

# Environmental and Development Economics

## Week 3 - How does the environment affect development?

Raahil Madhok  
UMN Applied Economics

2024-08-30

# Lecture 4

How does the environment affect economic development?

# Housekeeping

- ▶ State things that are due.

# Today

- ▶ **Guiding question:** how does the environment affect economic development?
- ▶ Today's focus: health effects
- ▶ Next time: productivity and direct income effects

## Recall our model of environment and development

$$MWTP_e = \frac{\lambda_e}{\lambda_y} = \frac{1}{\lambda_y} \left( \frac{\partial u}{\partial e} + \frac{\partial u}{\partial h} \frac{\partial h}{\partial e} \right) + \frac{\partial \Delta y}{\partial e} + \frac{\partial \Delta y}{\partial h} \frac{\partial h}{\partial e}$$

- ▶ Most research in environment/development economics seeks to identify  $\frac{\partial h}{\partial e}$ 
  - ▶ At this point,  $\frac{\partial h}{\partial e} > 0$  is unambiguous
  - ▶ Focus is on **magnitudes**
- ▶ What about indirect effects via productivity and health ( $\frac{\partial \Delta y}{\partial h} \frac{\partial h}{\partial e}$ )?
  - ▶ This is the research frontier
  - ▶ Next lecture

## Lets formalize our intuition

- ▶ Weather can harm health directly (physiological stress)
  - ▶ Or indirectly, e.g. by lowering incomes of those in weather-dependent (outdoor) jobs
- ▶ There is a market for defensive goods (e.g. medicine, AC, etc.) which  $\uparrow P(\text{survival})$
- ▶ But in developing countries, credit constraints restrict consumption smoothing
- ▶ Lets incorporate this into a model and study the comparative statics
  - ▶ Goal: explain how weather affects health, considering financial constraints
  - ▶ Health is endogenously determined by consumption and health investments

## Setup

- ▶ **Agents:** value consumption  $c_{it}$  and health  $h_{it}$ .
- ▶ **Objective:** Maximize lifetime utility

$$\mathbb{E}_0 \left[ \sum_{t=0}^{\infty} \beta^t u_i(c_{it}, h_{it}) \right]$$

- ▶ **Health Status:** Determined by health investments ( $q_{it}$ ) + exogenous weather ( $z_{it}$ )

$$h_{it} = h_i(\underbrace{q_{it}}_{(+)}, \underbrace{z_{it}}_{(-)})$$

# Setup

- ▶ **Production:**  $c_{it}$  produced by  $f(\cdot)$ , and  $q_{it}$  traded at price  $P_t$ 
  - ▶ Agent endowed with labor  $a_{it}$ , sold at wage  $\omega_t$
  - ▶ Labor can be affected by weather ( $z_{it}$ )

$$a_{it} = a_i(z_{it})$$

- ▶ **Financial Constraints:** Savings ( $x_{it}$ ) and borrowing are limited by:
  - ▶ Minimum savings:  $x_{it} \geq -b_i(r_t)$
  - ▶ Maximum savings:  $x_{it} \leq b_i(r_t)$
- ▶ **Budget Constraint:** Income is used for consumption and health investments:

$$x_{i,t+1} \leq r_t x_{it} + \omega_t a_i(z_{it}) - P_t q_{it} - c_{it}$$

## Setup

- ▶ **Market Equilibrium:** set optimal demand for labor/capital = aggregate supply
  - ▶ where  $K_t = \int x_{it} di$
  - ▶ wages  $w_t$  and interest rates  $r_t$  are determined in equilibrium as:

$$w_t = w(K_t), \quad r_t = r(K_t)$$

# Mechanisms relating weather to health

- ▶ **Direct health effect:** random shock  $z_{it}$  enters health directly:  $h_{it} = h_i(q_{it}, z_{it})$ 
  - ▶ Physiological channel: extreme temperature stresses cardiovascular/respiratory system
  - ▶ Impact can depend on region, or proximity to water (where pathogens thrive in hot weather), denoted by agent-specific function  $h_i(\cdot)$
- ▶ **Indirect health effect:** through productivity  $a_i(z_{it})$ 
  - ▶ Likely in rural areas, where ag workers are outside
- ▶ **Mitigation:** agent minimizes neg. utility shock from  $z_{it}$  by purchasing  $q_{it}$ .
  - ▶ This could also be opp cost of working inside (at lower  $w_{it}$ ) to avoid outdoors
  - ▶ Effect of mitigation limited by  $\frac{\partial h_i(q_{it}, z_{it})}{\partial q_{it}}$
- ▶ **Smoothing:** agent can also mitigate by reallocating resources b/w periods
  - ▶ Agents with near-binding borrowing constraints experience worse shocks

# Comparative Statics

- ▶ **Object of interest:** aggregate correlation b/w health  $h_{it}$  and weather  $z_{it}$ :

$$H_t = \left| \frac{\int h_{it} z_{it}^{(T)} di}{\int h_{it} di} \right|$$

- ▶ **Rural Areas:** Consider a set of changes that:
  - ▶ increases responsiveness of  $a_i(z_{it})$  to  $z_{it}^T$ .
  - ▶ raises prices for health goods  $P_t$  above that of urban
  - ▶ lower incomes by reducing  $a_i(z_{it})$  for any  $z_{it}$
  - ▶ Any one of these leads to increase in  $H_t$

**Prediction:** rural areas will have stronger correlation b/w weather and health

# Effect of Bank Access

- ▶ **Relaxing financial constraints** increases either or both of  $(-b_i(r_t), b_i(r_t))$ 
  - ▶ Partial credit modeled as increase in  $b_i(r_t)$  only (i.e. can save more)
  - ▶ Allows agents to smooth consumption; draw down savings to mitigate weather shock
  - ▶ Reduces the correlation between weather shocks  $z_{it}$  and health outcomes  $h_{it}$ .
- ▶ **Prediction:** Prediction:  $H_t \downarrow$  as  $b_i(r_t) \uparrow$ 
  - ▶ Areas with better financial access have smaller effect of  $z_{it}$  on health

# What is the effect on weather on health?

- ▶ What is the ideal experiment?
- ▶ What variation in environmental quality can approximate the experiment?
  - ▶ How would you measure environmental quality?
  - ▶ What might this measure miss?
  - ▶ What variation isolates a causal relationship?
  - ▶ What does that variation miss?

# Climate change versus weather

- ▶ Climate change is long term, weather is short term
- ▶ Studying impacts of weather is easier:
  - ▶ Usually use high dimensional fixed effects
  - ▶ E.g. with district and state-year FEs, exploit monthly deviations from average
- ▶ Studying impacts of climate is harder:
  - ▶ There is no good counterfactual for climate change
  - ▶ New method: long differences (Burke and Emerick, 2016)

## Deschenes et al. (2013): Weather and Death in India

- ▶ **Research question:** how do temperature shocks affect mortality?
  - ▶ How does that different between urban/rural areas?
  - ▶ What is the role of income as an intermediary?
- ▶ **Setting:** Indian districts from 1957-2000
- ▶ **Data:** daily gridded weather + annual infant mortality
- ▶ **Design:** panel fixed effects
  - ▶ Difference-in-differences with rural bank expansion
- ▶ **Result:** 1) heat ↑ mortality, 2) driven by rural populations, 3) credit access mitigates effect by facilitating consumption smoothing

# Empirical Implications

- ▶ Remember permanent income hypothesis?
  - ▶ Income shocks in  $t = 0$  caused by weather shock only affects contemporaneous survival via extent to which income shock reduces NPV lifetime income
- ▶ From model: if two agents face same shock at  $t = 0$ , and one has income stream that depends on weather
  - ▶ Then impact of shock on his death is stronger
- ▶ What if income shocks are at high frequency?
  - ▶ Will weather shock be enough to affect remaining lifetime income (and survival)?
  - ▶ This is why studying climate vs. weather is important

# Data

## Rural

- ▶ Yields: annual, district level output/cultivated area (27 crops) (World Bank)
- ▶ Ag price index: mean crop price weighted by value in district (World Bank)
- ▶ Ag wages: daily wage rate (World Bank)

## Urban

- ▶ Manufacturing productivity: output per worker (ASI)
- ▶ Urban CPI: NSS survey
- ▶ Manufacturing wages: ASI (nominal wage / CPI)

# Empirical Strategy

Flexible specification to model effect of daily temperature:

$$Y_{dt} = \sum_{j=1}^{11} \theta_j TMEAN_{dtj} + \sum_k \delta_k \mathbf{1}\{RAIN_{dt} \text{ in tercile } k\} + \alpha_d + \gamma_t + \lambda_r^1 t + \lambda_r^2 t^2 + \epsilon_{dt}$$

- ▶  $r$  is a climatic region
- ▶  $TMEAN$  is days in year  $t$  where:
  - ▶ daily mean temp  $< 70^\circ, \geq 97^\circ$ , or in the nine  $3^\circ$  bins in between
- ▶  $\theta_j$ 's capture non-linearities, despite having annual unit of analysis
- ▶  $RAIN$  accounts for correlation b/w temperature and rainfall

# Identifying Assumptions?

- ▶ Impact governed by daily mean alone
- ▶ Impact of days' mean temperature on annual mortality is constant within 3° intervals
  - ▶ This is reasonable
- ▶ By using number of days in each bin, assume that the sequence of hot and cold days is irrelevant.
- ▶ Annual deviation from district average is random
- ▶ Anything else? Reverse causality?

