epsilon_i$$

$$\Delta T_i = \omega + \tau_1 E_i + \tau_2 E_i I_i + \tau_3 M_i + \tau_4 M_i I_i + \mu I_i + \Gamma' X_i + \epsilon_i$$

where T_i is enrollment dummy, E_i is eligibility cutoff dummy, I_i is index, M_i is dummy for region where enrollment increases rapidly

Identification Assumptions

Results: OLS and Sharp RDD

	Tobit			OLS	
	% Polygon Deforested			Deforestation (0/1)	% Deforested (If 1)
	(1)	(2)	(3)	(4)	(5)
Eligible	.383	.549	.370	.013	.387
	(.181)**	(.295)*	(.217)*	(.008)*	(.190)**
Marginality index	.523	.753	.219	.031	.069
	(.041)***	(.077)***	(.189)	(.003)***	(.075)
Index^2		.069		.002	.060
		(.072)		(.003)	(.075)
Index^3		−.100		−.004	−.022
		(.037)***		(.001)***	(.025)
Index^4		−.002		−.0001	−.012
		(.015)		(.0005)	(.013)
Baseline area in forest, 2000	−.0004	−.0005	.004	.0006	.005
	(.001)	(.001)	(.002)**	(.0001)***	(.001)***
Ln(polygon area)	.947	.954	.728	.046	−.993
	(.042)***	(.042)***	(.068)***	(.002)***	(.062)***
Ln(total population in 1995)	.142	.144	.036	.010	−.040
	(.024)***	(.024)***	(.034)	(.001)***	(.025)
Ln(slope)	−.052	−.053	−.009	−.003	−.029
	(.005)***	(.005)***	(.010)	(.0002)***	(.006)***
Ln(road density)	−.059	−.056	.025	−.004	−.010
	(.026)**	(.026)**	(.053)	(.001)***	(.027)
Observations	58,587	58,587	15,758	58,587	5,545
Ecoregion controls	Yes	Yes	Yes	Yes	Yes

Results: Fuzzy RDD / Instrumental Variables

- Very strong first stage: eligiblity ↑ probability of enrollment ($\tau_1 = 0.8$)

	IV Tobit				IV OLS	
	Full Estimation Sample			Restricted Sample	Deforestation (0/1)	% Deforested (If 1)
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	.584	1.293		1.038	.031	1.264
	(.280)**	(.715)*		(.609)*	(.019)*	(.680)*
Proportion treated			3.453			
			(1.870)*			
Marginality index	.521	.641	.244	−.072	.028	−.005
	(.042)***	(.106)***	(.298)	(.339)	(.003)***	(.101)
Index^2		.177	.391		.004	.162
		(.116)	(.221)*		(.004)	(.119)
Index^3		−.091	−.053		−.003	−.036
		(.035)***	(.031)*		(.001)***	(.030)
Index^4		−.010	−.037		−.0003	−.019
		(.015)	(.022)*		(.0005)	(.014)
Baseline area in forest, 2000	−.0005	−.0008	−.001	.003	.0006	.004
	(.001)	(.001)	(.001)	(.002)**	(.0001)***	(.001)***
Ln(polygon area)	.963	.990	1.075	.756	.047	−.948
	(.043)***	(.047)***	(.079)***	(.070)***	(.002)***	(.065)***
Ln(total population in 1995)	.055	−.056	−.305	−.097	.005	−.262
	(.051)	(.116)	(.245)	(.086)	(.003)	(.120)**
Ln(slope)	−.054	−.057	−.064	−.012	−.003	−.033
	(.005)***	(.006)***	(.008)***	(.010)	(.0002)***	(.007)***
Ln(road density)	−.075	−.092	−.119	.016	−.005	−.049
	(.027)***	(.033)***	(.043)***	(.054)	(.001)***	(.036)
Observations	58,587	58,587	58,587	15,758	58,587	5,545
Ecoregion controls	Yes	Yes	Yes	Yes	Yes	Yes

Mechanisms

- RD results show that CCT increases deforestation
- Why? What changes at the household level?
- Approach 1: Use experimental sample (Progresa) with household survey data
  - Progresa had randomized rollout at village level for first 3 yrs
- Approach 2: Heterogeneity by road density to study role of market access

Mechanisms: Consumption channel

- Before (1997-1998) and after (2000) data on consumption
- 506 villages, 320 treated, 186 control (not-yet-treated)
- Treatment at the village level

Difference in differences:

$$y_{it} = \gamma_0 + \gamma_1 Treat_i + \gamma_2 Post_t + \gamma_3 Treat_i \times Post_t + \epsilon_{it}$$

- y_{it} = consumption
- $Treat_i = 1$ if household i in treated locality
- cluster at locality level bc randomization was at locality level
- How do we interpret γ_3?

