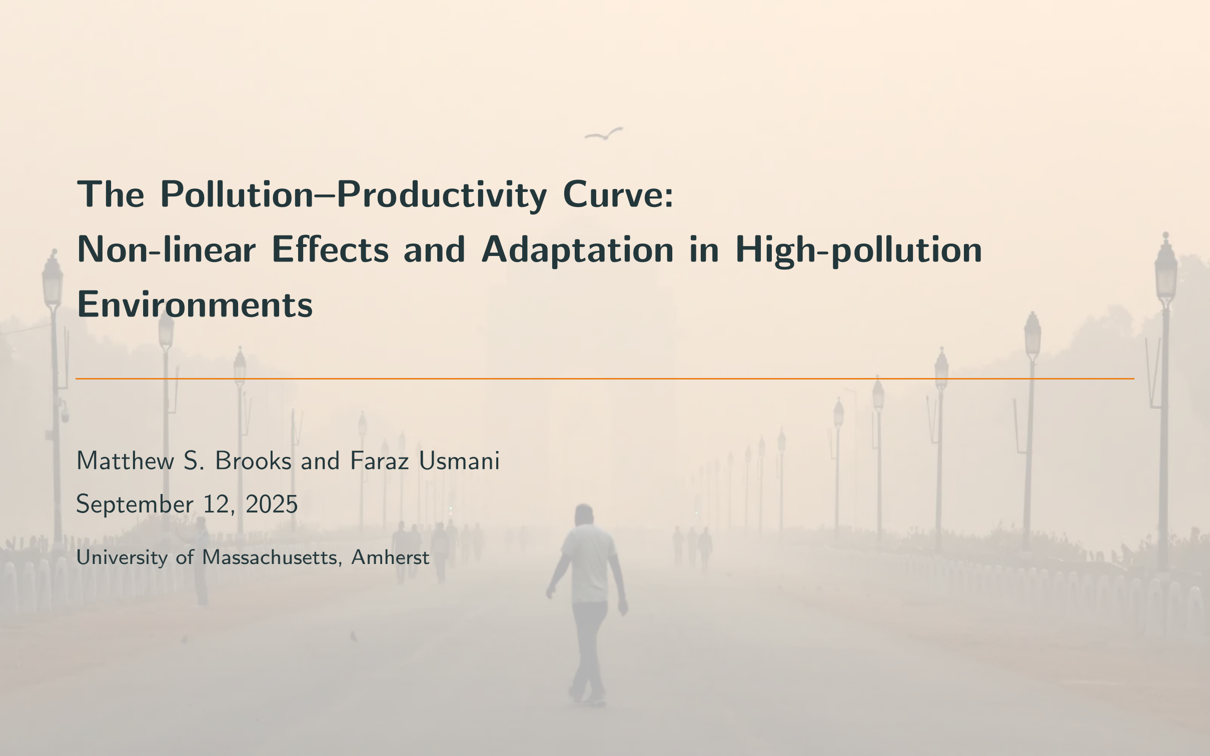


The Pollution–Productivity Curve: Non-linear Effects and Adaptation in High-pollution Environments

Matthew S. Brooks and Faraz Usmani

September 12, 2025

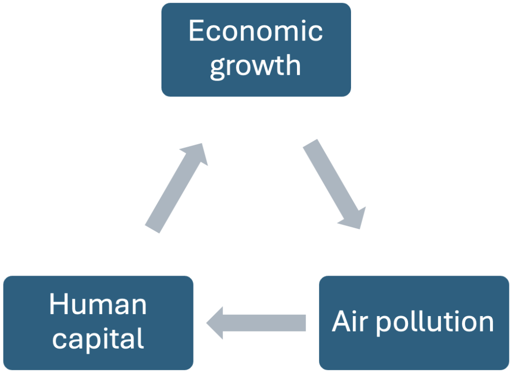
University of Massachusetts, Amherst

Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

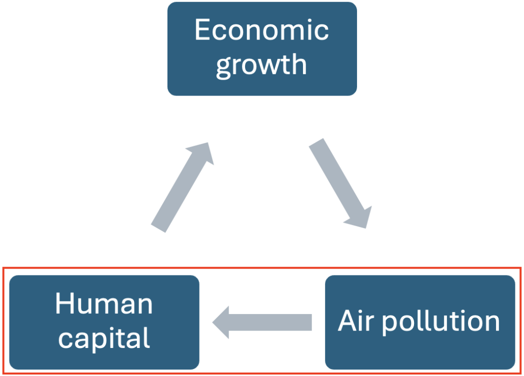
Why do we care about air pollution?

Air pollution lowers labor productivity (Adhvaryu et al., 2022; Archsmith et al., 2018; Behrer et al., 2023; Borgschulte et al., 2022; Chang et al., 2016; Chung et al., 2025; Dechezleprêtre et al., 2019; Fu et al., 2017; Guo and Fu, 2019; Hansen-Lewis, 2024; He et al., 2019; Hill et al., 2024; Hoffmann and Rud, 2024; Holub and Schündeln, 2023; Kahn and Li, 2020; Lichter et al., 2017; Merfeld, 2023; Mo et al., 2023)



Why do we care about air pollution?

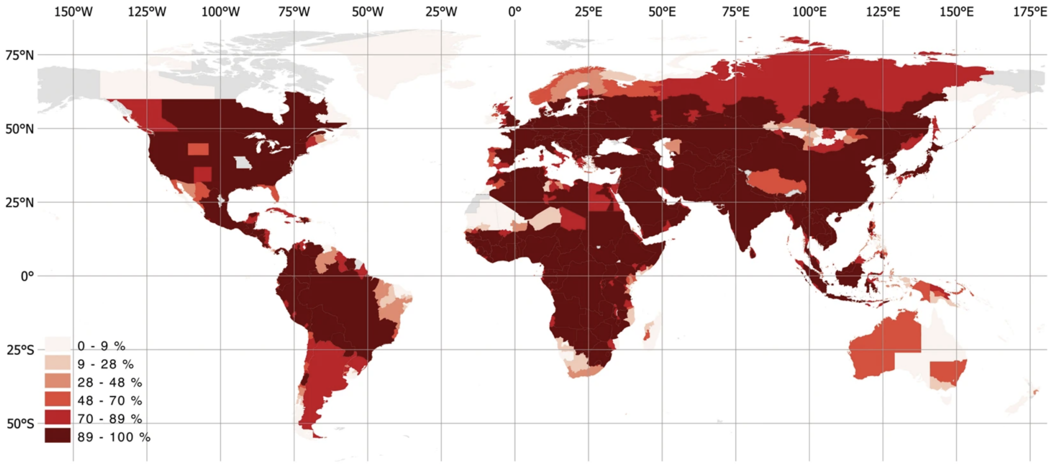
Air pollution lowers labor productivity (Adhvaryu et al., 2022; Archsmith et al., 2018; Behrer et al., 2023; Borgschulte et al., 2022; Chang et al., 2016; Chung et al., 2025; Dechezleprêtre et al., 2019; Fu et al., 2017; Guo and Fu, 2019; Hansen-Lewis, 2024; He et al., 2019; Hill et al., 2024; Hoffmann and Rud, 2024; Holub and Schündeln, 2023; Kahn and Li, 2020; Lichter et al., 2017; Merfeld, 2023; Mo et al., 2023)



Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

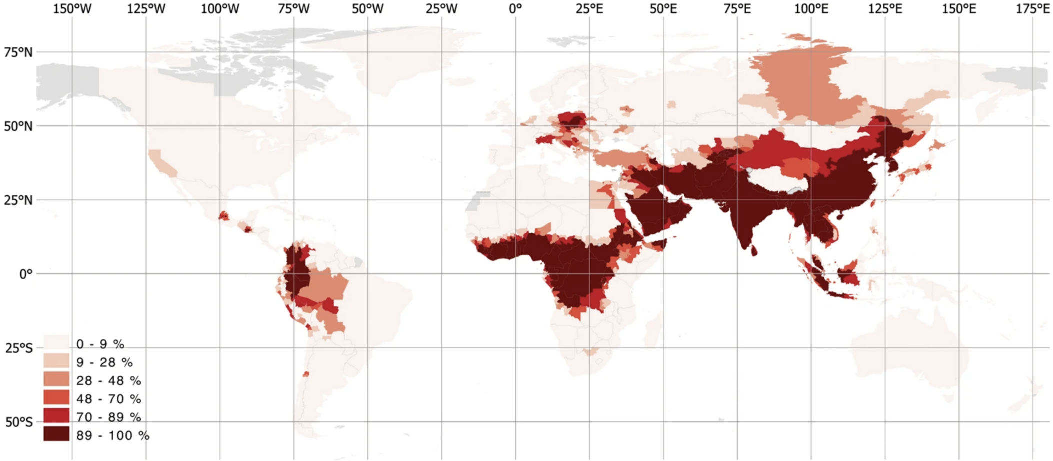
Figure 1: Percentage of the population exposed to annual average PM2.5 over **5** $\mu g/m^3$



Source: Rentschler and Leonova (2023).

Note: U.S. National Standard for PM2.5 is 9 $\mu g/m^3$, average levels in U.S. 7-9 $\mu g/m^3$ in 2013-2023.

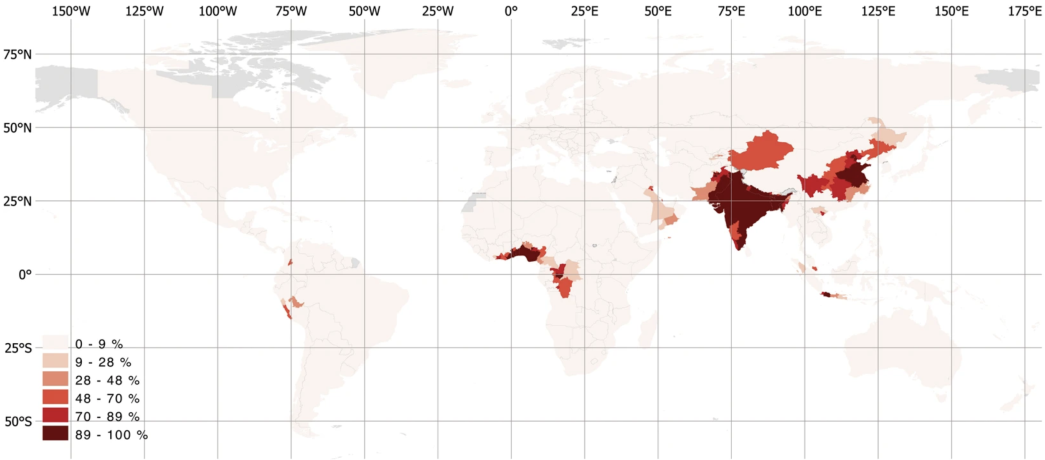
Figure 1: Percentage of the population exposed to annual average PM2.5 over **15** $\mu g/m^3$



Source: Rentschler and Leonova (2023).

Note: U.S. National Standard for PM2.5 is 9 $\mu g/m^3$, average levels in U.S. 7-9 $\mu g/m^3$ in 2013-2023.

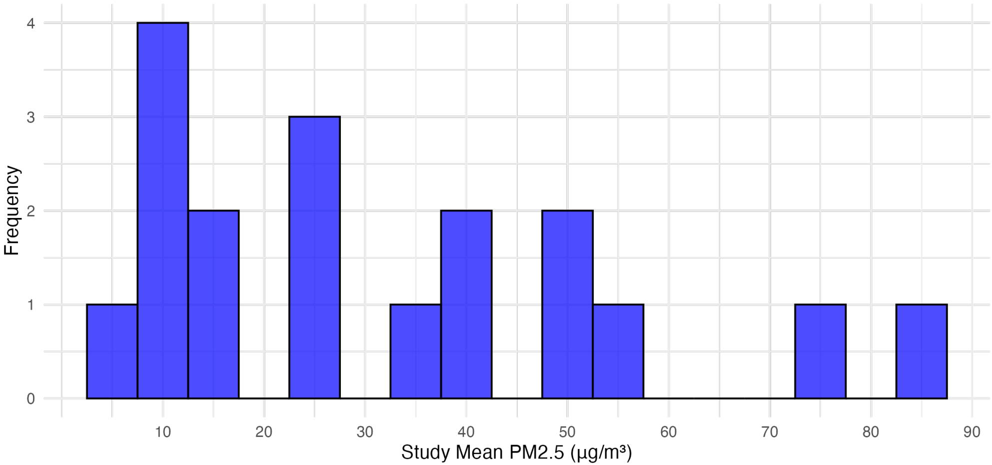
Figure 1: Percentage of the population exposed to annual average PM2.5 over **35** $\mu g/m^3$



Source: Rentschler and Leonova (2023).