## Secondary Approach

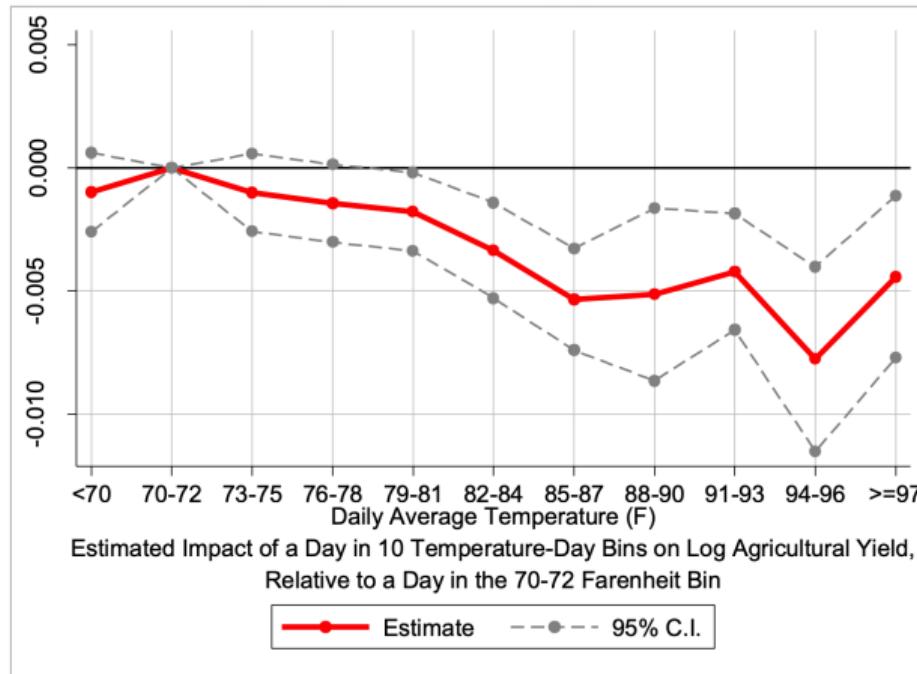
- ▶ consolidates same information
- ▶ workhorse model nowadays because more parsimonious

$$Y_{dt} = \beta CDD80_{dt} + \sum_{k=1}^3 + \sum_k \delta_k 1\{RAIN_{dt} \text{ in tercile } k\} + \alpha_d + \gamma_t + \lambda_r^1 t + \lambda_r^2 t^2 + \epsilon_{dt}$$

- ▶ where CDD80 is cumulative degree days in district that exceeded 80°.
- ▶ All bins  $j < 80^\circ$  are zero
  - ▶ The three above 80° restricted to be linearly increasing in average temp
- ▶ Assumption: days when mean daily temp below 80 irrelevant for determining Y

# Results: Yields

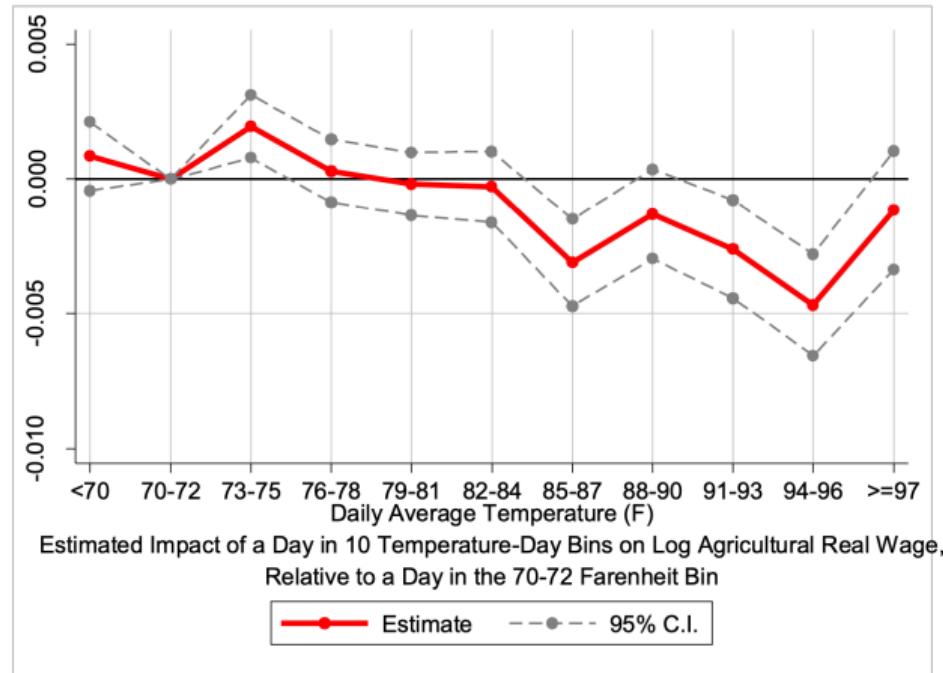
Figure 3: The Effect of Daily Average Temperatures on Log Agricultural Yields.



A single day with temperature over 85° reduces yields by 0.5%

# Results: Wages

Figure 4: The Effect of Daily Average Temperatures on Log Real Agricultural Wages.



A single day with temperature over 85° reduces ag wages by 0.25-0.5%

# Results: Urban vs. Rural

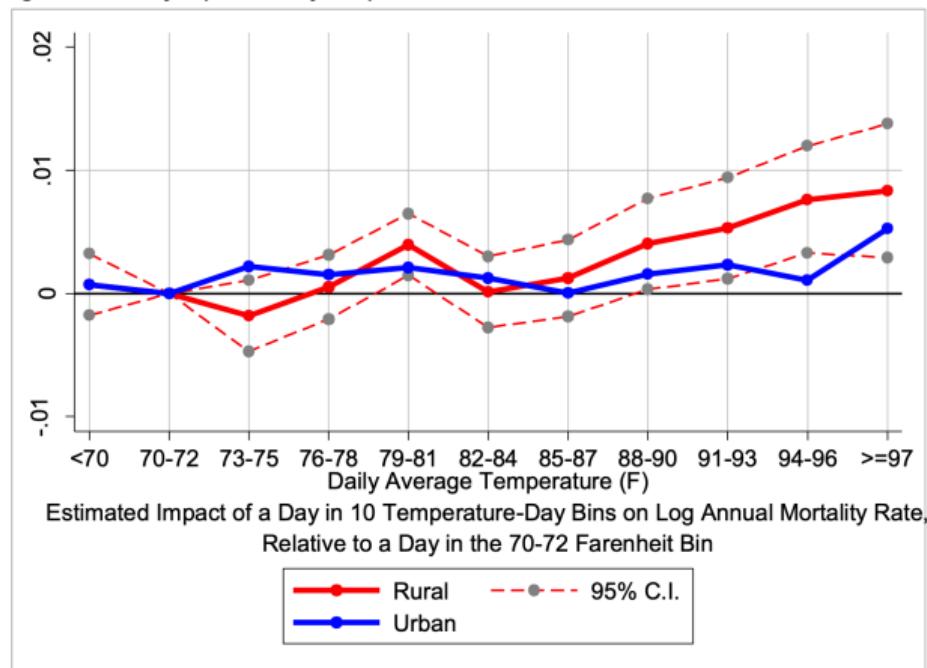
**Table 2: Weather and Incomes - Rural-Urban Differences, Exposure by Calendar Year**

Dependent Variable:	Rural		Urban	
	Log (Productivity)	Log (Real Wages)	Log (Productivity)	Log (Real Wages)
	(1)	(2)	(3)	(4)
<b>A. Temperature (degree-days over 90F)/10</b>	-0.0023*	-0.0023***	0.0003	0.0012
(std error)	(0.0009)	(0.0006)	(0.0033)	(0.0036)
[Effect of 1 std dev in CDD90]	[-0.023]	[-0.022]	[0.002]	[0.009]
<b>B. Temperature (degree-days over 80F)/10</b>	-0.0034***	-0.0026***	-0.0006	0.0006
(std error)	(0.0006)	(0.0004)	(0.0013)	(0.0013)
[Effect of 1 std dev in CDD80]	[-0.126]	[-0.098]	[-0.021]	[0.021]
<b>C. Temperature (degree-days over 70F)/10</b>	-0.0023***	-0.0018***	-0.0010	0.0005
(std error)	(0.0004)	(0.0003)	(0.0008)	(0.0008)
[Effect of 1 std dev in CDD70]	[-0.132]	[-0.101]	[-0.053]	[0.023]
Observations	8,304	8,304	512	592

What is happening? Direct (agonomical) effect should not vary

# Results: Mortality

Figure 6: Mortality Impact of Daily Temperature in Rural and Urban India.