Hypothesis: $\gamma_3 > 0$

Mechanisms: Market channel

- Demand shock must be met by supply, which drives land use change
  - e.g. $\uparrow$ consumption of milk met by $\uparrow$ in grazing land
- But γ_3 captures partial equilibrium
  - Part of demand shock supplied locally. What about rest?
- Hypothesis: If consumption is driving deforestation
  - Then effect should be larger when infrastructure quality is low
  - Low market access: demand met by local supply
  - High market access: demand shock propagates across markets

Triple Differences:

Heterogeneity by infrastructure quality:

$$
\begin{aligned}
y_{it} = \beta_0 &+ \beta_1 Treat_i + \beta_2 Post_t + \beta_3 (Treat_i \times Post_t) \\
&+ \beta_4 Road_i + \beta_5 (Road_i \times Treat_i) + \beta_6 (Road_i \times Post_t) \\
&+ \beta_7 (Road_i \times Treat_i \times Post_t) + \epsilon_{it}
\end{aligned}
$$

- $Road_i$ = inverse road density (km of road w/n 10km of village/polygon)

Hypothesis: $\beta_7 > 0$

Results: Increase in consumption of land intensive goods

	Rooms in Home		Days Ate Beef		Days Drank Milk	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	.014	.017	.114	.118	.337	.331
	(.033)	(.035)	(.030)***	(.031)***	(.081)***	(.087)***
Treatment × inverse road density		−.034		−.070		.183
		(.148)		(.097)		(.669)
Village chosen to receive Progresa	.0001	.002	−.025	−.031	−.133	−.143
	(.037)	(.038)	(.029)	(.030)	(.111)	(.118)
Posttreatment year	.053	.049	−.137	−.138	−.655	−.664
	(.028)*	(.029)*	(.024)***	(.025)***	(.061)***	(.065)***
Inverse of road density		.266		−.156		.051
		(.169)		(.069)**		(.499)
Village × inverse road density		.043		.102		.232
		(.236)		(.140)		(.682)
Posttreatment × inverse road density		.067		.016		.155
		(.140)		(.068)		(.252)
Observations	23,318	23,318	33,128	33,128	33,128	33,128
Mean dependent	1.557		0.388		1.440	
Variable in baseline	(0.930)		(0.661)		(2.367)	

Higher beef and milk demand (land intensive products)

Demand-side impacts do not vary with market access (we did't expect it to)

Results: No increase in local production

	Number of Plots		Log (1+ Total Hectares)		Number of Cows	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	.030	.031	−.014	−.015	.092	.036
	(.039)	(.040)	(.038)	(.039)	(.057)	(.057)
Treatment × inverse road density		−.107		.142		.936
		(.210)		(.223)		(.522)*
Village chosen to receive Progresa	.014	.037	−.004	.017	−.004	.058
	(.056)	(.057)	(.040)	(.040)	(.087)	(.085)
Posttreatment year	−.094	−.077	.312	.317	−.239	−.180
	(.032)***	(.033)**	(.033)***	(.033)***	(.046)***	(.046)***
Inverse of road density		.833		.820		2.122
		(.161)***		(.227)***		(.799)***
Village × inverse road density		−.263		−.217		−.760
		(.317)		(.258)		(.872)
Posttreatment × inverse road density		−.275		−.235		−.982
		(.149)*		(.128)*		(.402)**
Observations	45,087	45,087	32,631	32,631	34,248	34,248
Mean dependent	0.824		1.724		0.604	
Variable in baseline	(0.955)		(3.535)		(2.304)	

Supply-side impacts do not vary with market access

Results: Deforestation higher in places with poor market access

	Low Density		Medium Density		High Density		Interactions	
Dependent Variable	(%) (1)	(0/1) (2)	(%) (3)	(0/1) (4)	(%) (5)	(0/1) (6)	(%) (7)	(0/1) (8)
Treated	1.619	.075	.554	.019	1.818	.023	.778	.008
	(.868)*	(.037)**	(.836)	(.030)	(1.472)	(.021)	(.600)	(.015)
Treated × low							1.041	.059
							(.582)*	(.017)***
Low road density							−.550	−.028
							(.400)	(.012)**
Observations	19,529	19,529	19,529	19,529	19,529	19,529	58,587	58,587

More deforestation in places with less market access (low road density)

Supports mechanism of increased demand for land-intensive goods

Thoughts?

- Do we believe the story about increased demand for land-intensive goods?
  - Less market access means less access to inputs
  - Underinvest in technology, leading to higher deforestation
  - Maybe this is a story about technology, not consumption
- Can you think of an alternate explanation?
- Other concerns?
- This all points to shortcomings of RCTs in general
  - we cannot answer why

Same results in Colombia! Malerba (2020)

- Question: What is the impacts of CCT on consumption, energy, and deforestation?
- Context: Familias en Accion CCT project (2001-2005)
- Design: Matched Difference in Differences
- Result: Increased beef and milk (land intensive) consumption
    - Mediated by markets
    - Negligible $\uparrow$ in deforestation (counterintuitive)
    - No impact on CO2

Research Design

- CCT program launched in 2000
  - Non-random: Municipalities selected on amenities (banks, education, health)
  - 721 eligible: slow phase-in
- Study sample: 5,477 households
- Variation: compare outcome in treated hh before/after enrollment relative to control hh
- **Identification assumptions**

	Number of municipalities enrolled (annual)
2000	2
2001	360
2002	244
2003	6
2004	0
Total	612

Results: Increased beed and milk consumption

Household DiD:

$$\Delta Y_i = \alpha + \delta T_i + X_i + \epsilon_i$$

Variables	(1) Beef, days per week	(2) Beef, days per week	(3) Milk, days per week	(4) Milk, days per week
FA	0.416**	0.879**	0.720**	1.099**
	(0.166)	(0.333)	(0.282)	(0.547)
FA *distance to closest market		−0.010*		−0.008
		(0.005)		(0.008)
Households	2268	2268	2269	2269
R-squared	0.029	0.032	0.045	0.046

- Effect muted with low market access (very weak)
- Why is sample size declining from 5,477?