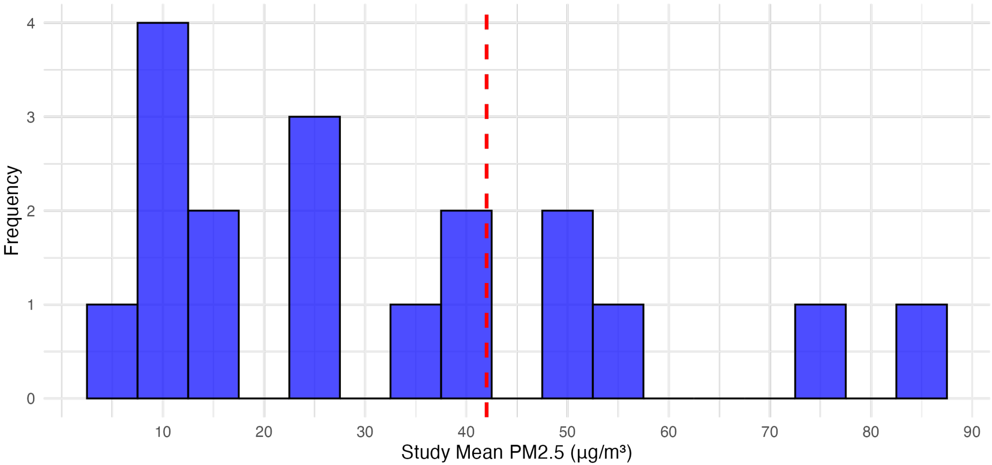
Note: U.S. National Standard for PM2.5 is 9 $\mu g/m^3$, average levels in U.S. 7-9 $\mu g/m^3$ in 2013-2023.

Figure 2: Mean PM2.5 in 18 Studies Assessing Effect of PM2.5 on Labor Productivity



Note: 5 $\mu g/m^3$ bins. Source: author's calculations from Adhvaryu et al. (2022); Archsmith et al. (2018); Behrer et al. (2023); Borgschulte et al. (2022); Chang et al. (2016); Chung et al. (2025); Dechezleprêtre et al. (2019); Fu et al. (2017); Guo and Fu (2019); Hansen-Lewis (2024); He et al. (2019); Hill et al. (2024); Hoffmann and Rud (2024); Holub and Schündeln (2023); Kahn and Li (2020); Lichter et al. (2017); Merfeld (2023); Mo et al. (2023) .

Figure 2: Mean PM2.5 in 18 Studies Assessing Effect of PM2.5 on Labor Productivity



Note: 5 $\mu g/m^3$ bins. Source: author's calculations from Adhvaryu et al. (2022); Archsmith et al. (2018); Behrer et al. (2023); Borgschulte et al. (2022); Chang et al. (2016); Chung et al. (2025); Dechezleprêtre et al. (2019); Fu et al. (2017); Guo and Fu (2019); Hansen-Lewis (2024); He et al. (2019); Hill et al. (2024); Hoffmann and Rud (2024); Holub and Schündeln (2023); Kahn and Li (2020); Lichter et al. (2017); Merfeld (2023); Mo et al. (2023) .

Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

This paper

Analyze cricket performance data from Indian Premier League (IPL) with plausibly exogenous exposure to PM2.5 to identify effect of PM2.5 on performance

Research question

What is marginal impact of PM2.5 exposure on labor productivity?

- How does this vary by long-term exposure?

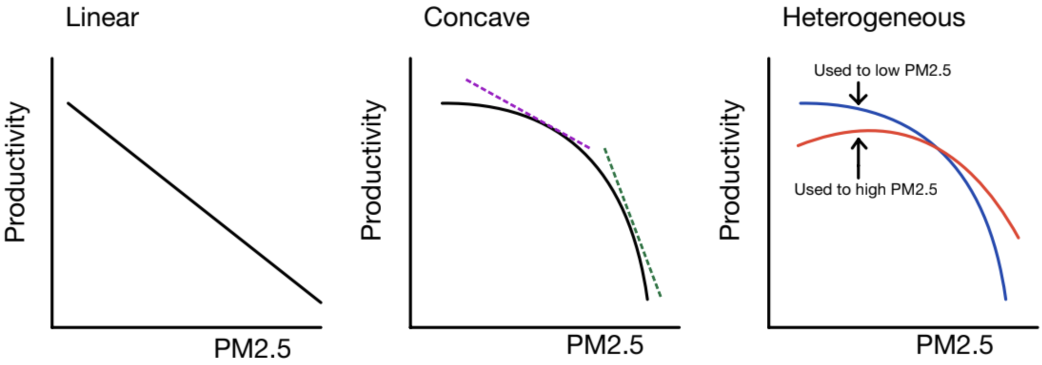
Antecedents

- Mullins (2018) examines how short-term effect of ozone varies by recent ozone exposure for college track athletes in U.S.

Contributions

- High pollution setting with novel data source (Wang et al., 2024)
- Trace out non-linearities in dose response function
- Examine how long-term exposure mediates short-term impacts

Figure 3: Non-Linear Dose-response of Productivity to PM2.5



Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

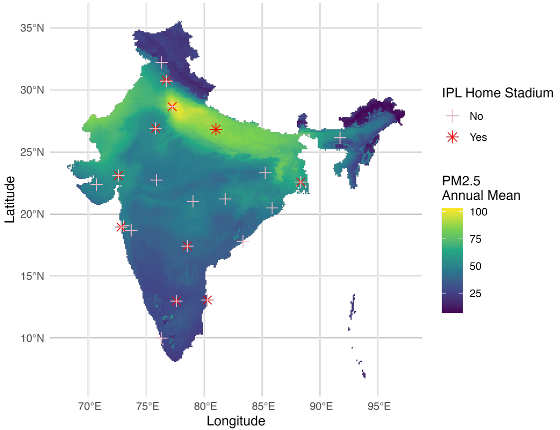
Empirical setting: cricket games in Indian Premier League

Source: Wikipedia.

- 183,572 deliveries (throws)
  Histogram
- 773 matches; 445 bowlers; 575 batters, 20 stadiums, 2008-2022.
- Bowlers (pitchers) likely to be more affected by PM2.5
  1. Higher respiration rates
  2. Tend to have longer exposure in game
     Density plot