Rural: years with more days with exceeding 87° have 0.5% higher mortality

## Bank Access

- ▶ If income → consumption smoothing, then it should mute relationship
  - ▶ Mechanism: lifetime utility maximized when  $u'(c)$  equated across  $t$  (Euler)
- ▶ Test: Between 1977-1990, central bank extends banking to the poor
  - ▶ Existing banks required to open four branches in unbanked areas
- ▶ Method: compare weather-death relationship in banked/unbanked areas
  - ▶ Difference in differences
  - ▶ Identification problem?
- ▶ Instrument: SEE BURGESS AND PANDE (2005)

# Equation

- ▶ See BURGESS AND PANDE (2005)

## Results: Credit Access

- ▶ On full sample, one more bank has no effect on weather-death relationship
  - ▶ Is single bank opening enough to enable smoothing?
- ▶ Make dummy equal to one for 75th pctile of banks, zero for 25th pctile
  - ▶ Moving from 25th to 75th pctile of banks eliminates effect of temperature!
- ▶ Mitigating effect of banks concentrated within growing season
  - ▶ Smoothing by farmers key mechanism
- ▶ **Takeaway: credit enables poor to withstand environmental shocks**

## Discussion

- ▶ aggregated data (district) with satellite temperature makes it hard to distinguish atmospheric/ambient means from experienced exposure
- ▶ How can we then measure direct effect through physiological channel
  - ▶ Leave it to the epidemiologists...

## Back to the question

- ▶ How does environmental quality affect economic development?
- ▶ Most reduced form studies identify  $\frac{\partial h}{\partial e}$
- ▶ Recall, this is an input into estimating benefits (MTWP) of  $\uparrow e$ 
  - ▶ which we need to know in order to set optimal environmental policy
  - ▶ optimal in the sense that agents decisions are first-best
- ▶ Is physical health the only thing that matters for estimating  $\frac{\partial h}{\partial e}$ ?

# Chen et al. (2024): Air Pollution and Mental Health

- ▶ **Research question:** What is the impact of air pollution on mental health?
  - ▶ Direct mechanism: oxidative stress
  - ▶ Indirect mechanism: labor productivity, income, physical activity
- ▶ **Setting:** China in 2014-2015
- ▶ **Data:** Kessler Psychological Distress Scale
  - ▶ PM2.5 from China National Environmental Monitoring Center
- ▶ **Design:** instrument pollution with thermal inversions
- ▶ **Results:** Pollution ↑ depression
  - ▶ Mechanisms: lower exercise, worse physical health
  - ▶ Total cost = \$22.88 billion, similar to mortality costs!

# Background

- ▶ Why do we care about mental health?
  - ▶ Poverty associated with life satisfaction and **stress**
  - ▶ These mechanisms affect productivity, which perpetuates poverty
  - ▶ See Haushofer and Salicath (2023); Haushofer and Fehr (2014)
- ▶ Frontier in development economics: RCTs on income and stress
  - ▶ Cash transfers ↑ subjective well-being (McGuire et al., 2022)
  - ▶ Cash transfers ↓ suicide in Indonesia (Christian et al., 2019)
  - ▶ Cash transfers → depression (Haushofer and Shapiro, 2016)
- ▶ Toxicology literature: pollution ↑ oxidative stress (Calderon-Garciduenas et al., 2003; Sørensen et al., 2003; MohanKumar et al., 2008, Power et al., 2015)

# Background

- ▶ Theory of change: **pollution** → **stress** → productivity
  - ▶ Poverty correlated with pollution and stress
  - ▶ Direct effect of pollution on productivity (Graff Zivin & Niedell, 2012; California)
  - ▶ Many parts of the chain are open areas for research!

# Air Pollution and Mental Health

What are the empirical challenges?

① OVB: pollution correlated with economic activity

- ▶ Richer counties more polluted
- ▶ Induces **negative** bias
- ▶ Solution: ?

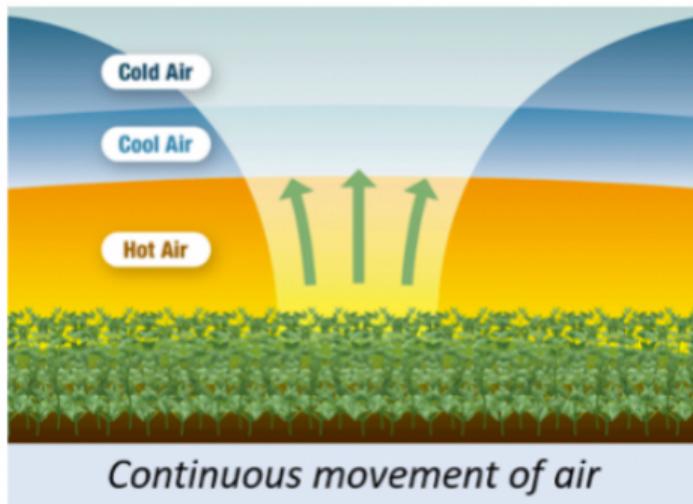
② ?

③ ?

④ ?

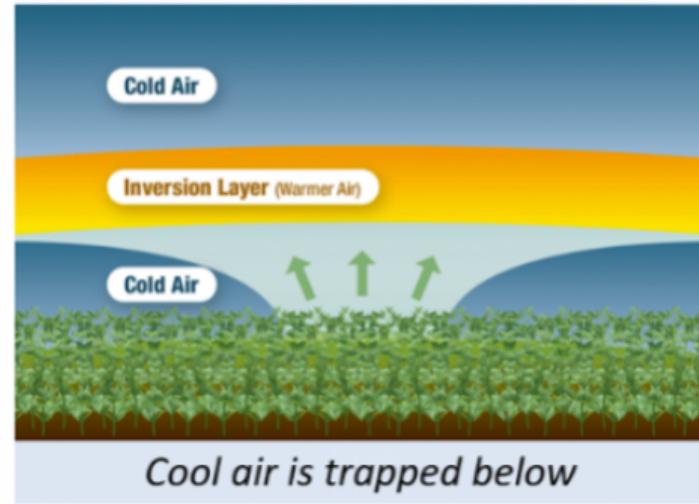
# Instrumental Variable: Thermal Inversion

During a typical day



Agweb.com

With an inversion



Agweb.com

# Instrumental Variable: Thermal Inversion



## Instrumental Variable: Thermal Inversion

- ▶ Air is warmer near surface, becomes cooler at higher altitudes
- ▶ Thermal inversion: temperature-altitude gradient is reversed
  - ▶ Rare phenomenon
  - ▶ Depends on **random** meteorological factors
- ▶ Cool air is denser and does not rise
- ▶ Pollution becomes trapped near surface

## What do you think?

- ▶ Pros and cons
- ▶ Inversions are short lived (last for a few hours)
  - ▶ cannot estimate impact of sustained exposure
  - ▶ Is mental health affected by rare short term events?
- ▶ correlated with weather, which can have direct effect on mental health
  - ▶ control for ground level temperature/rain
- ▶ If mechanism is through chronic condition, then TI may isolate wrong type of variation
- ▶ TI are very localized phenomena. Is this what we want to study?

# Data

- ▶ Mental health: 2014 China Family Panel Study Survey
  - ▶ 16,000 rural and 13,000 urban residents (162 counties out of 1,355)
  - ▶ Six questions about mental health on 0-24 Scale
  - ▶ Depression, nervousness, restlessness, hopelessness, effort, worthlessness
- ▶ Pollution: PM2.5
  - ▶ Webscraped
  - ▶ Validate with satellite data?
- ▶ Thermal inversion (MERRA-2 Air Temperature Layers)
  - ▶ Inversion = 1 if temperature of first layer (110m) < second layer (320m)
  - ▶ IV: number of inversions in month prior to each interview