Design: Environmental Impacts

- Unit of analysis: Municipality

- Design: municipalities enrolled gradually over time
  - Many were not enrolled by 2004

- Define $Treat_i = 1$ if enrolled by 2004
- $Post_t = 1$ after treatment, zero for 2000

$$y_{it} + \beta_0 + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 (Treat_i \times Post_t) + \beta_4 X_{it} + \epsilon_{it}$$

- Variation: Δ forest b/w 2000/05 for municipalities enrolled by 2004 rel. to non-enrolled

Results: No environmental impact

Variables	(1) % forest	(2) % forest, incl. ref.	(3) % forest
FA enrollment	0.510*	0.369	0.523*
	(0.271)	(0.303)	(0.293)
FA enrollment, number of years			
FA enrollment * distance to closest market			−0.000
			(0.001)
2005	−2.166***	−1.479***	−2.167***
	(0.266)	(0.309)	
Constant	22.361***	23.452***	22.358***
	(1.588)	(2.324)	(1.615)
R-squared	0.370	0.196	0.370
Observations	1440	1440	1440

- Weakly positive deforestation (counterintuitive)
- **Why?**

Take a 5 minute break

Back to the question

How does development affect the environment?

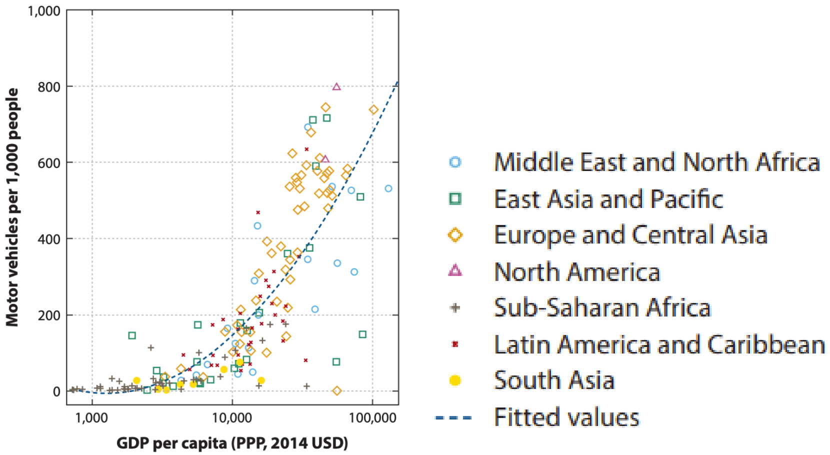
- Income
  - Land intensive consumption
  - **Energy**

- Next time: Technology and Infrastructure

Development, Energy, and the Environment

- Energy ladder: as income $\uparrow$, move from solid fuel $\rightarrow$ gas $\rightarrow$ electricity
  - Electricity may displace dirty energy (wood), which $\downarrow$ pollution
- But, electricity also increases total energy (fridge, AC)
  - ...which $\uparrow$ pollution
  - energy footprint increases with income
- Unless generated from solar, wind, etc
- People also buy cars as they get richer...

GDP and Car Ownership



Of course, these are correlations

But first, some theory

- Imagine we can exogenously vary income to poor households
- Will transfers to initially poor households yield **lumpy** investments?
  - Movements up the ladder involve big purchases
  - May explain why 1.3 billion people lack electricity (Gertler et. al, 2016)
- Credit constraints play an important role
  - Without credit markets, buy now but have nothing in next period
  - With savings, can save now, buy later (smooth consumption)
  - With borrowing, buy now, smooth consumption over time
- Lets formalize this

Model Set up

- Two periods with no discounting
- Agent i can consume two goods
  - non-durable gives per period utility: $u'(\cdot) > 0$, $u''(\cdot) < 0$
  - lumpy durable that gives static per period utility R, if owned
- Durable price = P; non-durable is numeraire
- Let Y_1, Y_2 be per period 1, 2 income
- Define $\bar{Y} = 1/2(Y_1 + Y_2)$ = average per period income

No credit constraints

Without credit constraints, if don't buy, total utility is:

$$u(\bar{Y}) + u(\bar{Y}) = 2u(\bar{Y})$$

If buy durable, spend P and spread cost equally across periods:

$$2u(\bar{Y} - \frac{P}{2}) + 2R$$

Purchase durable iff:

$$u(\bar{Y}) - u(\bar{Y} - \frac{P}{2}) \leq R$$

$u''(\cdot) < 0$ **implies acquisition increases in income**

1) With credit constraints: Buy in period 1

If don't purchase, total utility is:

$$u(Y_1) + u(Y_2)$$

If purchase in period 1:

$$u(Y_1 - P) + 2R + u(Y_2)$$

Purchase in period 1 iff:

$$\frac{u(Y_1) - u(Y_1 - P)}{2} \leq R$$

2) With credit constraints: Wait to buy in period 2

If don't purchase, total utility is:

$$u(Y_1) + u(Y_2)$$

If save in period 1, and buy in period 2:

$$2u(\bar{Y} - \frac{P}{2}) + R$$

Wait to buy in period 2 if:

$$u(Y_1) + u(Y_2) - 2u(\bar{Y} - \frac{P}{2}) \leq R$$

Intuition of Model

- $u''(\cdot) < 0$ means consumers gain from smoothing consumption
- Unconstrained household will buy in period 1
  - Use period 2 income (loan) to smooth consumption
- Credit constrained households cannot do this
  - If buy now, magnify consumption inequality across periods
  - Or, wait and buy in period 2, but delay utility gain
- Higher income in any period leads savers to buy more
- But, first period buyers respond only to first period income (lemma 1)
- But increase in *cumulative* income increases buying through delay/saving (lemma 2)