PM2.5 air pollution in India varies across space and time

Figure 4: Air pollution in India is high, but varies across space

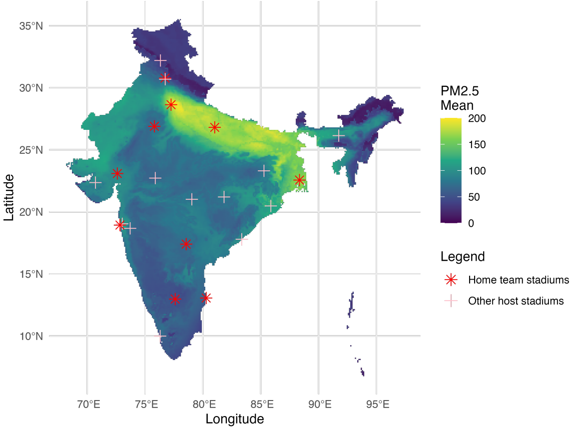


Notes. Annual mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: January

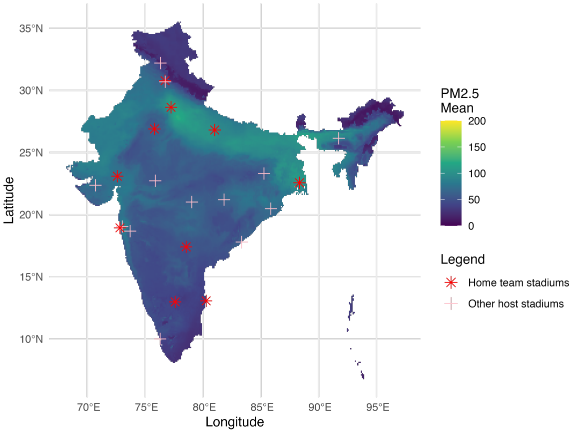


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: February

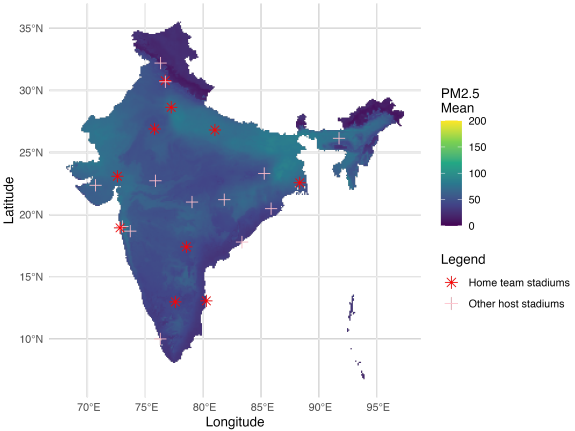


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: March

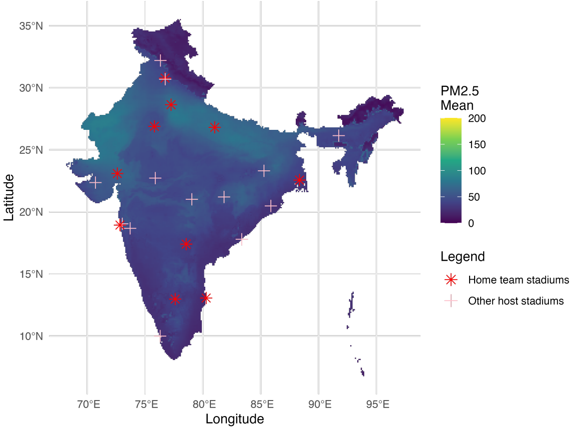


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: April

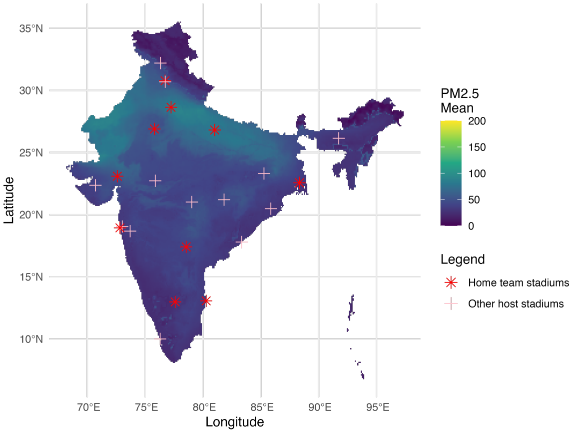


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: May

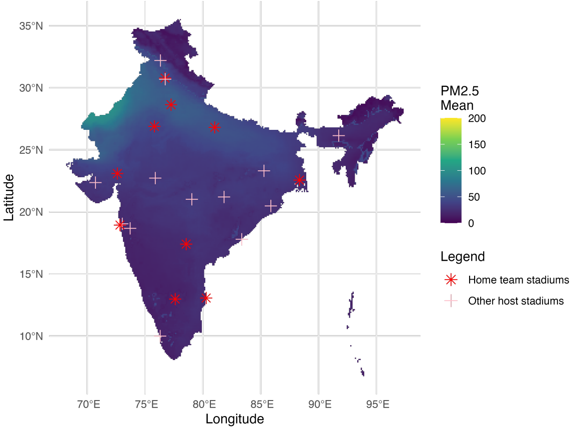


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: June

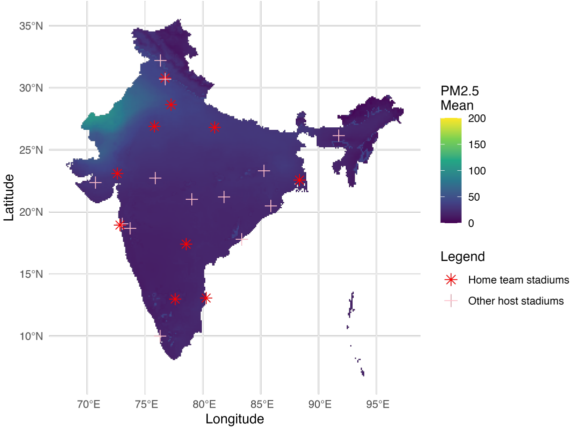


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: July

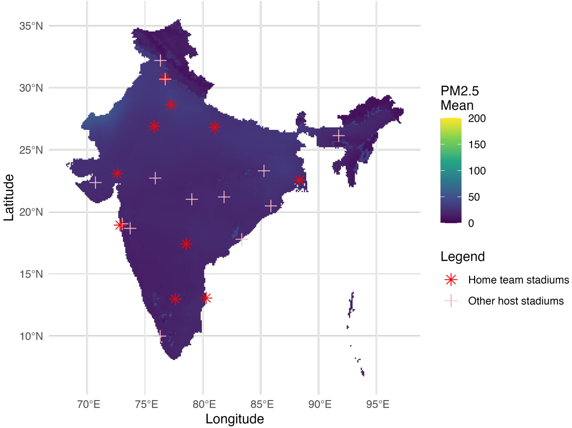


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: August

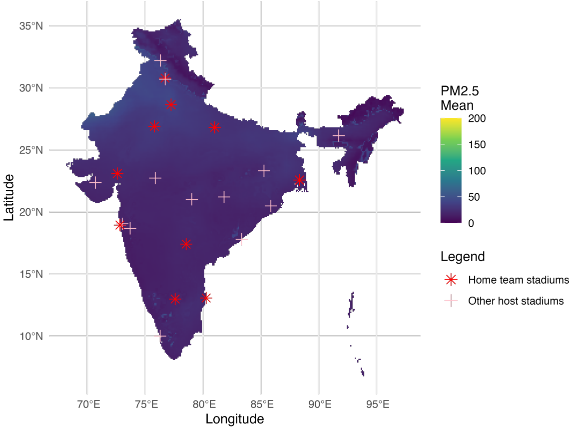


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: September

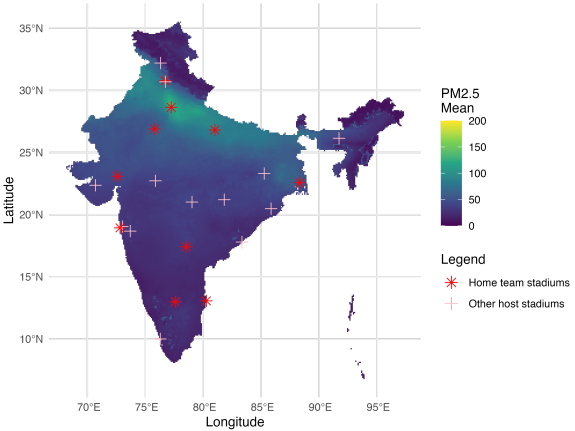


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: October

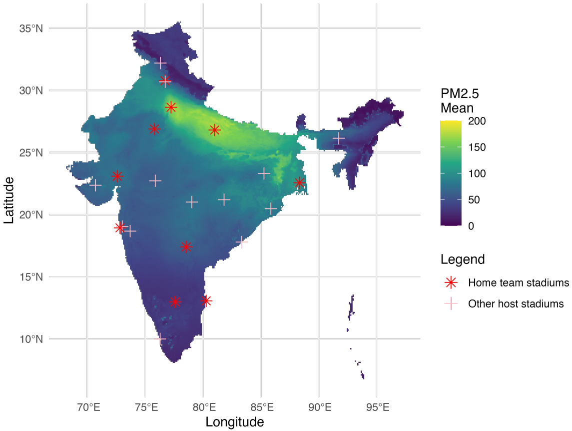


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: November

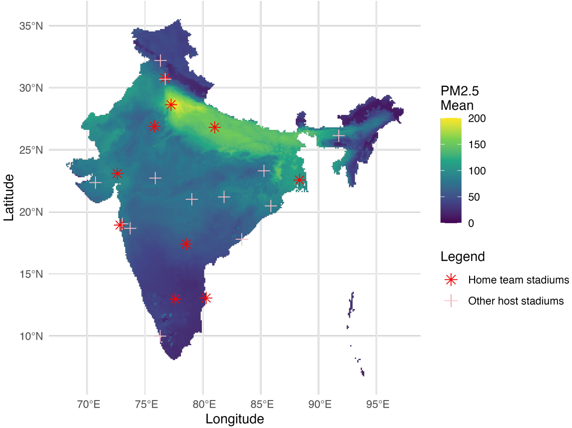


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: December

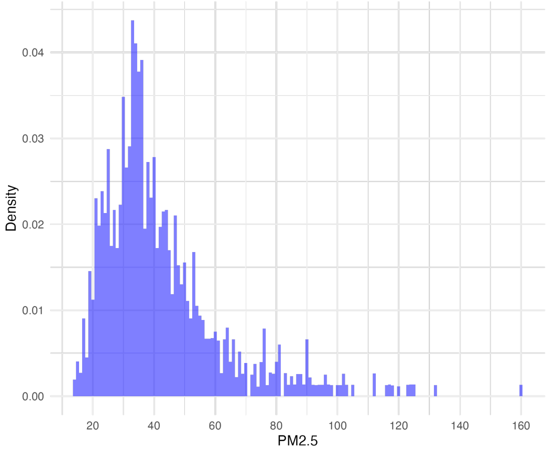


Notes. Monthly mean PM2.5 in 2019 and IPL stadiums.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

PM2.5 air pollution in India varies across space and time

Figure 4: IPL Game PM2.5 Distribution



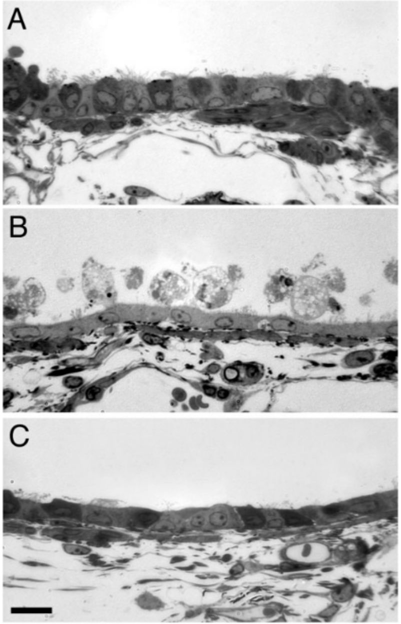
Notes. Distribution of PM2.5 at IPL games 2008-2022.

- India-specific ML algorithm to predict PM (Wang et al., 2024)
- Resolution: 10km × 10km daily gridded
- ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5)
- Validated with US Airnow monitors—R^2 ranges from 0.71-0.91 Scatter plots

Approach to measuring long-term PM2.5 exposure

- Measures of long-term exposure in the climate literature are typically place-based
- This makes sense for outcomes that are fixed in space, like agricultural yields
- But cricket players move:
  - Within a season (on the road for away games without returning home each time)
  - Across seasons (switching teams)
- Proposal: use mean PM2.5 exposure across all IPL games played by a given bowler
- Alternative definitions: Descriptive statistics Correlation matrix
  - Split panel in two, use bowler's mean from first half of panel in second half of panel (note: $\rho = 0.8$ between first half and full panel exposure)
  - Mean PM2.5 at bowler team's home-stadium during:
    - 3-week preseason training
    - Decade prior to IPL start (1998-2007)
    - Decade prior to IPL start, March-May (IPL months) only
    - Study period (2008-2022)
    - Study period (2008-2022), March-May (IPL months) only

Is there physiological evidence for adaptation to PM2.5 exposure?



This figure, reproduced from West et al. (2003), displays three panels of microscopic images of cells in the lungs of mice.

- **Panel A** shows cells from mice that were in the control group and breathed clean air throughout the experiment.
- **Panel B** shows cells from mice from the treatment group that was exposed to polluted air for one day, and had their cells imaged 24 hours after.
- **Panel C** shows cells from mice that were exposed to the same level of air pollution as those in Panel B, except they were exposed for 7 days instead of 1.

Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

Econometric specifications

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

- $R_{ij\ell t}$ = run scored (binary) on delivery t
- $\text{PM2.5}_{\ell d}$ = PM2.5 on day d of match at location ℓ (units: 10 $\mu g/m^3$)
- PM2.5_{j0} = PM2.5 for bowler j across all games in IPL (units: 10 $\mu g/m^3$)
- ψ_j = bowler fixed effects; ϕ_i = batter fixed effects; $\delta_{\ell y}$ = stadium-by-year fixed effects; θ_n = innings fixed effects; η_o = over fixed effects
- $\mathbf{X}_{\ell d}$ = weather controls (temperature, humidity, pressure, precipitation, wind)

Econometric specifications

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

$$R_{ij\ell t} = \sum_{k=2}^{5} \beta_k Q_k \left(\text{PM2.5}_{\ell d}\right) + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

- $R_{ij\ell t}$ = run scored (binary) on delivery t
- $\text{PM2.5}_{\ell d}$ = PM2.5 on day d of match at location ℓ (units: 10 $\mu g/m^3$)
- PM2.5_{j0} = PM2.5 for bowler j across all games in IPL (units: 10 $\mu g/m^3$)
- ψ_j = bowler fixed effects; ϕ_i = batter fixed effects; $\delta_{\ell y}$ = stadium-by-year fixed effects; θ_n = innings fixed effects; η_o = over fixed effects
- $\mathbf{X}_{\ell d}$ = weather controls (temperature, humidity, pressure, precipitation, wind)

Econometric specifications

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

$$R_{ij\ell t} = \sum_{k=2}^{5} \beta_k Q_k \left(\text{PM2.5}_{\ell d}\right) + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

$$R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t}$$

- $R_{ij\ell t}$ = run scored (binary) on delivery t
- $\text{PM2.5}_{\ell d}$ = PM2.5 on day d of match at location ℓ (units: 10 $\mu g/m^3$)
- PM2.5_{j0} = PM2.5 for bowler j across all games in IPL (units: 10 $\mu g/m^3$)
- ψ_j = bowler fixed effects; ϕ_i = batter fixed effects; $\delta_{\ell y}$ = stadium-by-year fixed effects; θ_n = innings fixed effects; η_o = over fixed effects
- $\mathbf{X}_{\ell d}$ = weather controls (temperature, humidity, pressure, precipitation, wind)

Econometric specifications

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \tag{1}$$

$$R_{ij\ell t} = \sum_{k=2}^{5} \beta_k Q_k \left(\text{PM2.5}_{\ell d}\right) + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \tag{2}$$

$$R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \tag{3}$$

Identification in (1) and (2):

- Identification comes from differences in PM2.5 across games for a given bowler in the same stadium throughout the year after controlling for the batter's average performance, strategy that could be influenced by the stage of game, and weather
- Identifying assumption: this residual variation in PM2.5 is not correlated with any factors that could also affect bowler performance

Econometric specifications

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \quad (1)$$

$$R_{ij\ell t} = \sum_{k=2}^{5} \beta_k Q_k \left(\text{PM2.5}_{\ell d}\right) + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \quad (2)$$

$$R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \quad (3)$$

Identification in (3)

- Identification comes from comparing bowlers who face the same match pollution but have differing long-run exposures to pollution
- Identifying assumption: this variation in long-run exposure to pollution is not correlated with any factors that affect bowler performance

Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

Table 1: PM2.5 exposure and run-scoring probability

Full table | Unforced errors | Hist. w/quintiles | Hist. w/residual PM2.5

	(1)	(2)
	$\mathbb{1}$ (At least one run scored)	
Match-day PM2.5	0.0041**	
	(0.0017)	
Q2 (Match-day PM2.5)		0.0074
		(0.0060)
Q3 (Match-day PM2.5)		0.010
		(0.0069)
Q4 (Match-day PM2.5)		0.014
		(0.0086)
Q5 (Match-day PM2.5)		0.028***
		(0.0099)
Weather controls	✓	✓
All FE	✓	✓
N	183,556	183,556