# Inversions and Pollution

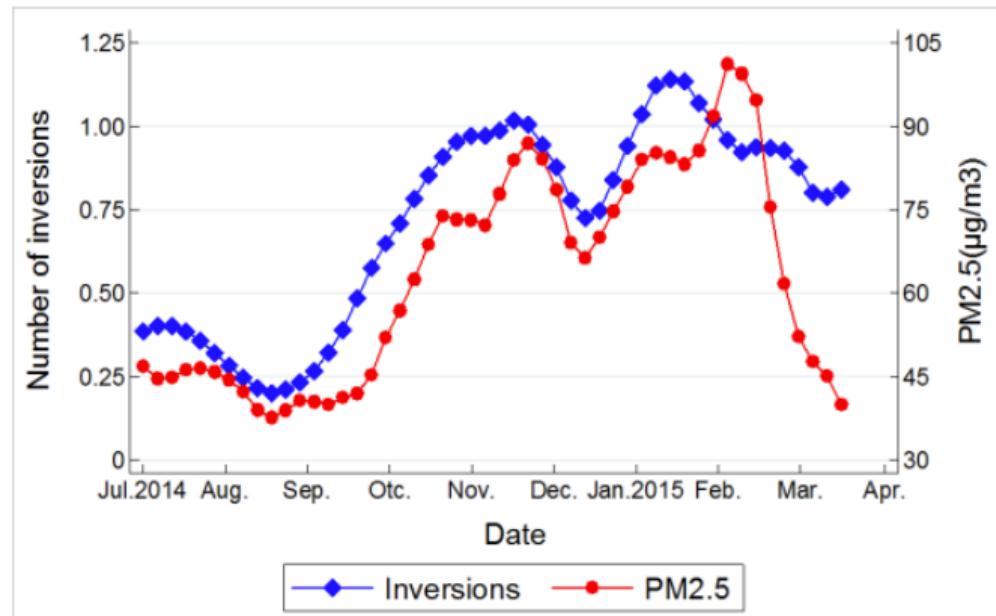


Figure 1. Time Trend of PM<sub>2.5</sub> and Thermal Inversion

Do thermal inversions seem rare and random?

# Empirical Strategy

First Stage:

$$P_{ict} = \alpha_0 + \alpha_1 I_{ict} + f(W_{ict}) + \gamma_c + \sigma_t + \mu_{ict}$$

Second Stage:

$$H_{ict} = \beta_0 + \beta_1 P_{ict} + f(W_{ict}) + \gamma_c + \sigma_t + \epsilon_{ict}$$

- ▶  $P_{ict}$  is PM2.5 in month prior to interview
- ▶  $I_{ict}$  is instrument, number of thermal inversions in past month
- ▶  $H_{ict}$  is total mental illness score (0-24)
- ▶  $f(W_{ict})$  are flexible weather controls
- ▶  $\sigma_t$  are year-month FE $s$  (**important!**)

## Main Results: Mental Health

TABLE 1—FIRST-STAGE AND IV RESULTS

Dep. var.	First stage		Second-stage estimation			
	PM <sub>2.5</sub> (1)	PM <sub>2.5</sub> (2)	K6 score (3)	K6 score (4)	Severe mental illness (5)	Severe mental illness (6)
Thermal inversions	0.2118 (0.0469)	0.2638 (0.0357)				
PM <sub>2.5</sub>			0.1229 (0.0500)	0.1164 (0.0388)	0.0047 (0.0025)	0.0042 (0.0019)
Mean [SD] of dep. var.	47.7132	[18.0450]	2.9556	[3.7598]	0.0438	[0.2047]
Weather controls	No	Yes	No	Yes	No	Yes
KP F-statistics	20.36	54.56				

- ▶ 1  $\mu/m^3$  PM2.5 ↑ mental illness by 0.1164 units (4% of mean;  $0.4\sigma$ )
- ▶ 1  $\mu/m^3$  PM2.5 ↑ prob. of several mental illness by 0.42pp (9.5% of mean)

## Main Results: Symptoms

Table 4. Effect of Air Pollution on Mental Health: By Symptom

	About how often do you feel					
	Depression (1)	Nervousness (2)	Restlessness (3)	Hopelessness (4)	Difficulty (5)	Worthlessness (6)
PM <sub>2.5</sub>	0.0155** (0.0073)	0.0045 (0.0062)	0.0134*** (0.0044)	0.0150*** (0.0048)	0.0183*** (0.0055)	0.0123** (0.0049)
Observations	12,657	12,659	12,660	12,638	12,657	12,649
County <i>FE</i>	Yes	Yes	Yes	Yes	Yes	Yes
Year-by-month <i>FE</i>	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample weights	Yes	Yes	Yes	Yes	Yes	Yes
KP <i>F</i> -statistic	36.17	36.19	36.20	36.11	36.15	36.09

- ▶ Not driven by any one particular symptom

## Main Results: Persistence

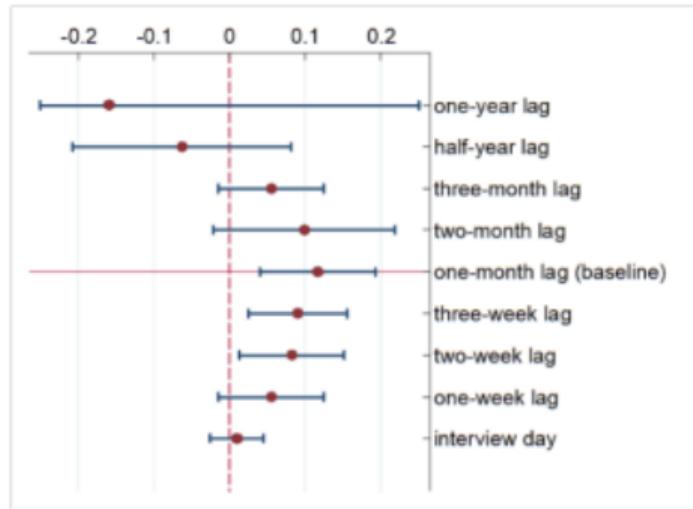


Figure A.1: Impact of PM2.5 on the K6 score

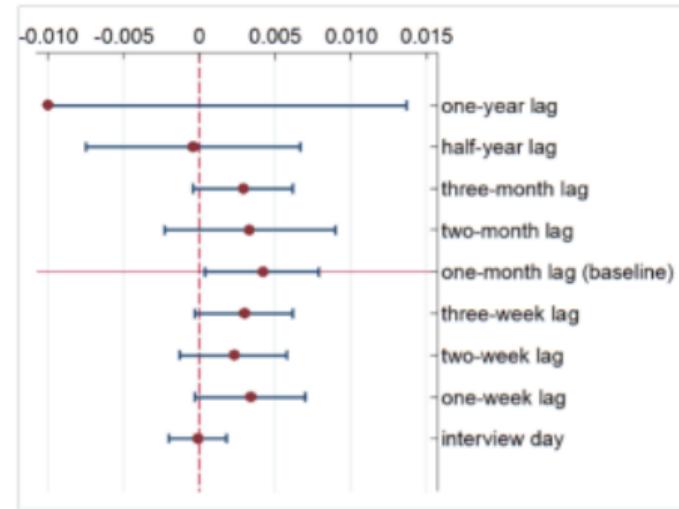


Figure A.2: Impact of PM2.5 on  $K6 \geq 12$

- ▶ Effect of pollution exposure on mental health dissipates over time

# Mechanisms

- Focus on indirect effects (physical health, productivity, behavior)

Table 6. Mechanism Tests

	Exercise (past 1 week)			Sickness (past 2 weeks)		Self-rated health (past 1 month)	
	(1) 1-yes, 0-no	(2) times	(3) hours	(4) 1-yes, 0-no	(5) 1-not serious, 5-very serious	(6) 1-very healthy, 5-very unhealthy	
PM <sub>2.5</sub>	-0.0049** (0.0022)	-0.0068* (0.0035)	-0.0448** (0.0209)	0.0020 (0.0019)	0.0089* (0.0052)	0.0092* (0.0056)	
Observations	12,664	12,663	12,670	12,670	3,806	12,668	
Mean of Dep. Var.	0.4670	2.4008	3.3681	0.2917	3.1140	2.9698	
S.D. of Dep. Var.	0.4989	3.1324	6.4268	0.4545	1.4053	1.1816	
County <i>FE</i>	Yes	Yes	Yes	Yes	Yes	Yes	
Year-by-month <i>FE</i>	Yes	Yes	Yes	Yes	Yes	Yes	
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	
Sample weights	Yes	Yes	Yes	Yes	Yes	Yes	
KP <i>F</i> -statistic	48.73	48.57	48.72	44.92	59.36	36.30	

- Driven by reduction in physical exercise, not other illness
- Very small impact compared to pollution effect ( $0.002\sigma$  vs.  $0.4\sigma$ )

# Heterogeneity

	The K6 Score			
	(1)	(2)	(3)	(4)
	Male	Female	Age 16-60	Age >=60
PM <sub>2.5</sub>	0.0986*** (0.0268)	0.0575* (0.0342)	0.0611** (0.0298)	0.1499** (0.0599)
Observations	6,093	6,522	9,446	3,169
Mean of Dep. Var.	2.6378	3.2799	2.9504	3.0281
S.D. of Dep. Var.	3.5378	3.9405	3.5819	4.2683
County FE	Yes	Yes	Yes	Yes
Year-by-month FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Sample weights	Yes	Yes	Yes	Yes
KP F-statistic	25.14	43.22	32.11	27.20

- ▶ Why are effects larger for men?