Testable Predictions

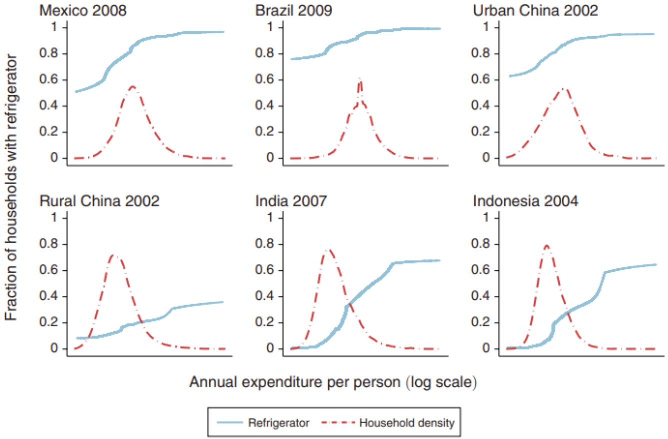
- Prediction 0: Acquisition increasing in income
- Prediction 1: S-shaped curve in acquisition

Gertler et al. (2016): Demand for energy-using assets among middle class

- **Question:** How does income affect energy demand?
- **Motivation:** 1.3 billion people live without electricity
- Should we expect linear climb up the energy ladder as incomes rise?
  - what about credit markets?

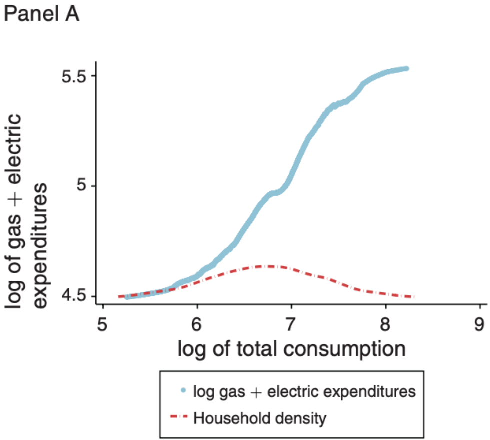
	Electricity access (percent of population)	Refrigerators (share of households)
Brazil	98.7	0.93
China	99.7	0.69
India	75.0	0.13
Indonesia	73.0	0.17
Mexico	97.9	0.83
Sub-Saharan Africa	32.5	0.11
Total	70.8	0.38
United States	100.0	0.99

Non-linear relationship between income and fridges



- Are fridges representative of energy-intensive assets?
- Can aggregating non-linearities $\rightarrow$ linearity?

S-shape appears robust



Empirical Setting: Mexico Oportunidades (again)

- 320 randomly selected communities given early treatment (April 1998)
  - 52% were "eligible" for cash transfer
- 186 randomly selected communities given late treatment (October 1999)
- This paper: N=506 communities, 10,000 households surveyed (1997-2007)
  - Surveys done in 1998, 1999, 2000, 2003, and 2007
- Compare early and late treatment households (**not treatment vs. control**)
  - Similar b/c both groups selected on "vulnerability" characteristics
  - Strongly balanced on covariates
  - Assumes timing is random

Empirical Design

- **Prediction 0:** Probability of asset purchase $\uparrow$ in income

$$h(a_{it}) = Pr(a_{it} = 1 | a_{it-1} = 0) = \gamma_0 + \alpha_1 CI_{it} + \beta X_i + \beta_t F_i + R_{rt} + \epsilon_{it}$$

- where h(a) is the prob. that i buys a in time t conditional on not having it in $t-1$.
- CI_{it} is cumulative income
- X_i includes baseline household controls
- F_i are family structure variables interacted w/ survey round
- R_{rt} are state-year fixed effects

Hypothesis: $\alpha_1 > 0$

Empirical Design

- **Prediction 1:** S-shaped curve in acquisition

$$h(a_{it}) = Pr(a_{it} = 1 | a_{it-1} = 0) = \gamma_0 + \alpha_1 CI_{it} + \alpha_2 CI_{it}^2 + \beta X_i + R_{rt} + \epsilon_{it}$$

Hypothesis: $\alpha_2 \lessgtr 0$?

Measurement issues

- Cumulative Income is sum of wage, farm, business income, CCT transfers
  - Reporting biases in wage data
  - Income reported in survey year, not cumulative
  - Authors interpolate across years, adding noise
- Double counting problem if households invest CCT money into business
- **Solution:** use potential CCT **transfers** to instrument cumulative **income**

Identifying Variation

Table 2—Oportunidades Bimonthly Support Levels in 2003 (*Pesos*)

Basic Support	155	
Educational scholarship		
Grade	Boys	Girls
Third	105	105
Fourth	120	120
Fifth	155	155
Sixth	205	205
Seventh	300	315
Eighth	315	350
Ninth	335	385
Tenth	505	580
Eleventh	545	620
Twelfth	575	655

Use maximum cumulative transfer as **instrument** for actual (cumulative) transfer

Plus variation from early/late enrollment

Identification Assumptions

- What is the exclusion restriction?
- Is it reasonable?

Results: Prediction 0

TABLE 4—BASIC RESULTS: REFRIGERATOR (*Income Effects*)

	Discrete time hazard		Household FE	Discrete time hazard		Household FE
	OLS	IV	IV	OLS	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative Transfers	0.018***	0.020***	0.047***			
	[0.005]	[0.007]	[0.008]			
Cumulative Income				0.003***	0.016***	0.034***
				[0.001]	[0.005]	[0.007]
Observations	30,414	30,414	30,258	30,414	30,414	30,258
R^2	0.103			0.104		
Kleibergen-Paap Wald F-Stat on excluded variables		2,503	2,060		92	108
Number of households			6,655			6,655

- For every 10,000 pesos, probability of buy fridge ↑ by 4.7 pp.