S.e.'s clustered two-way by match & bowler. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Interpretation

- 142.3 of 237.5 deliveries per game (60%) result in a run
- 0.41 ppt $\uparrow$ in run-scoring probability per delivery $\Longrightarrow$
  - 1.0% $\uparrow$ in run-scoring
  - $\approx 1 \uparrow$ runs/game
  - $\approx 0.5 \uparrow$ runs/team
  - $\approx 0.1 \uparrow$ runs/bowler
- Average win-margin is 16.9 runs (median 8.0)

Spline specification

- Linear spline model with bowler fixed effects with knots at WHO thresholds

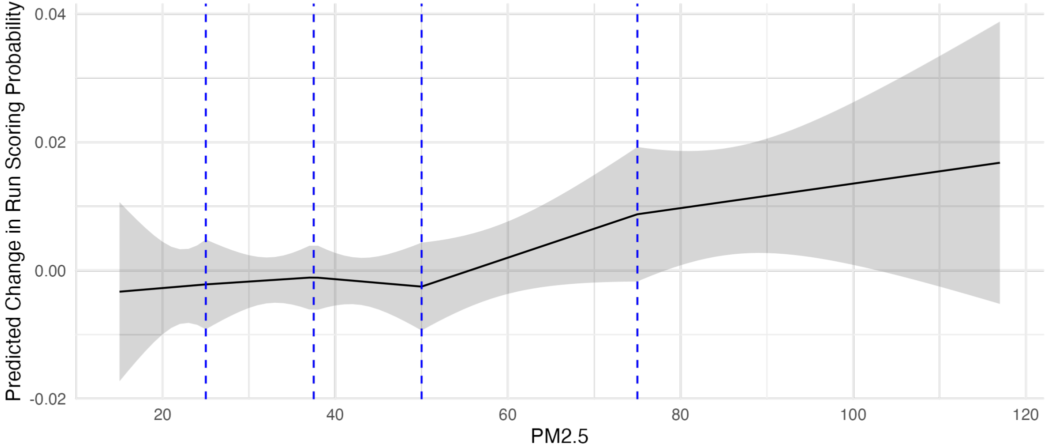
$$R_{ij\ell t} = \beta_1 \mathrm{PM2.5}_{\ell d} + \sum_{k=1}^{4} \beta_k (\mathrm{PM2.5}_{\ell d} - \tau_k) \mathbf{1}\{\mathrm{PM2.5}_{\ell d} \geq \tau_k\} + \mathbf{X}'_{\ell d}\phi + \psi_j + \varepsilon_{ij\ell t} \quad (4)$$

  where

- $k \in \{25, 37.5, 50, 75\} \mu g/m^3$ location of knots (WHO target air pollution levels)

Effect is concentrated in highest PM levels

Figure 5: Linear Spline with $N = 4$ knots at WHO Thresholds



Notes. This figure reports the change in run scoring probability conditional on PM2.5 relative to the mean run scoring probability for a given bowler.

Table 2: Evidence of Acclimatization to Air Pollution Robustness

	(1)	(2)
	$\mathbb{1}$ (At least one run scored)	
Match PM2.5	0.0041**	0.013**
	(0.0017)	(0.0052)
Match PM2.5 X Bowler PM2.5		-0.0019*
		(0.0011)
Weather controls	✓	✓
All FE	✓	✓
N	183,556	183,556
R^2	0.052	0.052

Notes. Standard errors clustered two-way by match and bowler.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

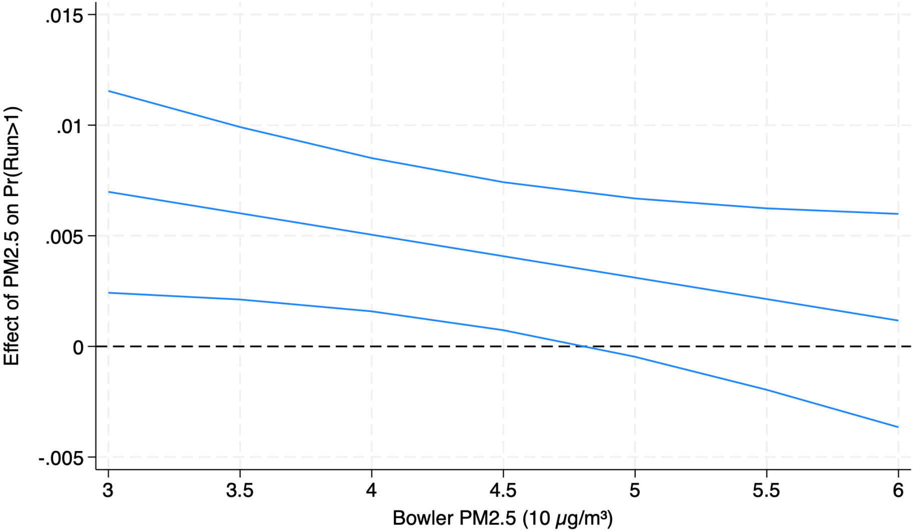
Interpretation

Table 3: Marginal effects of PM2.5 at different percentiles of baseline exposure P_{j0}

Percentile of P_{j0}	**Value of P_{j0}**	**Marginal Effect $\beta_1 + \beta_2 \times P_{j0}$**
1	3.0	$0.013 + (-0.0019) \times 3 = 0.0073$
50	4.2	$0.013 + (-0.0019) \times 4.2 = 0.0050$
99	5.9	$0.013 + (-0.0019) \times 5.9 = 0.0018$

- PM2.5 impairs performance for everyone
- But the effect is 64% smaller ($\frac{0.0050-0.0018}{0.0050} = 0.64$) for bowlers in the 99th percentile of exposure compared to those with median exposure histories

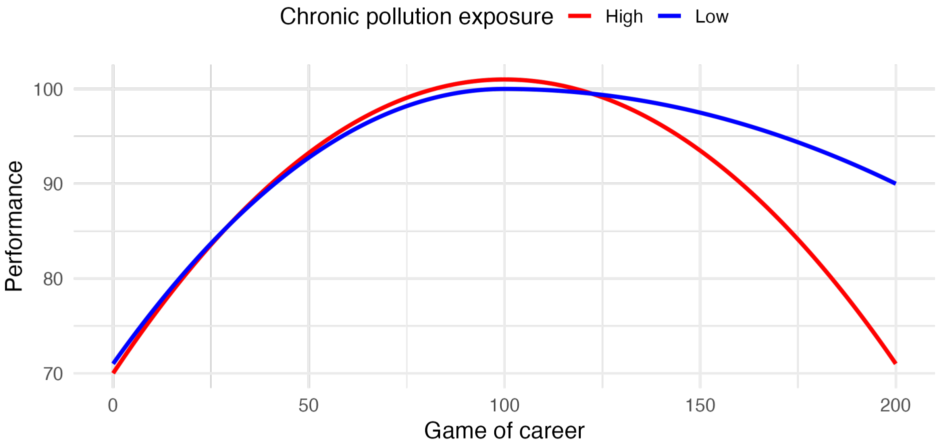
Figure 6: Marginal Effect of PM2.5 on Pr(Run) by Bowler Long-term PM2.5



Notes. This figure plots the marginal effect of PM2.5 on the likelihood of run-scoring by varying levels of $PM2.5_{j0}$. PM2.5 histogram

What is the downside to chronic pollution exposure?

Figure 7: Performance degradation across career trajectory



Notes. This is simulated data showing plausible career trajectories for bowlers with high and low chronic exposure to air pollution.

Outline

1. Motivation
2. Research question
3. Empirical setting and data
4. Econometric specifications
5. Preliminary results
6. Discussion

Discussion and next steps

Summary

- Pollution negatively impacts performance
- The effect is non-linear, with higher pollution impacting performance more
- Long-term exposure appears to mitigate—but not eliminate—these impacts

Next steps

- Explore alternative definitions of long-term exposure
- Explore performance degradation over time during career (instead of game-specific performance)
- Explore number of games exposed above certain threshold (may see parabolic shape to optimal exposure)

Adhvaryu, A., N. Kala, and A. Nyshadham (2022): "Management and Shocks to Worker Productivity," *Journal of Political Economy*, 130, 1–47, publisher: The University of Chicago Press.

Archsmith, J., A. Heyes, and S. Saberian (2018): "Air Quality and Error Quantity: Pollution and Performance in a High-Skilled, Quality-Focused Occupation," *Journal of the Association of Environmental and Resource Economists*, 5, 827–863.

Behrer, A. P., R. Choudhary, and D. Sharma (2023): *Air Pollution Reduces Economic Activity: Evidence from India*, World Bank, Washington, DC.

Borgschulte, M., D. Molitor, and E. Y. Zou (2022): "Air Pollution and the Labor Market: Evidence from Wildfire Smoke," *The Review of Economics and Statistics*, 1–46.

Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016): "Particulate Pollution and the Productivity of Pear Packers," *American Economic Journal: Economic Policy*, 8, 141–169.

Chung, S., C. Persico, and J. Liu (2025): "The Effects of Daily Air Pollution on Students and Teachers," .

Dechezleprêtre, A., N. Rivers, and B. Stadler (2019): "The economic cost of air pollution: Evidence from Europe," Tech. rep.

Fu, S., B. Viard, and P. Zhang (2017): "Air Quality and Manufacturing Firm Productivity: Nationwide Estimates for China," .

Guo, M. and S. Fu (2019): "Running With a Mask? The Effect of Air Pollution on Marathon Runners' Performance," *Journal of Sports Economics*, 20, 903–928, publisher: SAGE Publications.

Hansen-Lewis, J. (2024): "Air Pollution and Firm Productivity: Understanding the Roles of Technology and Adaptation," *Unpublished Manuscript*.

He, J., H. Liu, and A. Salvo (2019): "Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China," *American Economic Journal: Applied Economics*, 11, 173–201.

Hill, A. E., J. Burkhardt, J. Bayham, K. O'Dell, B. Ford, E. V. Fischer, and J. R. Pierce (2024): "Air pollution, weather, and agricultural worker productivity," *American Journal of Agricultural Economics*, 106, 1329–1353, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajae.12439.

Hoffmann, B. and J. P. Rud (2024): "The Unequal Effects of Pollution on Labor Supply," *Econometrica*, 92, 1063–1096, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA20484.

Holub, F. and M. Schündeln (2023): "Pro-environmental voting when climate change is made salient: Evidence from high-resolution flooding data," *PLOS Climate*, 2, e0000219, publisher: Public Library of Science.

Kahn, M. E. and P. Li (2020): "Air pollution lowers high skill public sector worker productivity in China," *Environmental Research Letters*, 15, 084003, publisher: IOP Publishing.

Lichter, A., N. Pestel, and E. Sommer (2017): "Productivity effects of air pollution: Evidence from professional soccer," *Labour Economics*, 48, 54–66.

Merfeld, J. D. (2023): "Air Pollution and Agricultural Productivity in a Developing Country," *IZA Discussion Paper*.

Mo, J., Z. Wu, and Y. Yuan (2023): "Air pollution kills competition: Evidence from eSports," *Journal of Environmental Economics and Management*, 122, 102886.

Mullins, J. T. (2018): "Ambient air pollution and human performance: Contemporaneous and acclimatization effects of ozone exposure on athletic performance," *Health Economics*, 27, 1189–1200, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.3667.

Muñoz Sabater, J. (2019): "ERA5-Land monthly averaged data from 1950 to present," Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Available at www.doi.org/10.24381/CDS.68D2BB30 (accessed on November 16, 2024).

Palacios-Huerta, I. (2025): "The Beautiful Dataset," *Journal of Economic Literature*.

Rentschler, J. and N. Leonova (2023): "Global air pollution exposure and poverty," *Nature Communications*, 14, 4432, publisher: Nature Publishing Group.