## Economic Cost? Back of the Envelope

- ▶  $1\sigma \uparrow$  in PM2.5 leads to 9.5%  $\uparrow$  in prob. of severe mental illness
  - ▶ About  $0.095 \times 1.4 \text{ billion} = 13.3 \text{ million}$  additional people
- ▶ Annual cost per patient of mental illness in China is \$USD 3,665 (Xu et. al 2016)
- ▶ Assume 8% of patients with mental illness get treated (Phillips et al., 2009)
- ▶ **Economic cost of mental illness:**  $13.3 \times 3665 \times 0.008 \times 12 = 46.7 \text{ billion USD}$
- ▶ Mortality cost of  $1\sigma$  PM2.5 is \$USD 30 billion in USA (Deryugina et al., 2016)

**Issues with the back of the envelope calculation?**

## Discussion

- ▶ small sample
- ▶ few counties
- ▶ Small time window
  - ▶ cannot include year FE's or control for annual macro trends
- ▶ exploit short term variation (daily)
- ▶ mechanism unclear

# Lecture 5

How does the environment affect economic development?

# Housekeeping

- ▶ State things that are due

# Today

- ▶ **Guiding question:** how does the environment affect economic development?
  
- ▶ Today's focus: productivity and income
  - ▶ worker productivity
  - ▶ absenteeism

# Pollution and Productivity

- ▶ We know temperature is neg. associated w/ output (Dell, Jones, and Olken 2012)
- ▶ We also know high temperatures are associated with lower crop yields
  - ▶ Mendelsohn and Dinar (1999), Auffhammer et al. (2006), Schlenker and Roberts (2009), Lobell et al. (2011), and Gupta et al. (2017)
- ▶ What are the mechanisms?
  - ▶ heat stress, mortality, conflict
- ▶ Open and policy-relevant research area
  - ▶ If we know importance of each mechanism, we know where to target policy

- ▶ **Research question:** What is the effect of temperature on output?
  - ▶ How does heat stress mediate the temperature-output relationship
- ▶ **Setting:** India's manufacturing sector (cloth weaving, sewing, steel)
- ▶ **Data:** High-frequency (15 year panel, daily)
  - ▶ Worker-level, sewing line level, team level, plant-year, district-year
- ▶ **Design:** panel fixed effects
- ▶ **Result:** 1) heat ↓ plant-level annual output, 2) driven by output elasticity of labor, not capital 3) AC used in labor-intensive tasks (pareto efficient)

# Background

- ▶ Workers should respond quickly when working in uncomfy temperatures
  - ▶ Heat impacts on labor can be identified from daily or aggregated data
  - ▶ Distinguishes heat stress from alternative mechanisms which materialized over long time scales (i.e. conflict)
- ▶ Two channels through which temperatures affect workers
  - ▶ Produce less (productivity)
  - ▶ Absenteeism
- ▶ Can we expect lumpy investments in climate control in developing countries?
  - ▶ Depends on value-added per worker
  - ▶ If too low, then electricity cost not worth it

# Data: Worker + Cloth Weaving

- ▶ Worker data
  - ▶ Surveyed output and attendance from cloth-weaving, garment sewing, and steel firms
  - ▶ Three cloth-weaving factories in Surat,
  - ▶ Eight garment plants (Delhi, Hyderabad, Chhindwara)
  - ▶ Steel plant (one large plant in Central India)
- ▶ Cloth data (2012): 147 workers
  - ▶ Daily meters woven (for each worker)
  - ▶ Worker pay slips each day (output + attendance)

## Data: Garment Sewing

- ▶ Garment sewing (2012-2013): Line level
  - ▶ Lines of 10-20 workers
  - ▶ Productivity: hourly output relative to target
  - ▶ Daily production from 103 sewing lines
- ▶ Climate control
  - ▶ Phased rollout of cooling equipment
  - ▶ 84 of 103 sewing lines in Delhi, of which 74 had cooling

## Data: Rail and Steel Plant

- ▶ Supplies rectangular blocks of steel to Indian Railways
- ▶ Three shifts per day
- ▶ Workers assigned to one of three teams, which rotate across shifts
- ▶ Data:
  - ▶ Team and number of blocks for each shift during 1999-2008
  - ▶ 9,172 shifts over 3,337 working days
  - ▶ Team-level absences over 857 days for 2000-2003

# Garment factory and cloth weaving



# Rail/Steel Mill



## Data: Panel

- ▶ Annual Survey of Industries (1999-2013)
  - ▶ Census of large manufacturing plants (> 100 workers)
  - ▶ Random sample of small plants
  - ▶ Output, fixed assets, input expenses, employment
  - ▶ Unbalanced panel of 58,377 plants
- ▶ District Manufacturing GDP (1998-2009)
  - ▶ Planning Commission of India
  - ▶ ASI plants + unregistered plants
- ▶ Weather: public weather stations + gridded data (IMD)
  - ▶ Daily maximum at closest station to factory
  - ▶ Gridded data for panel data

## Data: Diamond Firms

- ▶ Survey 150 diamond cutting plants in Surat
  - ▶ Random sample from 500 registered units
- ▶ Five operations: 1) grading 2) marking, 3) rounding, 4) cutting, 5) polish
- ▶ Ask about number of workers and use of AC in each operation
- ▶ Ask to rate on 1-5 scale importance of each operation to quality of final output
  - ▶ Why?

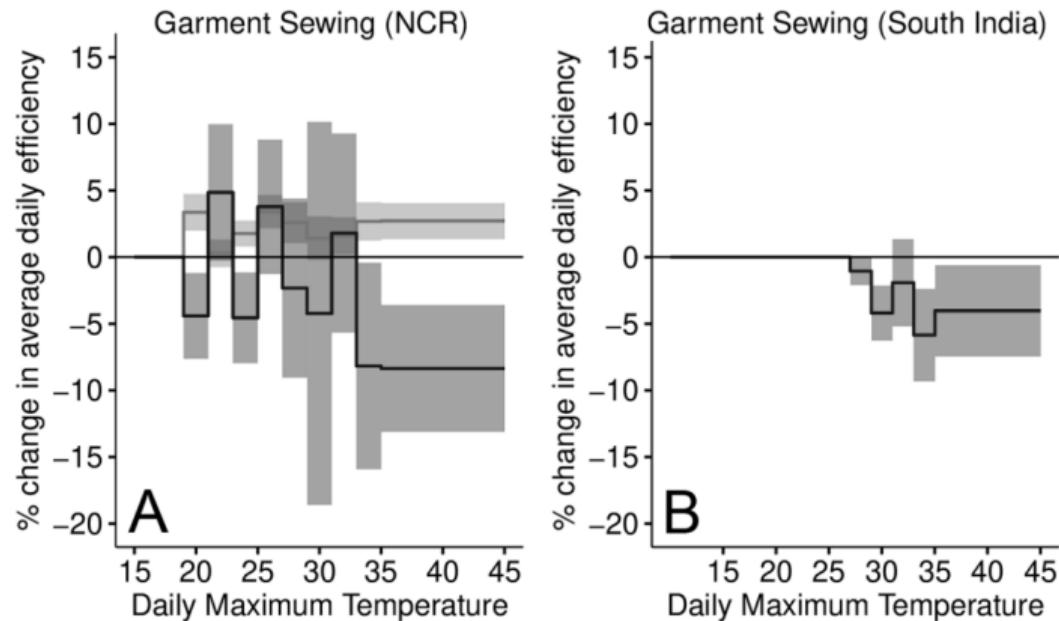
# Empirical Strategy

Temperature-output estimated with weekly data as follows:

$$y_{iw} = \alpha_i + \gamma_M + \gamma_t + \sum_{j=2}^J \beta_j T_{iw}^j + \theta R_{iw} + \lambda X_{iw} + \epsilon_{iw}$$