Results: Prediction 1

TABLE 5—BASIC RESULTS: REFRIGERATOR (*Nonlinear Income Effects*)

	Discrete time hazard		Household FE
	OLS	IV	IV
	(1)	(2)	(3)
Cumulative Income	−0.0007	−0.0059	0.0132
	[0.0011]	[0.0108]	[0.0132]
Cumulative Income ^{2}	0.0001***	0.0009**	0.0008**
	[< 0.0001]	[0.0004]	[0.0004]
Observations	30,414	30,414	30,258
R^2	0.105		
Kleibergen-Paap Wald F-Stat on excluded variables		22	23
Number of households			6,655

- $\alpha_2 > 0$ implies convex relationship b/w cumulative income and asset purchase
- Since all households are poor, don't expect any to be past second inflexion point

Thoughts on paper?

Research gaps

- Does earned income have different effect than transfers?
- Can we trace out the full S-curve?
- Is there an S-curve for other technologies?
- Which externalities have highest income elasticities?

Next week

- How does development affect the environment?
  - **Technology and Infrastructure**

Lecture 3

How does development affect the environment?

Housekeeping

- Research proposal first draft: Oct 3rd
  - If you find idea w/n 10 days, gives 2 weeks to write
- Class presentations: separate Canvas upload each week
- Replication project?

Recap: Alix-Garcia et al. (2013)

- Do we believe the story about increased demand for land-intensive goods?
  - Less market access means less access to inputs
  - Underinvest in technology, leading to higher deforestation
  - Maybe this is a story about technology, not consumption
- Can you think of an alternate explanation?
  - Talk to your neighbor for five minutes
- Other concerns?

Recap

- Higher income leads to land-intensive **consumption**
  - Beef, milk, etc.
  - This requires clearing land for agriculture
  - Effect depends on market access

- Higher income leads to energy-intensive **consumption**
  - But the relationship is non-linear
  - Role of credit constraints

Today

- Development increases access to **productive** capital
  - How do we think about environmental implications?
- Case I: Development increases agricultural productivity
  - Abman et al. (2023)
- Case II: Development spurs infrastructure development
  - Garg et al. (2023)

Technology and infrastructure

- As economies develop, technology and infrastructure improve
- Productivity boost means firms produce at lower cost
  - Outward shift in supply curve $\rightarrow$ bad for environment
  - But, technological progress reduces resource intensity $\rightarrow$ good for environment
- Most active areas of research:
  - agricultural productivity
  - infrastructure

Agricultural productivity and the environment

- **Boserup's hypothesis** : extensive margin
  - improvements in land productivity $\uparrow$ ag land value
  - leads to pressure on forests to clear land for ag
- **Borluag's hypothesis**: intensive margin
  - *Under factor market constraints*, $\rightarrow$ productivity spurs intensification
  - "spares' ' land for nature and puts less pressure on forests
- **Jevon's paradox**: general equilibrium
  - technology for increasing yield also lowers MC of producing given amount
  - Ag supply curve shifts out and may dominate gains in yield
  - Leads to net increase in land demand

Takeaway: this is an empirical question!

Existing Work

- Theory: (Angelsen and Kaimowitz,2001; Balsdon, 2007; Goldstein et al., 2012; Green et al., 2005; Phalan et al.,2016; Takasaki,2006).
  - Most models yield ambiguity
- Most work on agricultural productivity and deforestation
  - Intensive margin dominates: Abman et al. (2024), Abman and Carney (2020a), Abman and Carney (2020b), Assuncao et al. (2023)
- New work: market access and deforestation
  - Abman and Lundberg (2024)
- Open Topics:
  - Effect of agriculture on deforestation in the first place? (Madhok et al. 2024)
  - Role of factor markets? Labor vs. input markets
  - Mechanisms behind market access effects
  - Agriculture, migration and deforestation

Conceptual framework

- Continuum of agents with heterogeneous outside options
- Farmers allocate labor b/w farming on existing land, or cultivate new land
- Let γ denote household labor allocated to new land
- Production function for new and existing land are $f(\cdot)$ and $g(\cdot)$
- Exogenous productivity shock, Ω, improves productivity of existing land more

$$\frac{\partial g(l)}{\partial \Omega} \geq \frac{\partial f(l)}{\partial \Omega} \geq 0 \qquad \forall l$$

Conceptual framework

- Farmers are factor market constrained (Conning and Udry, 2007)
  - Cannot hire labor from market
- Total deforestation given by new land cleared for agriculture:

$$D^* = n^* \cdot \gamma^*$$

- Differentiating with respect to Ω gives:

$$\frac{\partial D^*}{\partial \Omega} = \underbrace{n^* \frac{\partial \gamma^*}{\partial \Omega}}_{\lesseqgtr 0} + \underbrace{\gamma^* \frac{\partial n^*}{\partial \Omega}}_{\lesseqgtr 0}$$

- What does each term represent and what is the sign?

Abman et al. (2023): Agricultural Productivity and Deforestation

- **Goal:** To test $\frac{\partial D^*}{\partial \Omega}$, we need random variation in Ω
- **Question:** what is the impact of ag productivity on deforestation?
- **Context:** Large scale ag extension program in Uganda (2008-2013)
- **Design:** Leverage Spatial discontinuity in village eligibility
- **Results:** improvement in agricultural productivity; reduction in deforestation

Background: Why Uganda?