Wang, S., M. Zhang, H. Zhao, P. Wang, S. H. Kota, Q. Fu, C. Liu, and H. Zhang (2024): "Reconstructing long-term (1980–2022) daily ground particulate matter concentrations in India (LongPMInd)," *Earth System Science Data*, 16, 3565–3577, publisher: Copernicus GmbH.

West, J. A. A., L. S. Van Winkle, D. Morin, C. A. Fleschner, H. J. Forman, and C. G. Plopper (2003): "Repeated inhalation exposures to the bioactivated cytotoxicant naphthalene (NA) produce airway-specific Clara cell tolerance in mice," *Toxicological Sciences: An Official Journal of the Society of Toxicology*, 75, 161–168.

What can sports do for economics?

“Natural experiments, clean observability, precise measurement, high stakes, expert subjects, unimaginable detail, large datasets, exogenous rule changes, quasi-experimental variation, observable social effects, and no Hawthorne effects. These and other desirable attributes for empirical work are found in sports settings.” (Palacios-Huerta, 2025)

Three candidate datasets for air pollution

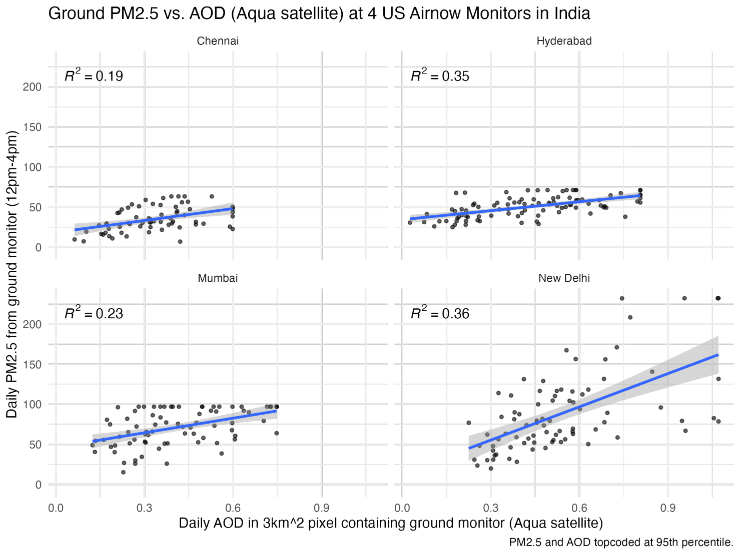
1. US Airnow ground monitors
   - Pro: high quality, available 2008-2024, hourly, rarely missing
   - Con: only in 5 cities (IPL in 20 stadiums)
2. MODIS AOD
   - Pro: available in all stadiums, across entire study time period
   - Con: low correlation with daily PM2.5 levels
3. Local reconstruction (Wang et al., 2024)
   - India-specific reconstruction dataset (Wang et al., 2024)
   - Resolution: 10km $\times$ 10km daily gridded Map
   - ML (gradient boosting decision tree) model incorporating ground monitors, remotely sensed products (MERRA-2, ERA5) Input table
   - Validated with US Airnow monitors—R^2 ranges from 0.71-0.91

Figure 8: ERA5, MERRA-2, and ground observation data used in Wang et al. (2024) Back

Type	Variable	Description	Spatial resolution	Temporal resolution
ERA5	SSRD	Surface solar radiation	$0.1^\circ \times 0.1^\circ$	Hourly
	BLH	Boundary layer height	$0.25^\circ \times 0.25^\circ$	Hourly
	EVAP	Evaporation	$0.1^\circ \times 0.1^\circ$	Hourly
	TEMP2	2 m air temperature	$0.1^\circ \times 0.1^\circ$	Hourly
	DEWP2	2 m dew point temperature	$0.1^\circ \times 0.1^\circ$	Hourly
	SP	Surface pressure	$0.1^\circ \times 0.1^\circ$	Hourly
	TPREC	Total precipitation	$0.1^\circ \times 0.1^\circ$	Hourly
	TCLOUD	Total cloud cover	$0.25^\circ \times 0.25^\circ$	Hourly
	UWIND10	10 m u component of wind	$0.1^\circ \times 0.1^\circ$	Hourly
	VWIND10	10 m v component of wind	$0.1^\circ \times 0.1^\circ$	Hourly
MERRA-2	BCSMASS	Black carbon surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	OCSMASS	Organic carbon surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	DUSMASS25	Dust $PM_{2.5}$ surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	DUSMASS	Dust surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	SO2SMASS	Sulfur dioxide surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	SO4SMASS	Sulfate surface mass concentration	$0.5^\circ \times 0.625^\circ$	Hourly
	TOTEXTTAU	Total aerosol extinction [550 nm]	$0.5^\circ \times 0.625^\circ$	Hourly
Observation	$PM_{2.5}$, PM_{10}	Particulate matter	Point	Hourly

Notes. Source: Wang et al. (2024)

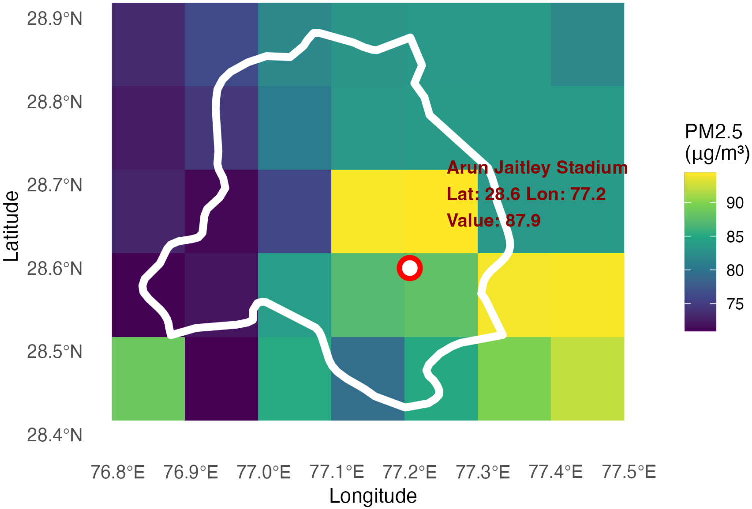
Figure 9: U.S. Airnow vs. MODIS AOD Back



Notes. This figure compares the reading of AOD from the MODIS Aqua satellite (passing over India at approximately 1:30pm) in a 3km radius around the location of USAirnow ground monitors in the 4 hour window centered at 1:30pm. Values are topcoded at the 95th percentile within a location.

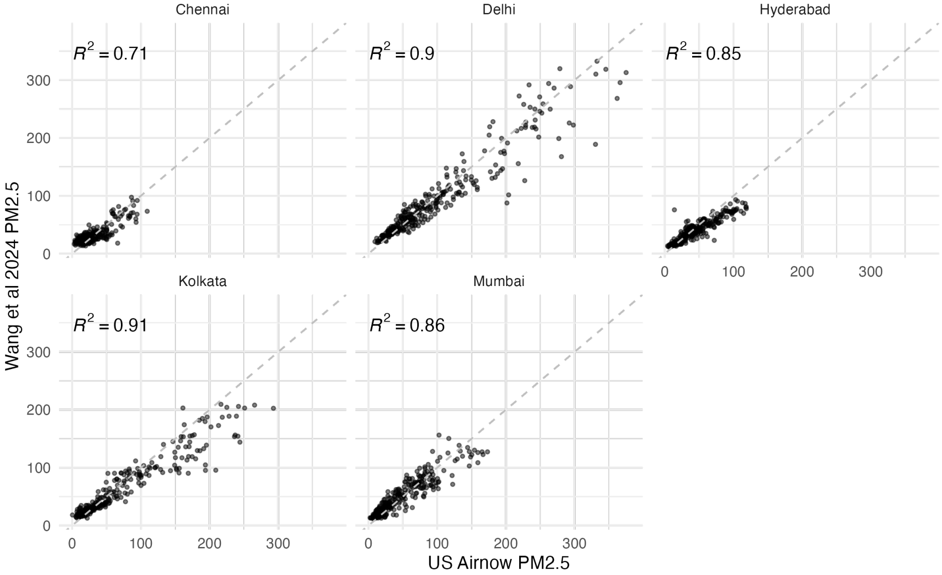
Figure 10: PM2.5 Raster Extraction Example - Arun Jaitley Stadium, Delhi on March 1, 2019

Back



Notes. This plot shows the raster of Wang et al. (2024) on March 1, 2019 in the vicinity of Delhi. Delhi's city boundaries are indicated in white and the location of the cricket stadium is encircled in red.

Figure 11: Wang et al. (2024) vs US Airnow PM2.5 in 2019 Back



Notes. This figure compares daily averages of PM2.5 concentrations from the LongPMInd model in Wang et al. (2024) with ground monitor readings from the U.S. Airnow network at U.S. consulates and embassies in India in the year 2019.

PM2.5 and weather

- Weather data: ERA-5 (Muñoz Sabater, 2019), daily 11km $\times$ 11km gridded
- Variables: temperature, relative humidity, atmospheric pressure, precipitation, solar radiation, wind speed
- Positive correlation between run-scoring and temperature/wind; negative correlation with humidity/precipitation Correlation matrix

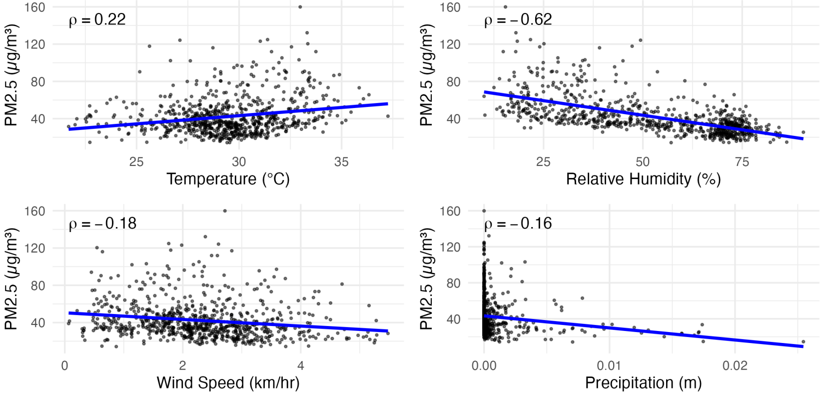
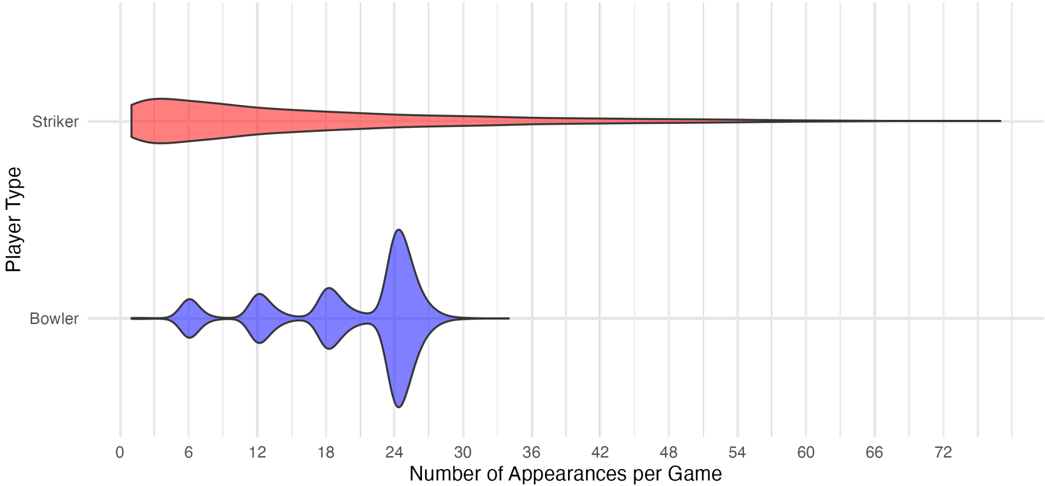


Table 4: PM2.5 and Weather Correlation Matrix Back

	Run	PM2.5	Temperature	Precipitation	Radiation	Wind	Humidity
Run	1	0.017	0.007	-0.002	0.003	0.001	-0.014
PM2.5	0.017	1	0.218	-0.153	0.230	-0.182	-0.619
Temperature	0.007	0.218	1	-0.254	0.232	0.234	-0.366
Precipitation	-0.002	-0.153	-0.254	1	-0.520	-0.070	0.253
Radiation	0.003	0.230	0.232	-0.520	1	0.108	-0.385
Wind	0.001	-0.182	0.234	-0.070	0.108	1	0.155
Humidity	-0.014	-0.619	-0.366	0.253	-0.385	0.155	1

Figure 12: Number of balls per game Back



Notes. This figure displays the count of number of balls per game for each type of player.