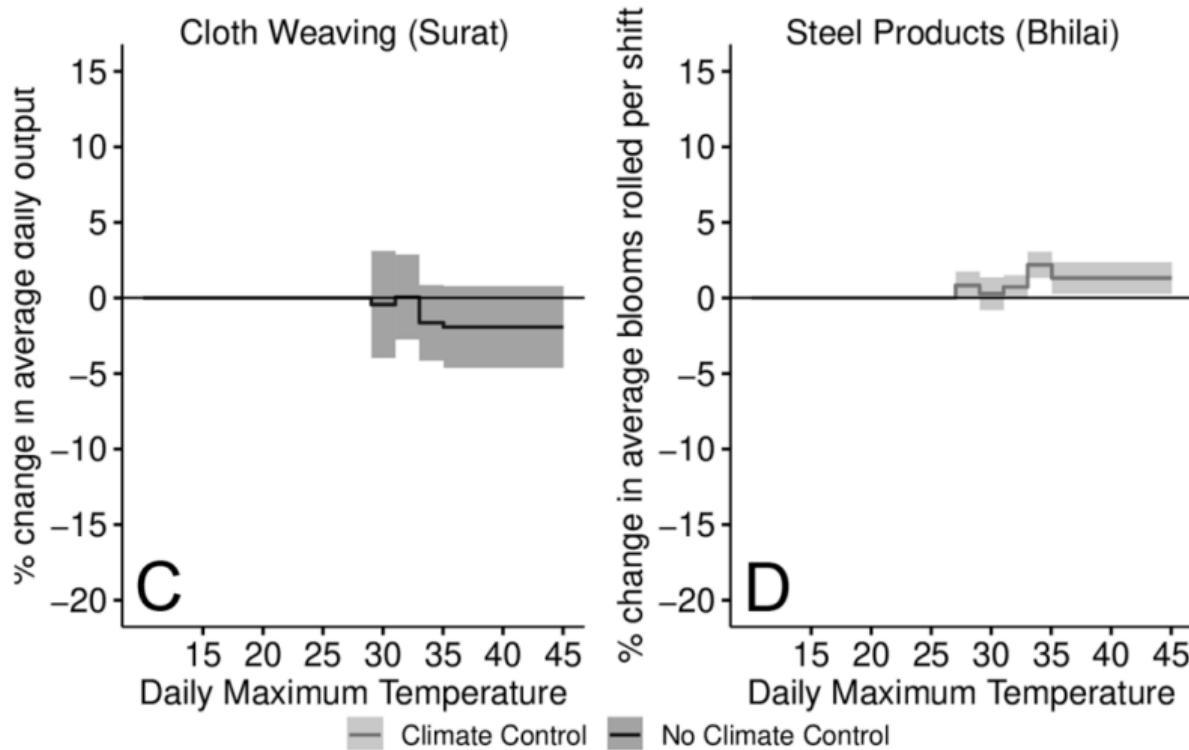
- ▶  $y_{iw}$  is physical output by worker  $i$  in week  $w$
- ▶ If absent, output=0 for that day before aggregating to week
- ▶  $\alpha_i$  is individual FE (what does it account for?)
- ▶  $\gamma_{M/t}$  are month/year FEs (what does it account for?)
- ▶  $T_j$  counts days falling into temperature bin  $j$  (captures nonlinearities)

## Results: Garment Sewing



- ▶ output is lower in hotter weeks **without climate control**
- ▶ Moving from lowest to highest temperature bin ↓ productivity by 8%
- ▶ Note: these are 90% CI's!

## Results: Cloth and Steel



## Lagged Effects

- ▶ Previously,  $\beta_j$  combined productivity and absenteeism channel
- ▶ Now we use daily data

$$y_{id} = \alpha_i + \gamma_M + \gamma_t + \sum_j \beta_j T_{id}^j + \sum_j \omega_j L_{id}^j + \theta R_{id} + \lambda X_{id} + \epsilon_{id}$$

- ▶ where  $T^j$  is now an indicator for day  $d$  falling into temperature bin  $j$
- ▶  $X_{id}$  includes day-of-week FE
- ▶  $L_{id}^j$  counts number of days in bin  $j$  during six preceding days

## Discussion

- ▶ Is lag specification plausible?
  - ▶ Is six days a plausible window to detect cumulative exposure?
- ▶ What are the identification assumptions?
- ▶ What are the threats?

# Lagged Effects: Productivity

TABLE 2  
EFFECT OF HOT DAYS ON WORKER OUTPUT

	CLIMATE CONTROL		NO CLIMATE CONTROL		
	Garments (Log Efficiency) (1)	Steel (Log Blooms Rolled) (2)	Weaving (IHS Meters) (3)	Garments (Log Efficiency) (4)	Garments (Log Efficiency) (5)
T (33°–35°C)	.025** (.010)	.028* (.017)	-.040** (.019)	-.129*** (.042)	-.007 (.037)
T (above 35°C)	.035*** (.014)	.020** (.009)	.011 (.022)	-.154*** (.041)	.008 (.046)
L (33°–35°C)	-.004 (.005)	.005 (.004)	-.033*** (.011)	-.009 (.012)	.004 (.010)
L (above 35°C)	-.011** (.005)	-.002 (.005)	-.027*** (.009)	-.019 (.027)	.015 (.018)
Climate control	Yes	Yes	No	No	No
Number of units	74 lines	9 teams	147 workers	10 lines	19 lines
Time span (days)	730	3,337	365	730	730

- ▶ Lower output on hot days in sites w/o climate control
- ▶ Weavers: additional day > 35° in past week ↓ daily output by 3%

# Lagged Effects: Absenteeism

TABLE 3  
EFFECT OF HOT DAYS ON WORKER ABSENTEEISM

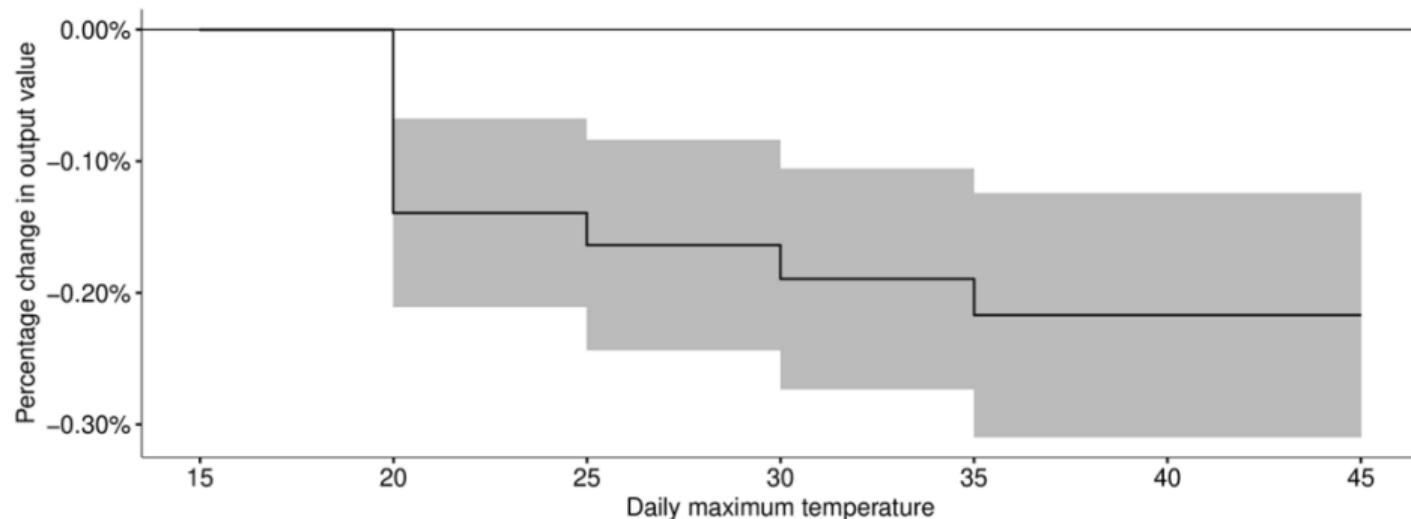
	CLIMATE CONTROL				NO CLIMATE CONTROL	
	Garments		Steel (All) (3)	Weaving (All) (4)	Garments	
	Paid (1)	Unpaid (2)			Paid (5)	Unpaid (6)
T (33°–35°C)	.082*** (.022)	−.083 (.065)	−.011 (.048)	.003 (.004)	−.001 (.128)	.796 (.678)
T (above 35°C)	.115*** (.027)	.031 (.049)	.051 (.068)	−.004 (.004)	−.034 (.117)	1.001 (.862)
L (33°–35°C)	−.018 (.011)	−.047 (.032)	.044*** (.014)	.006*** (.002)	.017 (.077)	.772 (.686)
L (above 35°C)	.021** (.010)	−.001 (.022)	.045** (.020)	.005*** (.002)	.078 (.083)	.567 (.426)
Number of units	224 lines		9 teams	147 workers	42 lines	
Time span (days)	730		3,337	365	730	

- ▶ How do you interpret these results?