- Forest cover shrunk from 24% to 9% from 1990-2015
- Most forest land is **privately owned**
  - Land owners allowed to covert forest (Land Act, 2010)
  - Cannot use state powers to limit deforestation
- Need an indirect mechanism to incentivize conservation

Background: BRAC Extension Program

- Launched in 2008
- Aims to extend modern cultivation techniques to smallholders
- Two step treatment:
  - Training: "model farmers' ' trained in modern cropping, irrigation, weeding, pest control. Then, set up demo plot and train others
  - Extension: Promoters selected to sell subsidized HYV seeds in their villages
- **Key feature: Program limited to villages within 6km of each BRAC branch**

Data

- Deforestation from GFC (Hansen et al., 2013)
  - Pro: high resolution (30m)
  - Con: 1) only captures loss not gain, 2) everything relative to 2000
- Lat/lon of villages within 12km of BRAC office (N=807 villages)
- Attribute forest pixel to village if w/n 400m of village centroid
  - Rationale: median household distance to village center = 400m
  - Problems?
- BRAC agricultural survey on 7,781 households (451 villages) in 2011
  - Detailed cropping practices from last two seasons

Empirical Design: Spatial Regression Discontinuity

- Estimate IIT since we do not know true treatment designation
- Local linear regression to estimate left and right side of discontinuity

$$\beta = lim_{z\uparrow 0}E[Y|z_i = z] - lim_{z\downarrow 0}E[Y|z_i = z]$$

- where running variable, z_i, is distance of village from 6km cutoff
  - $z \leq 0$ means village had access to extension
- For estimation, use 2km bandwidth
  - Higher regression weights on points nearer to cutoff (triangle kernel)

Aside: RD with local linear regressions

- Left side of cutoff

$$Y_i = \alpha_L + \beta_L(z_i - c), \qquad \text{for } z_i \leq \text{c}$$

- Right side of cutoff

$$Y_i = \alpha_R + \beta_R(z_i - c), \qquad \text{for } z_i \geq \text{c}$$

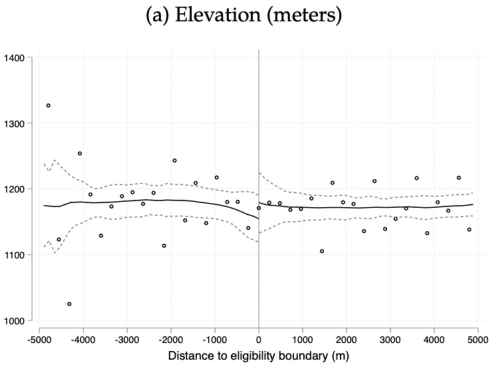
- Treatment effect at threshold:

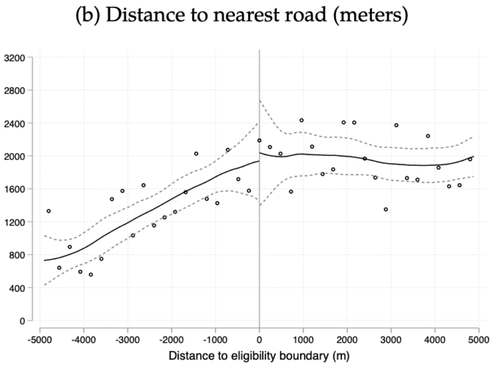
$$\beta = \alpha_R - \alpha_L$$

since α are the intercepts on each side of cutoff

Identification Assumptions?

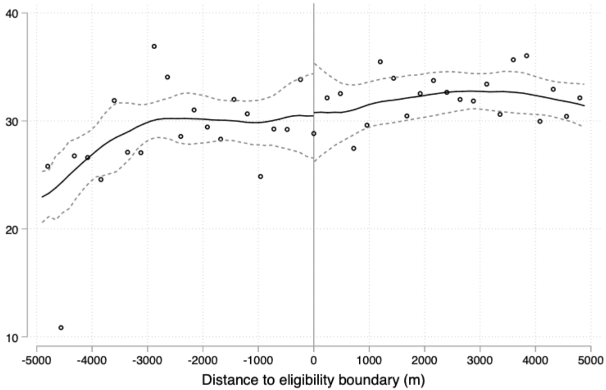
Identification Assumptions



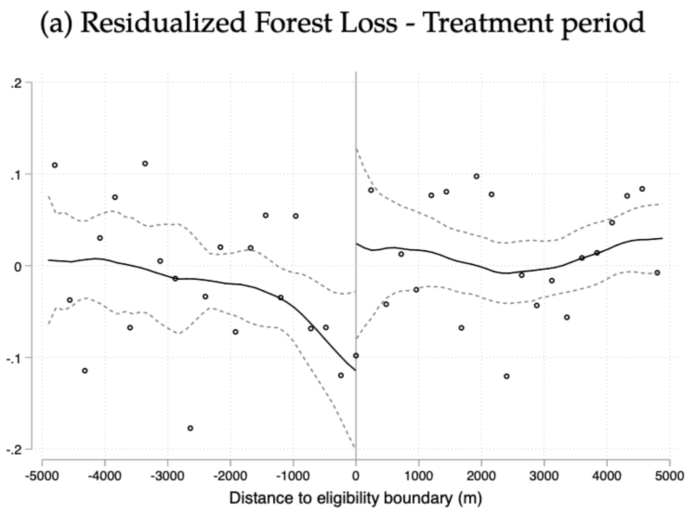


Identification Assumptions

(c) Baseline treecover



Results: 14% less deforestation in treated villages



Mechanisms: Intensification

Outcome	(1) Manure Use	(2) Intercrop	(3) Crop Rotation	(4) Irrigation	(5) Weeding	(6) Fertilizer Use	(7) HYV Seeds	(8) Perennial crops
Program Eligible	0.0977***	0.0590*	0.0737***	0.0326***	0.0644**	-0.0171	-0.0440	-0.0264
	(0.0249)	(0.0308)	(0.0251)	(0.00829)	(0.0310)	(0.0161)	(0.0323)	(0.0351)
Obs	2912	2912	2912	2912	2912	2912	2912	2912
Control mean	0.0731	0.796	0.797	0.0266	0.693	0.0725	0.356	0.351