Back

Table 5: Summary Statistics of Long-run PM2.5 Variables

Definition	Distinct values	Mean	S.d.	Min	1^{st} perc.	99^{th} perc.	Max
Bowler PM2.5[a]	445	42.62	9.66	17.83	23.29	77.31	103.16
Bowler PM2.5[b]	107	40.35	7.58	22.76	25.20	63.17	69.13
Team stadium PM2.5[c]	13	46.94	18.33	25.11	25.11	83.59	83.59
Team stadium PM2.5[d]	13	51.66	20.71	28.27	28.27	96.23	96.23
Team stadium PM2.5[e]	13	47.54	15.85	26.66	26.66	77.70	77.70
Team stadium PM2.5[f]	13	43.40	15.21	22.28	22.28	70.49	70.49

[a] All IPL seasons. [b] IPL seasons 2008-2014. [c] 1998-2007. [d] 2008-2022.
[e] 2008-2022, IPL months. [f] 1998-2007, IPL months.

Notes. This table reports summary statistics for six alternative definitions of long-run PM2.5.

Alternative measures of long-run stadium PM2.5 exposure are highly correlated.

Figure 13: Correlation Matrix of Various Definitions of Long-term PM2.5

Back

	1998-2007	2008-2022	2008-2022 (Mar-May)	1998-2007 (Mar-May)
1998-2007	1.00	0.99	0.95	0.92
2008-2022		1.00	0.93	0.90
2008-2022 (Mar-May)			1.00	0.99
1998-2007 (Mar-May)				1.00

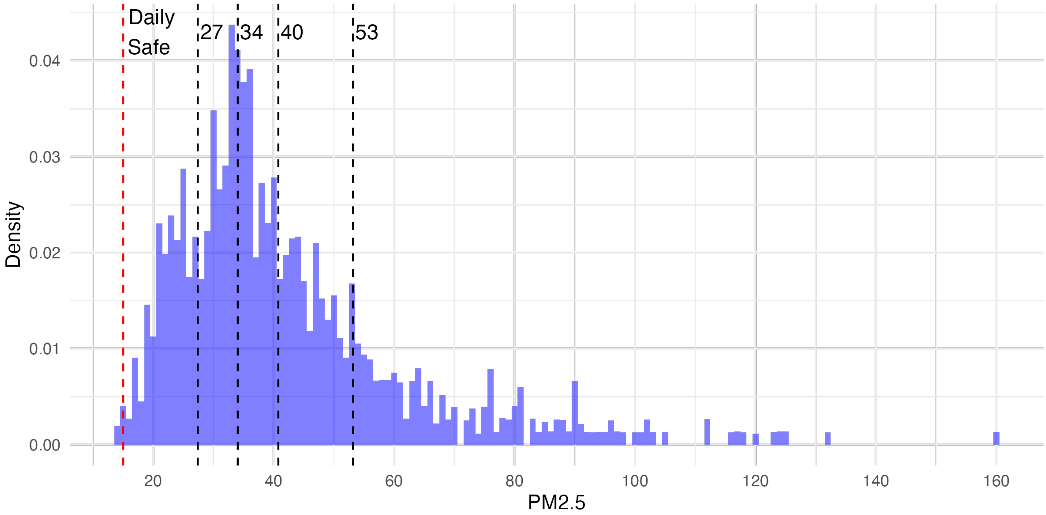
Notes. This plot is a correlation matrix of 4 measures of long-run stadium PM2.5 exposure.

Table 6: Match PM2.5 exposure and run-scoring probability Back

	(1)	(2)	(3)	(4)	(5)	(6)
			$\mathbb{1}$ (At least one run scored)			
Match-day PM2.5	0.0041***	0.0027*	0.0041**			
	(0.0010)	(0.0014)	(0.0017)			
Q2 (Match-day PM2.5)				0.0027	0.0038	0.0074
				(0.0064)	(0.0069)	(0.0060)
Q3 (Match-day PM2.5)				0.00015	-0.00065	0.010
				(0.0065)	(0.0073)	(0.0069)
Q4 (Match-day PM2.5)				0.0086	0.0047	0.014
				(0.0076)	(0.0096)	(0.0086)
Q5 (Match-day PM2.5)				0.023***	0.017	0.028***
				(0.0069)	(0.010)	(0.0099)
Weather controls		✓	✓		✓	✓
All FE			✓			✓
N	183,572	183,572	183,556	183,572	183,572	183,556

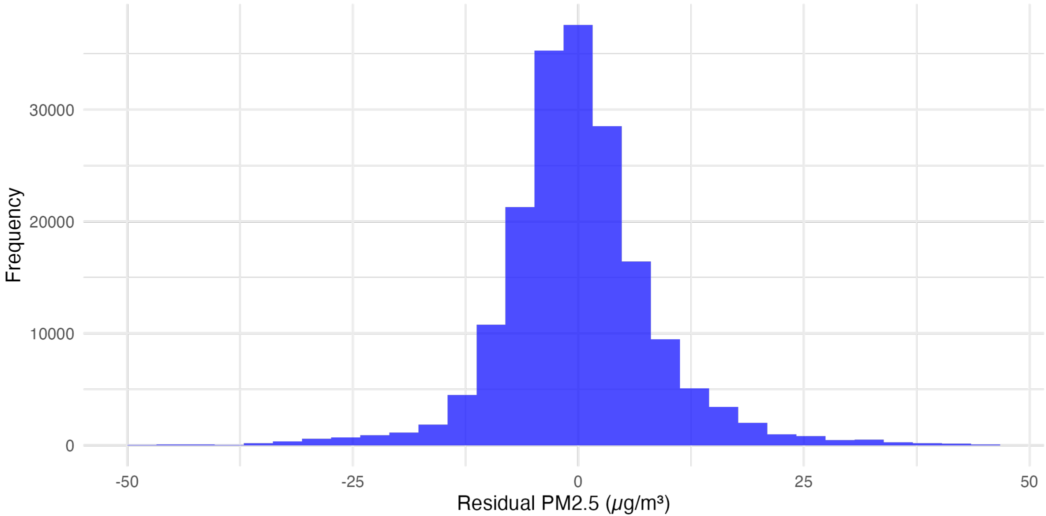
Notes. Standard errors clustered two-way by match and bowler.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 14: Histogram of Game PM2.5 with Quantiles Indicated Back



Notes. This figure displays the distribution of game PM2.5 with quantiles and WHO safe daily limit (15 mug/m^3) indicated.

Figure 15: Histogram of Residual PM2.5 Back



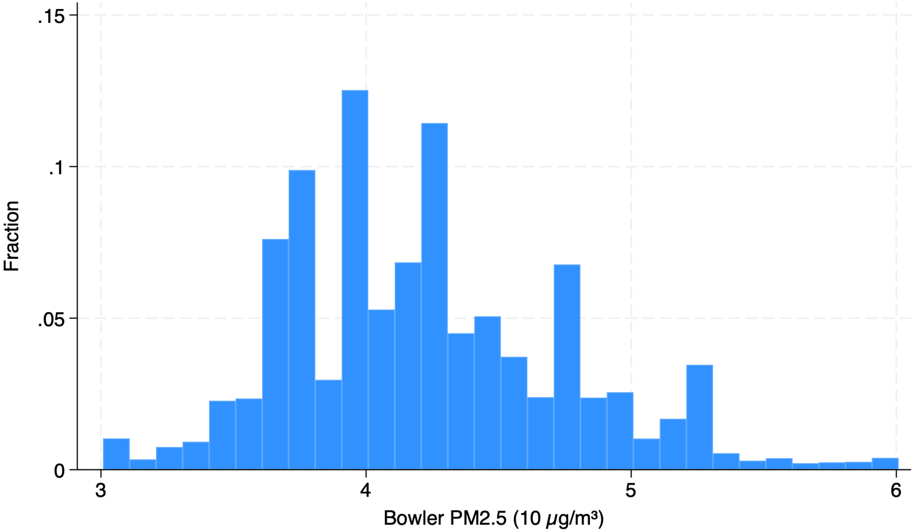
Notes. This figure displays a histogram of residual PM2.5 after controlling for all weather controls and fixed effects in Equation (1).

Table 7: PM2.5 exposure and unforced error probability Back

	(1)	(2)	(3)	(4)	(5)	(6)
			$\mathbb{1}$ (At least one run scored)			
Match-day PM2.5	-0.000077	0.00087*	0.00039			
	(0.00035)	(0.00050)	(0.00057)			
Q2 (Match-day PM2.5)				0.00094	0.0013	-0.00059
				(0.0019)	(0.0020)	(0.0019)
Q3 (Match-day PM2.5)				-0.0020	-0.00039	-0.0026
				(0.0020)	(0.0022)	(0.0021)
Q4 (Match-day PM2.5)				0.00017	0.0031	0.00070
				(0.0022)	(0.0026)	(0.0026)
Q5 (Match-day PM2.5)				-0.0017	0.0031	0.00043
				(0.0021)	(0.0028)	(0.0027)
Weather controls		✓	✓		✓	✓
All FE			✓			✓
N	183,572	183,572	183,556	183,572	183,572	183,556

Notes. Standard errors clustered two-way by match and bowler.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 16: Histogram of Bowler PM2.5 Mean in IPL Games 2008-2022 Back



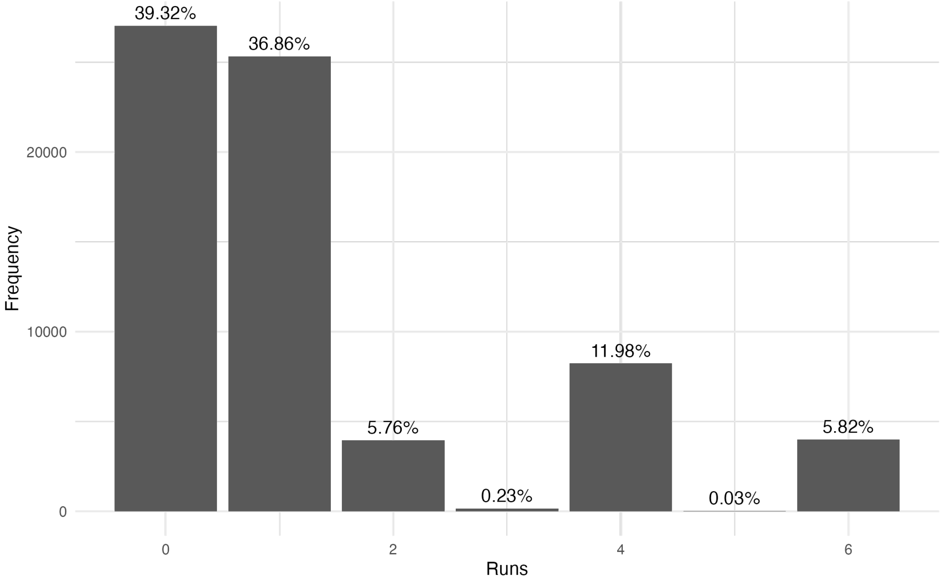
Notes. This figure displays the distribution of bowler mean PM2.5 (truncated to 1st through 99th percentiles) defined as the mean PM2.5 exposure across all games for a given bowler in the IPL 2008-2022.