## 15-year plant-level panel results

$$y_{it} = \alpha_i + \gamma_t + \sum_{j=2}^5 \beta_j T_{it}^j + \theta R_{it} + \epsilon_{it}$$

outcome is log value of plant  $i$  output in year  $t$



## Recap of Results

- ▶ Daily worker data: workers are less productive on hot days
- ▶ Daily worker data: (kind of) more absenteeism on hot days
- ▶ Plant panel: **aggregate** output is lower on hot days
  - ▶ This is the object we are interested in
  - ▶ Remember  $\frac{\partial \Delta y}{\partial h} \frac{\partial h}{\partial e}$
- ▶ **How much of the temperature-induced  $\Delta Y$  is due to  $\Delta MP_L$ ?**

## Lets make a Cobb-Douglas model

Consider a production function:

$$Y = A \cdot K^\omega \cdot L^\beta$$

Now modify by letting elasticities be a function of temperature,  $T$

$$Y = A(T) \cdot K^{\omega(T)} \cdot L^{\beta(T)}$$

where:

$$A(T) = A_0 + \sum a_j T_j$$

$$\omega(T) = \omega_0 + \sum \omega_j T_j$$

$$\beta(T) = \beta_0 + \sum \beta_j T_j$$

# Estimating Equation

- Log-linearizing, we get Equation (7) from the paper:

$$y = \alpha_0 + \sum_{j=2}^5 \alpha_j T^j + \omega_0 \cdot k + \sum_{j=2}^5 \omega_j T^j k + \beta_0 \cdot l + \sum_{j=2}^5 \beta_j T^j l$$

- Data:
  - k: opening value of equipment and machinery at start of financial year
  - l: number of full-time workers
- Results
  - temperature effects on output elasticity of L is negative and significant
  - temperature effects on output elasticity of K are (weakly) positive
  - Controlling for interactions, residual temperature-output effect is zero
  - Since overall temperature-output relationship is negative, temperature-induced declines in labor productivity drive the effect

# Results

TEMPERATURE INTERACTIONS WITH FACTOR INPUTS

	(1)	(2)	(3)	(4)	(5)
$T^a$				.02324*	
				(.01345)	
$T^2$	.00256** (.00097)	.00008 (.00211)	-.00008 (.00162)		
$T^3$	.00147 (.00103)	-.00205 (.00229)	-.00009 (.00165)		
$T^4$	.00081 (.00108)	-.00094 (.00237)	-.00028 (.00170)		
$T^5$	.00003 (.00118)	-.00499* (.00259)	-.00171 (.00185)		
$l$	.8612*** (.0957)		.91426*** (.09660)	.36520*** (.05910)	
$k$		.20433** (.05674)	.06629 (.04114)		
$l \times T^2$	-.00098*** (.00027)		-.00134*** (.00034)	-.00056** (.00022)	
$l \times T^3$	-.00067** (.00027)		-.00104*** (.00027)	-.00038*** (.00017)	
$l \times T^4$	-.00052* (.00027)		-.00077*** (.00027)	-.00030* (.00017)	
$l \times T^5$	-.00036 (.00029)		-.00075*** (.00029)	-.00039** (.00018)	
$k \times T^2$		-.00009 (.00015)	.00028* (.00015)		
$k \times T^3$		.00005 (.00016)	.00022* (.00012)		
$k \times T^4$		-.00003 (.00016)	.00016 (.00011)		
$k \times T^5$		.00022 (.00018)	.00024** (.00012)		

## Discussion

- ▶ What do you think about the paper
- ▶ Critique
- ▶ How can they improve the identification?
- ▶ This data is as good as it gets: did they do everything they could?
  - ▶ What about long term effects/adaptation?
- ▶ Alternative explanations

## Back to the model

$$MWTP_e = \frac{\lambda_e}{\lambda_y} = \frac{1}{\lambda_y} \left( \frac{\partial u}{\partial e} + \frac{\partial u}{\partial h} \frac{\partial h}{\partial e} \right) + \frac{\partial \Delta y}{\partial e} + \frac{\partial \Delta y}{\partial h} \frac{\partial h}{\partial e}$$

- ▶ Most research in environment/development economics seeks to identify  $\frac{\partial h}{\partial e}$ 
  - ▶ At this point,  $\frac{\partial h}{\partial e} > 0$  is unambiguous
  - ▶ Focus is on **magnitudes**
- ▶ What about indirect effects via productivity ( $\frac{\partial \Delta y}{\partial h} \frac{\partial h}{\partial e}$ )?
  - ▶ This is the research frontier
- ▶ Next: what about  $\frac{\partial \Delta y}{\partial e}$ , the direct effect of environment on incomes
  - ▶ Almost no research on this

## How might the environment **directly** affect incomes?

- ▶ Assume landholders vary in distance to a central market
- ▶ Each chooses land and labor to maximize farm profits
- ▶ New PA imposed that overlaps existing productive land and restricts land use
- ▶ Assuming initially efficient allocation, what happens to net welfare?
- ▶ What if land rents increase outside the PA?

## Background: Conservation and Development

- ▶ On other hand, PAs may generate tourism income and other positive spillovers
- ▶ Surprisingly little evidence on income channel
- ▶ Main challenge: non-random placement of PAs

## Sims (2010): Conservation and Development: Evidence from Thailand

**Research Question:** How does PA policy affect local poverty outcomes? - What are the mechanisms?

**Setting:** Thailand in 2000

**Data:** Household consumption surveys + PA locations

**Empirical Strategy:** Instrument PA location with priority watershed status

**Results:** PAs increase consumption and lower poverty - Higher tourism is the main mechanism

# Data

- ▶ Wildlife sanctuaries (N=31) and national parks (N=57)
  - ▶ Source: World Database on Protected Areas
  - ▶ Includes year of establishment
- ▶ Poverty Mapping
  - ▶ Model household income in 2000 Thai Economic Survey as a function of household characteristics and assets
  - ▶ Use coefficient to predict household income in census households
- ▶ Poverty Measure

$$FGT(\alpha) = \frac{1}{n} \sum_{y_i < p} \left( \frac{p - y_i}{p} \right)^\alpha$$

where  $p$  is poverty line and  $y_i$  is income

# Empirical Strategy: Ordinary Least Squares

$$y_{ij} = \beta_1 PA_{ij} + \beta_2 G_{ij} + \beta_3 X_{ij} + \alpha_j + \epsilon_j$$

- ▶  $y_{ij}$  is socioeconomic outcome for locality  $i$  in district  $j$
- ▶  $PA_{ij}$  is share of locality land designated as PA prior to 2000
- ▶  $G_{ij}$  are fixed geographic controls (slope, elevation, water, etc.)
- ▶  $X_{ij}$  is a measure of pre-treatment (historic forest cover, distance to major roads)
- ▶  $\alpha_j$  are district fixed effects

**What variation is exploited?**

**What are the identification assumptions?**

**What are potential omitted characteristics?**

# OLS Results

Outcomes		(1)	(2)	(3)	(4)
	Log consumption (Baht/month)		Poverty headcount ratio	Poverty gap	Squared poverty gap
<b>Share protected</b>	0.135*** (0.033)		-0.066*** (0.017)	-0.017*** (0.006)	-0.006** (0.002)
$R^2$	0.593		0.702	0.627	0.563
N localities	4113		4113	4113	4113

- Median land share protect is 1/3
- Expected ↑ in consumption is about 4.5% (divide by coeff. by 3)

# Instrumental Variables Solution

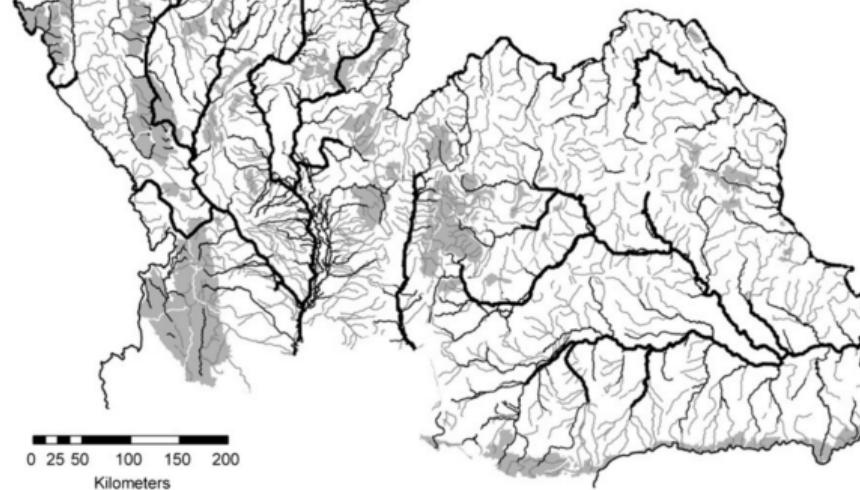
- ▶ Instrument PA placement with **priority watershed status**
- ▶ Upper watersheds (headwaters) protected to improve **downstream** water quality
  - ▶ Watershed protected not driven by local (headwater) economic needs
- ▶ First Stage:

$$PA_{ij} = \pi_1 Z_{ij} + \pi_2 G_{ij} + \pi_3 X_{ij} + \pi_4 W_{ij} + \delta_j + u_{ij}$$

- ▶  $Z_{ij}$  includes a)  $i$  overlaps major tributary, b) distance to nearest major tributary .
- ▶  $\pi_1$  picks up variation in PA driven by differences in downstream characteristics

# Watershed Instrument

a



# Discussion

Is the IV plausible? Why or why not?