- Techniques to address nutrient depletion **on existing land**
- No evidence of increased fertilizer, pesticides, or HYV seed adoption

Intensification vs. Extensification

Outcome	(1) Revenue per acre	(2) IHS Cultiv ag area
Program Eligible	0.325*	0.0262
	(0.189)	(0.0514)
Obs	2843	2907
Control mean	11.59	1.414

- Revenue/acre is a proxy for yields
- No extensification → binding constraint on land clearing

Discussion

- Empirical issues
- Alternative explanations
- External validity

Back to the Question

How does development affect the environment?

- Income
  - Land intensive consumption
  - Energy

- Access to capital
  - Technology
  - **Infrastructure**

Infrastructure and the Environment

- Roads can increase deforestation by:
  - $\downarrow$ MC (transport costs) of forest products, increasing supply
  - $\uparrow$ market access for agricultural products, causing extensification
  - $\uparrow$ land value, leading to agglomeration effects

- Roads can increase forest cover by:
  - $\uparrow$ access to forest product substitutes (i.e. clean fuel vs. firewood)
  - $\uparrow$ access to external labor/capital markets, lowering need for land-clearing

- Net effect is ambiguous

What about Infrastructure and Air Pollution?

- Lets build a conceptual framework
- Talk to your neighbor for 10 mins

Garg et al. (2023): Rural Roads, Labor Exits, and Crop Fires

- **Question:** what is the impact of roads on air quality?
  - what is the role of labor reallocation as a mechanism?
- **Context:** India builds last-mile roads using population cutoff
- **Design:** Regression discontinuity with crop fires as outcome
- **Results:** Roads $\uparrow$ crop fires and PM2.5 levels in a village
- **Mechanism:** Roads $\rightarrow$ labor exit $\rightarrow$ higher wages $\rightarrow$ labor-saving technology

Background

- Labor is misallocated in agriculture in developing countries
  - MP_{labor} lower in agriculture than other sectors
- Sources of misallocation is a big question in dev econ
- Many policies aimed at reducing misallocation
  - e.g. road-building (improve market access)
- Do these have externalities?
  - **innovation**: where do the externalities come from?

Context: India

- 14 of the 20 most polluted cities in the world are in India
- Important pollution source: crop burning
- Pros:
  - Clear rice straw residue & undergrowth b/w cropping seasons
  - Cheaper than labor/capital alternative
- Cons:
  - Carries civil/criminal penalty (weakly enforced)
  - Reduces soil nutrients and decreases ag productivity
  - Damages human health
- Roads may facilitate agricultural labor loss
  - Will farmers turn to crop burning, despite the costs?

Research Design

- PMGSY aims to provide all-weather roads to all unconnected villages in India
- Prioritize larger villages first
  - Connect villages with population > 1000 by 2003
  - Connect villages with population > 500 by 2007
  - Connect villages with population > 250 afterwards
- State-by-state basis
  - If connected all large villages, can proceed to smaller ones immediately

Data

- Roads: SHRUG database
  - N=11,151 villages w/o paved road at baseline (2001)
  - Main variable: road construction date in village
- Fires: NASA EODIS
  - # fires w/n 10km of village polygon
- Pollution: Van donkelaar et al. (2016) measure PM2.5 at 10km resolution
  - Black Carbon and Organic Carbon from biomass/non-biomass sources
- Infant mortality: NFHS-IV (2015-16)
  - child-birthyear panel from 2001-2013
  - Identify survey clusters (villages) within 50km of each PMGSY village

Empirical Strategy: Fuzzy RDD

First Stage:

$$
\begin{aligned}
Road_{vdst} &= \gamma_0 + \gamma_1 1(pop_{vds} \geq T) + \gamma_2(pop_{vds} - T) \\
&\quad + \gamma_3(pop_{vds} - T) \times 1(pop_{vds} \geq T) + \theta X_{vds} + \mu_{d,h} + \rho_t + \epsilon_{vdst}
\end{aligned}
$$

Second Stage:

$$
\begin{aligned}
Y_{vdst} &= \beta_0 + \beta_1 Roads_{vdst} + \beta_2(pop_{vds} - T) \\
&\quad + \beta_3(pop_{vds} - T) \times 1(pop_{vds} \geq T) + \delta X_{vds} + \eta_{d,h} + \omega_t + \epsilon_{vdst}
\end{aligned}
$$

- $Road_{vdst} = 1$ if village v received road by 2013
- Y_{vdst} is outcome (fires, pollution, mortality)
- pop_{vds} is village population in 2001
- T is treatment threshold (500 or 1000, depending on state)

Covariate Balance

Table 1: Main analysis sample: summary statistics, and balance and falsification tests