Table 8: Robustness of Long-term PM2.5 Measures Back

	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$ (At least one run scored)					
Match PM2.5	0.0052	0.0067**	0.0062**	0.0067**	0.0058**	0.0071***
	(0.0060)	(0.0030)	(0.0025)	(0.0028)	(0.0024)	(0.0027)
PM2.5 X Bowler PM2.5 (IPL Seasons 2008-2014)	-0.0012					
	(0.0013)					
PM2.5 X Bowler Team Stadium PM2.5 (Preseason)		-0.00043				
		(0.00046)				
PM2.5 X Bowler Team Stadium PM2.5 (1998-2007)			-0.000038			
			(0.000038)			
PM2.5 X Bowler Team Stadium PM2.5 (1998-2007, IPL Months)				-0.000049		
				(0.000046)		
PM2.5 X Bowler Team Stadium PM2.5 (2008-2022)					-0.000026	
					(0.000032)	
PM2.5 X Bowler Team Stadium PM2.5 (2008-2022, IPL Months)						-0.000063
						(0.000048)
Weather controls	✓	✓	✓	✓	✓	✓
Bowler FE	✓					
Bowling Team-by-year FE		✓	✓	✓	✓	✓
All other FE	✓	✓	✓	✓	✓	✓
N	53,607	183,558	183,558	183,558	183,558	183,558

Notes. Standard errors clustered two-way by match and bowler. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 17: Histogram of Runs Back



Outline

Do bowlers perform worse as they are exposed to PM2.5 longer?

1. Theory: bowlers are more affected by air pollution than batters because they are exerting physical effort for longer than batters
    - $\implies$ **Testable implication:** probability of a bowler conceding a run weakly increases as they bowl more balls

2. Theory: bowlers who are more adapted to air pollution will exhibit a less pronounced decline in performance over time than ones who are less adapted
    - $\implies$ **Testable implication:** probability of conceding a run increases at a slower rate (slope is flatter) for bowlers from high-pollution areas

Rules for bowlers in IPL cricket

Terminology:

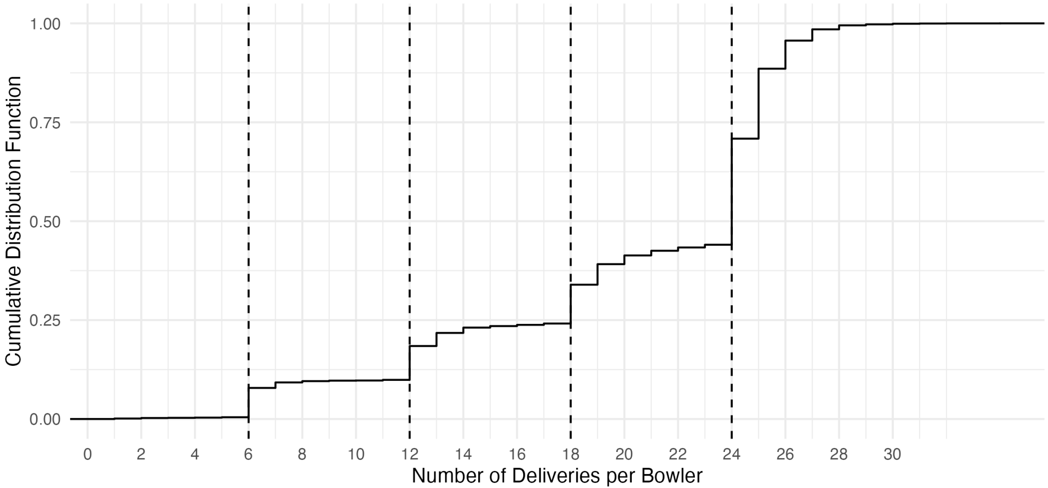
- Delivery: when a bowler bowls (pitches) the ball to the batter
- Illegal delivery: a ball that is delivered unfairly (i.e., no-ball, wide, dead ball)
- Extras: points that the striking team gets when the bowler delivers illegally
- Over: a series of 6 legal deliveries (note: illegal deliveries not counted)
- Change of ends: after an over, bowl from the opposite end of pitch
- Inning: consists of 20 overs; 2 innings per game

Rules:

- Bowlers can bowl up to 4 overs (maximum $4 \times 6 = 24$ *legal* deliveries)
- No bowler can bowl two consecutive overs (due to change of ends)
- Bowlers (typically) stop bowling within a match at the end of an over

Bowlers switch out after an over (6 deliveries).

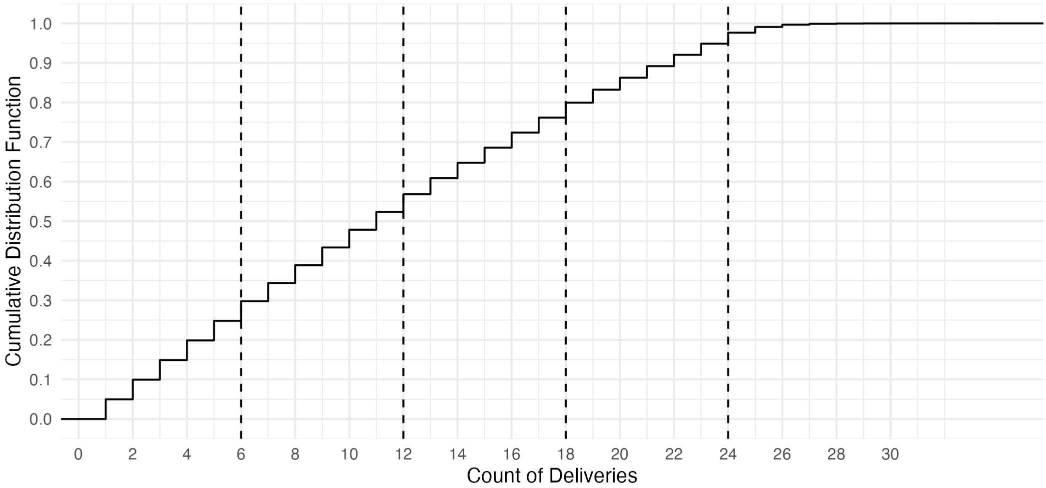
Figure 18: CDF of Number of Balls Bowled Per Bowler Within Game Histogram



Notes. Number of deliveries per bowler for 445 bowlers across 733 games in IPL. Bowlers are typically required to switch after an over (consisting of 6 deliveries); every 6 deliveries indicated with dashed line. Note that illegal deliveries (essentially a foul ball) are included in this graph, leading to some bowlers switching out after 6 deliveries (since illegal deliveries do not count toward the 6 deliveries per over.)

This results in 90%+ of bowler-ball obs. ≤ 24.

Figure 19: CDF of Count of Balls Bowled



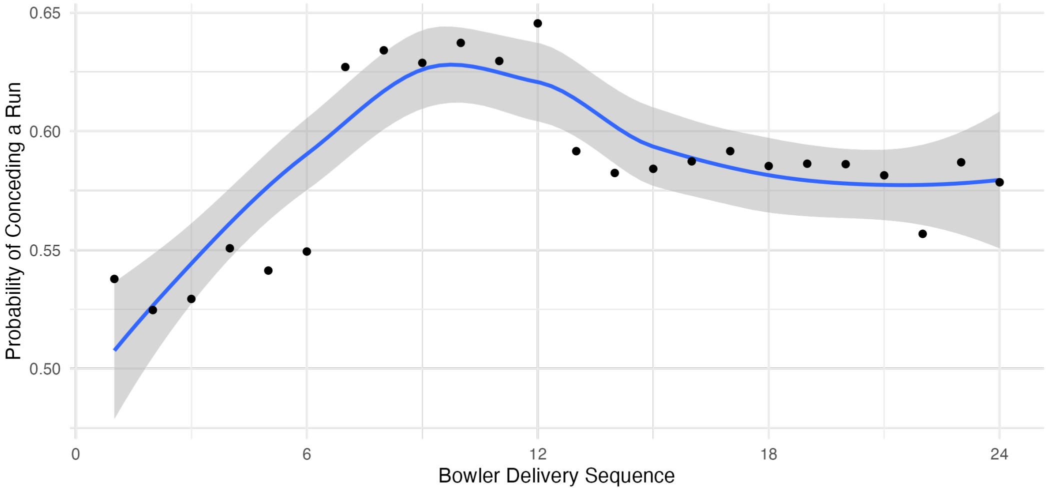
Notes. This figure plots the cumulative distribution function of number of deliveries each bowler makes within a game.

Takeaway

- Most bowlers bowl their maximum allotment of 4 overs (24 deliveries).
- A small fraction of deliveries ($< 10\%$) are beyond the 24th delivery—this is a result of illegal deliveries
- I analyze bowler performance over the first 24 deliveries to cover the vast majority of bowler performance
- I do not analyze bowler performance at deliveries > 24 since there is a smaller sample size and estimates are therefore noisier
- To account for composition effects (pulling out underperforming bowlers sooner), I restrict to bowlers who bowl all 4 of their overs
  - I also conduct this analysis with all bowlers without this restriction and the patterns are similar

Probability of conceding a run has an inverted-U shape.

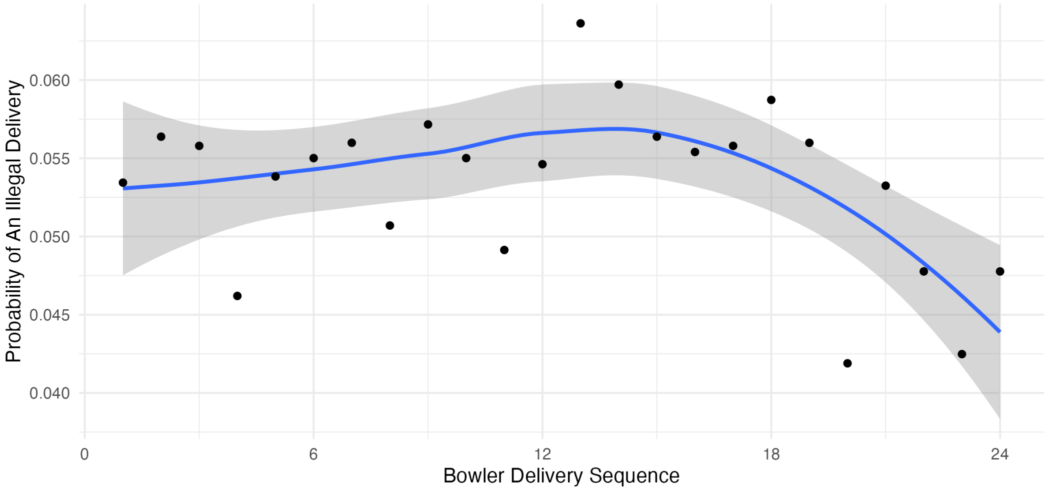
Figure 20: Probability of Conceding a Run by Delivery



Notes. This plot shows the mean probability of conceding a run for each of a bowler's deliveries.

Probability of an illegal delivery is roughly flat, but dips off after 18 deliveries.

Figure 21: Probability of an Illegal Delivery by Delivery



Notes. This plot shows the mean probability of an illegal delivery for each of a bowler's deliveries.

The probability of conceding a run has an inverted-U shape

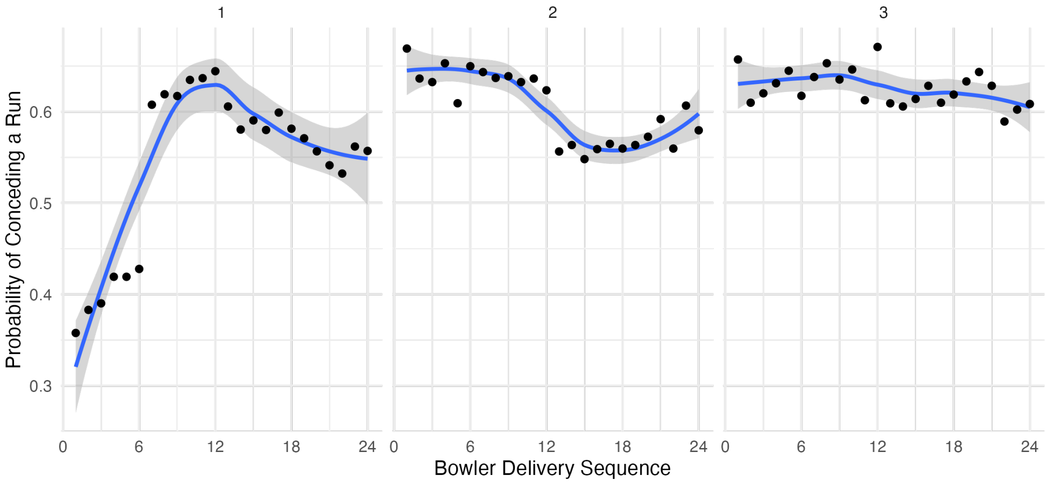
- It is not monotonically increasing
- Note: this is also the case when disaggregating by inning (though there is a level effect for second inning)
- It also appears for expected score Expected Score
- It also appears for the probability of scoring 6 ("home run") Six
- However, probability of scoring 4 is roughly flat then increases after 3 overs (unclear why) Four
- The probability of an illegal delivery is very small ($< 0.07\%$) across all deliveries
- It does not seem to increase as the bowler is on the field longer; in fact, it decreases at higher delivery numbers

There are factors other than bowler fatigue and PM2.5 exposure.