## IV Results

	Outcomes			
	(1) Log consumption (Baht/month)	(2) Poverty headcount ratio	(3) Poverty gap	(4) Squared poverty gap
<b>IV estimates</b>				
A. Intersects with tributary				
<b>Share protected</b>	0.372** (0.184)	-0.183** (0.092)	-0.063* (0.034)	-0.030* (0.017)
B. Distance to major tributary				
<b>Share protected</b>	0.358** (0.172)	-0.187** (0.093)	-0.069** (0.035)	-0.033* (0.017)
C. Both instruments				
<b>Share protected</b>	0.365** (0.170)	-0.185** (0.088)	-0.066** (0.033)	-0.031* (0.016)

- Signs agree with OLS, but magnitudes are more than 2x as large
- Consistent with idea that OLS negatively biased

## Mechanisms: Is tourism responsible?

- ▶ PAs restricts land use, so why does income increase?
- ▶ Test 1: Test whether impacts in national parks > in wildlife sanctuaries
  - ▶ Idea: NP's encourage tourism whereas sanctuaries only permit ecotourism/research
- ▶ Test 2: Back-of envelope calculation
  - ▶ if we know costs/benefits of tourism, is it enough to explain net ↑ in consumption?
- ▶ Test 3: Are benefits higher where tourism is higher
  - ▶ Park statistics show that visitation is highest near major cities

# Test 1: National parks vs. Wildlife Sanctuaries

Socioeconomic outcomes and share protected by IUCN category.

	Outcomes			
	(1) Log consumption (Baht/mo)	(2) Poverty headcount ratio	(3) Poverty gap	(4) Squared poverty gap
<b>Share WLS (I)</b>	0.054 (0.050)	-0.019 (0.032)	-0.008 (0.013)	-0.004 (0.006)
<b>Share NP (II)</b>	0.086** (0.038)	-0.057*** (0.018)	-0.015*** (0.005)	-0.005** (0.002)
<b>Share FR (IV)</b>	-0.062*** (0.016)	0.022*** (0.008)	0.005* (0.003)	0.002 (0.001)
$R^2$	0.870	0.596	0.703	0.628
N localities	4113	4113	4113	4113

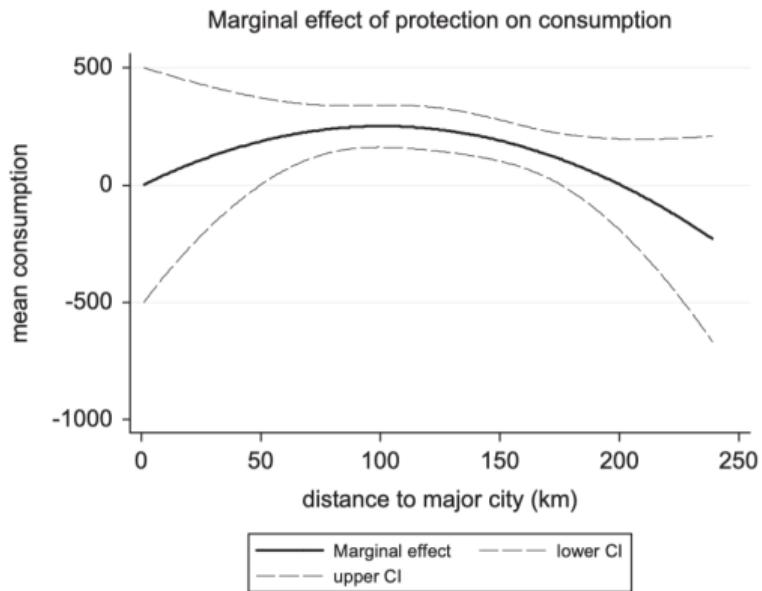
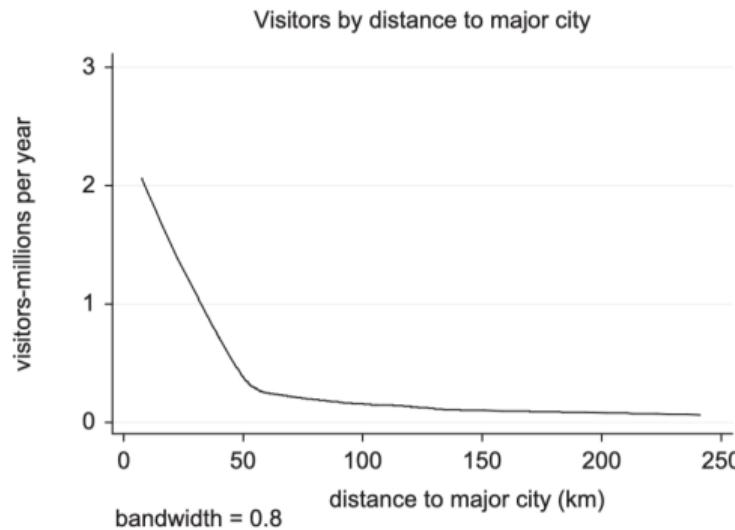
- ▶ Effects larger where tourism is allowed (NPs)
- ▶ Difference in coefficients not statistically significant

## Test 2: Back of the envelope calculation

- ▶ Assume income is agricultural and reduced proportionally to land lost
- ▶ Assume credit constraints (no savings) → income losses = consumption losses
- ▶ Consumption cost of PAs is 11% of income (forest cover coeff. - see paper)
- ▶ To achieve 5% *net* ↑ in consumption, tourism must generate 16% consumption
  - ▶ Equivalent to 233 Baht per month ( $0.16 \times \text{mean consumption}$ )
- ▶ 10.81 mil. tourists in 2020 and 172,978 hh in PA localities = 61 tourists/hh
- ▶ Each tourist would need to contribute:  $(233*12)/61 = 46 \text{ Baht } (\$/\text{USD } 1.20)$  of net income to generate observed increases in consumption

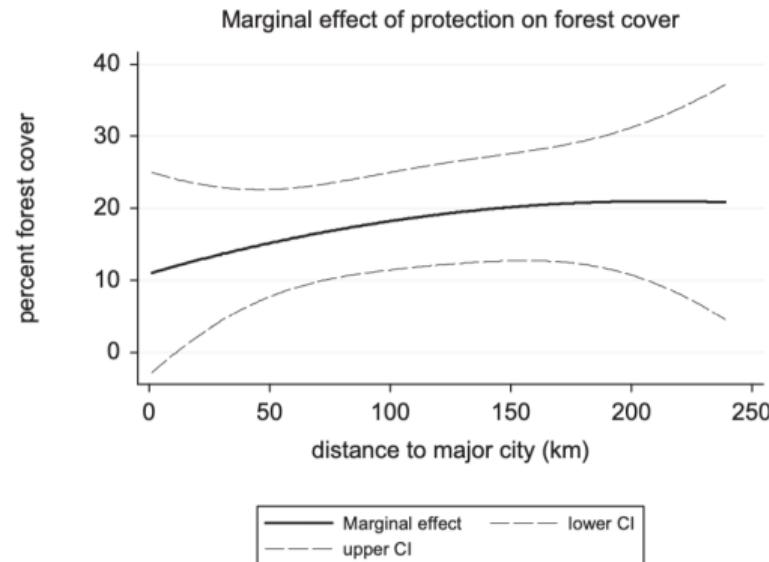
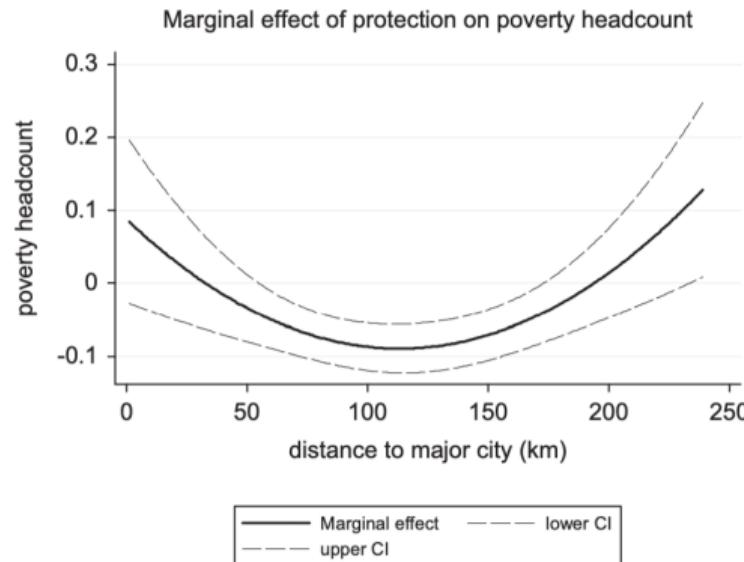
**Is this reasonable?**

## Test 3: Heterogeneity by Distance to City



- ▶ Most tourists near city, but effects concentrated at intermediate distance
- ▶ Why?

# Test 3: Heterogeneity by Distance to City



## Discussion

- ▶ What do you think of the paper?
- ▶ Do you believe the instrument?
- ▶ What are some alternative explanations?