Variable	Full sample	Below threshold	Over threshold	Difference of means	p-value on difference	RD estimate	p-value on RD estimate
Primary school	0.959	0.955	0.964	0.01	0.02	-0.018	0.59
Medical center	0.166	0.155	0.177	0.02	0.00	-0.097	0.14
Electrified	0.430	0.414	0.447	0.03	0.00	-0.014	0.87
Distance from nearest town (km)	26.490	26.379	26.613	0.23	0.58	-3.426	0.34
Land irrigated (share)	0.281	0.276	0.287	0.01	0.05	-0.025	0.59
Ln land area	5.151	5.093	5.215	0.12	0.00	-0.103	0.33
Literate (share)	0.457	0.454	0.461	0.01	0.01	-0.012	0.62
Scheduled caste (share)	0.143	0.141	0.145	0.00	0.24	-0.020	0.52
Land ownership (share)	0.733	0.733	0.732	-0.00	0.75	0.013	0.72
Subsistence ag (share)	0.435	0.438	0.432	-0.01	0.25	0.024	0.58
HH income > INR 250 (share)	0.754	0.752	0.757	0.00	0.37	-0.023	0.63
Outcomes at baseline:							
Annual fires (count)	0.685	0.699	0.670	-0.03	0.40	0.194	0.50
Annual PM2.5 ($\mu g/m^3$)	35.886	35.821	35.958	0.14	0.41	-0.029	0.94
Downwind infant mortality (0/1)	0.067	0.068	0.067	-0.00	0.28	0.002	0.63
Other directions infant mortality (0/1)	0.072	0.072	0.071	-0.00	0.23	-0.002	0.72
Downwind PM 2.5 ($\mu g/m^3$)	36.038	35.879	36.252	0.37	0.04	0.431	0.36
Other directions PM 2.5 ($\mu g/m^3$)	39.254	38.646	40.237	1.59	0.00	0.901	0.15
N	11151	5859	5292				

Results

	New road	Annual fire activity		Annual average PM 2.5	
	(1)	(2)	(3)	(4)	(5)
	1^{st} *stage*	RF	IV	RF	IV
Above threshold pop.	0.230***	0.366**		0.106**	
	(0.017)	(0.152)		(0.045)	
Road built			1.567**		0.470**
			(0.689)		(0.207)
N	133,788	133,788	133,788	133,788	133,788
Control group mean	0.10	2.77	2.77	42.69	42.69

- Note: optimal bandwidth = +/- 84 around population threshold
- Road access increases annual crop fires by 60%
- Road access increase PM2.5 level by 0.5 $\mu g/m^3$ (1.1%)

Does pollution result also capture increased driving on new roads?

Panel A: Winter harvest and post-harvest months							
	Fires	All sources		Biomass burning		Other sources	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		BC	OC	BC	OC	BC	OC
Road built	1.051***	0.285**	3.146*	0.320***	2.791**	-0.010	-0.010
	(0.391)	(0.130)	(1.777)	(0.103)	(1.180)	(0.009)	(0.009)
N	133,788	133,788	133,788	133,788	133,788	133,788	133,788
Control group mean	1.45	30.47	120.81	1.17	13.46	29.30	107.35

Panel B: Rest of the year							
	Fires	All sources		Biomass burning		Other sources	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		BC	OC	BC	OC	BC	OC
Road built	0.276	-0.219	-0.115	0.192	0.886	-0.010	-0.010
	(0.438)	(0.351)	(3.120)	(0.246)	(2.680)	(0.009)	(0.009)
N	133,788	133,788	133,788	133,788	133,788	133,788	133,788
Control group mean	1.27	30.83	124.70	1.53	17.34	29.31	107.36

Mechanism: Roads increase ag wage rate

- Data: REDS Survey (1999 and 2006)
- Design: Diff-in-Diff (221 villages, 100 districts, 17 states)
  - Treat = 1 if v received road b/w 1999-2006; Post=1 in 2006

$$Y_{vt} = \alpha + \beta_1(Treat_v \times Post_t) + \gamma_v + \theta_t + \epsilon_{vt}$$

	Wage rate			Log wage		
	(1) Male	(2) Female	(3) Average	(4) Male	(5) Female	(6) Average
Treat X Post	0.722*	0.918***	0.856**	0.012*	0.020**	0.016**
	(0.420)	(0.349)	(0.352)	(0.007)	(0.008)	(0.007)
N	442	442	442	442	442	442
Control group mean	62.99	47.65	55.25	62.99	47.65	55.25
R^2	0.76	0.78	0.77	0.75	0.76	0.76

Mechanism: Roads induce labor exit

- Back to RDD design
- Data from Socioeconomic and Caste Census (2013), provided in SHRUG
- Outcome: ag. labor share in districts with high/low ag wage

	Share of labor in agriculture		Share of non-agricultural manual labor	
	(1) High rel. ag. wage	(2) Low rel. ag. wage	(3) High rel. ag. wage	(4) Low rel. ag. wage
Road built	-0.029	-0.245***	0.030	0.210**
	(0.046)	(0.093)	(0.046)	(0.092)
N	5,402	5,483	5,402	5,483
Control group mean	0.49	0.46	0.45	0.46

- More labor exit in districts with lower baseline ag wages

Mechanism: places with more labor exit burn more biomass

	High rel. ag. wage		Low rel. ag. wage	
	(1)	(2)	(3)	(4)
	Fires	PM 2.5	Fires	PM 2.5
Road built	0.778	0.378*	3.195**	0.617*
	(0.721)	(0.220)	(1.439)	(0.367)
N	62,880	62,880	67,740	67,740
Control group mean	2.68	45.38	2.88	40.09

- Fire/pollution impact greater in places with lower baseline ag wage
- Road generates more labor exit in these places

Discussion

- Do you believe the story?
- Alternative explanations

Next Time

- Presentations
- How does the environment affect development?
  - **Health**
  - Productivity