- batter strategy
- Bowler strategy (conservative, then more aggressive)
- Bowler composition effects - can rule this out due to restricting to bowlers who bowl their max allotment
- Stage of game effects - earlier in inning, batters are more conservative since don't want to get "out"
    - Examine this in next slide

Could stage of game be influencing strategy?

Figure 22: Probability of Conceding Run by Stage of Game



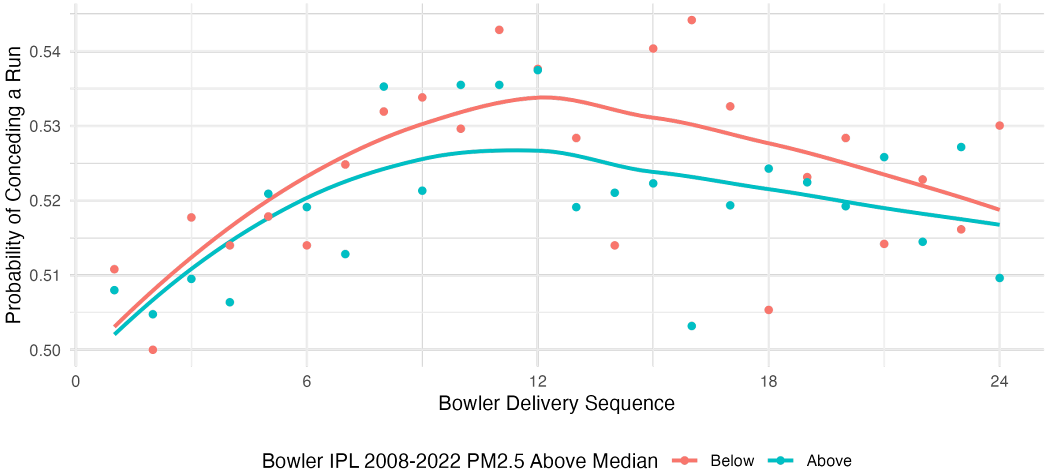
Notes. This plot shows the mean probability of conceding a run for each of a bowler's deliveries, disaggregated by when the bowler started bowling within an inning. Panel 1 shows bowlers who started in over 1-2; panel 2, 3-6; panel 3, 7-19. These categories are formed by dividing the bowler-starts into the 1-33%, 34-66%, and 67% + percentiles.

Are bowlers with higher long-run exposure PM2.5 less detrimentally affected?

- My initial hypothesis was that bowler performance would degrade over time within a game, perhaps in part a result of PM2.5 exposure
- This does not seem to be the case
- However, there may be differential patterns depending on bowler typical PM2.5 exposure

Bowlers from higher PM2.5 teams less likely to concede run.

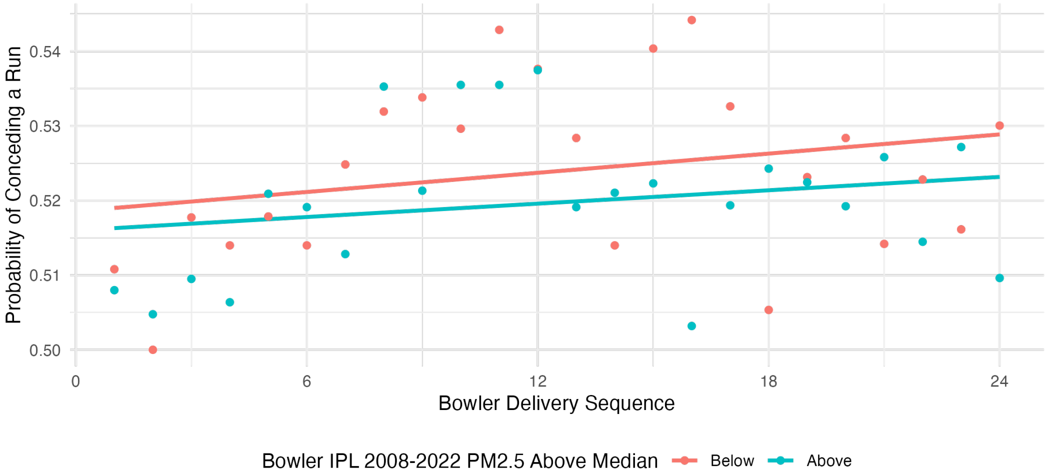
Figure 23: Probability of a Conceding Run by Delivery, Disaggregated



Notes. This plot shows the mean probability of conceding a run for each of a bowler's deliveries, disaggregated by whether bowler's mean PM2.5 exposure in IPL is below or above the median for all bowlers in IPL.

Bowlers from higher PM2.5 teams less likely to concede run (linear).

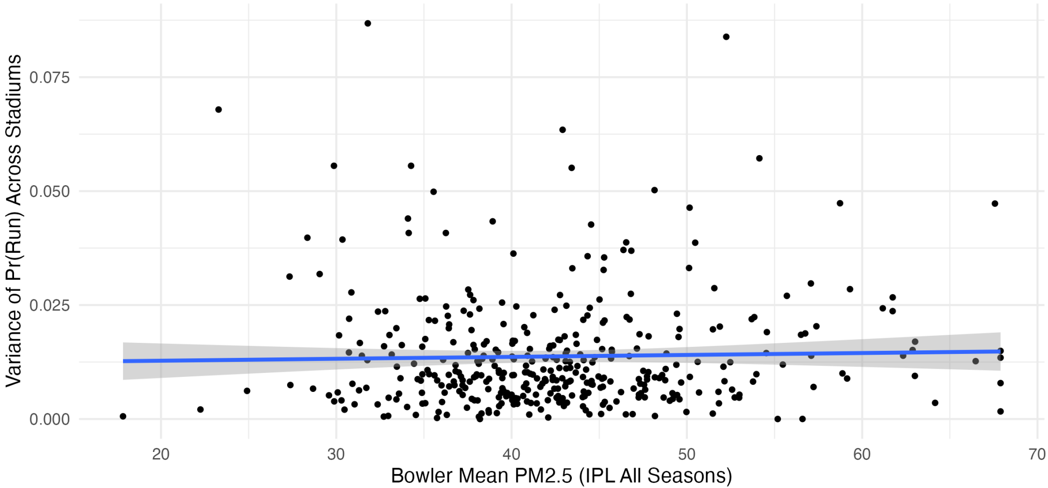
Figure 24: Probability of a Conceding Run by Delivery, Disaggregated



Notes. This plot shows the mean probability of conceding a run for each of a bowler's deliveries, disaggregated by whether bowler's mean PM2.5 exposure in IPL is below or above the median for all bowlers in IPL.

There does not appear to be a pattern across space.

Figure 25: Distribution of Bowler Heterogeneity Across Stadiums



Notes. For each bowler, this plot shows the variance of their performance across stadiums (defined as the variance of the binary outcome of conceding a run) by their typical PM2.5 exposure during the IPL.

Outline

Daily AOD correlates somewhat with daily PM2.5 (pixel)

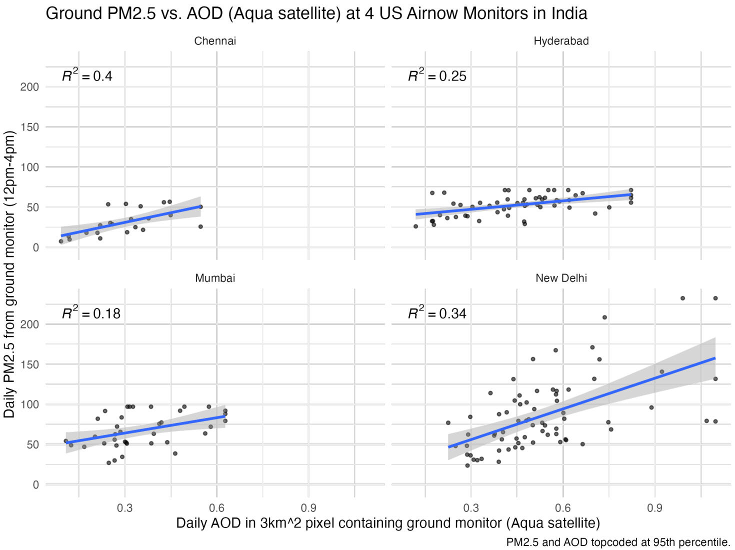


Figure 26: Ground PM2.5 vs. AOD

Daily AOD correlates somewhat with daily PM2.5 (3km radius)

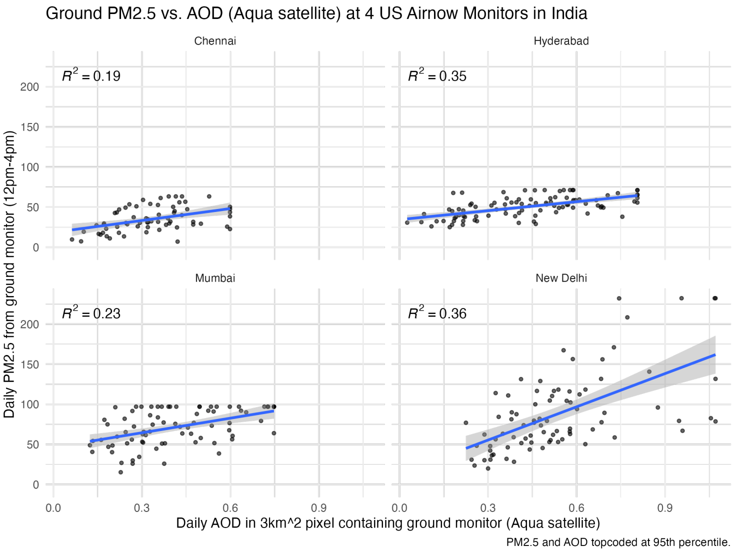


Figure 27: Ground PM2.5 vs. AOD

The agreement looks better when binning observations (pixel)

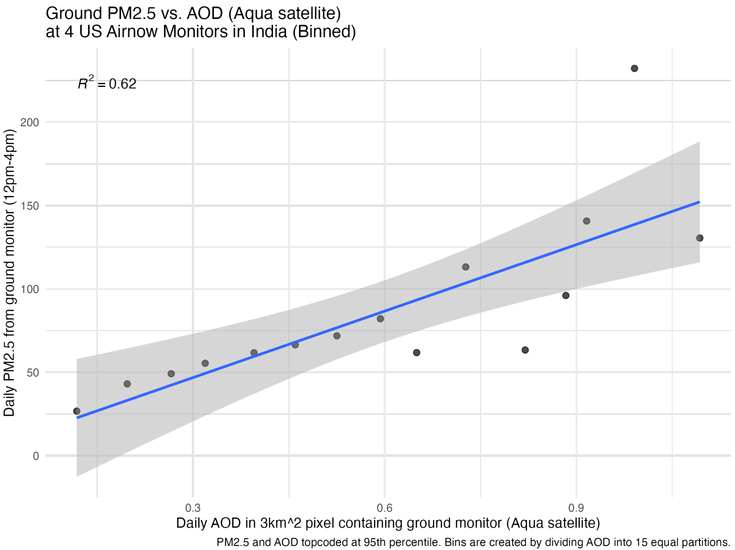


Figure 28: Ground PM2.5 vs. AOD

The agreement looks better when binning observations (3km radius)

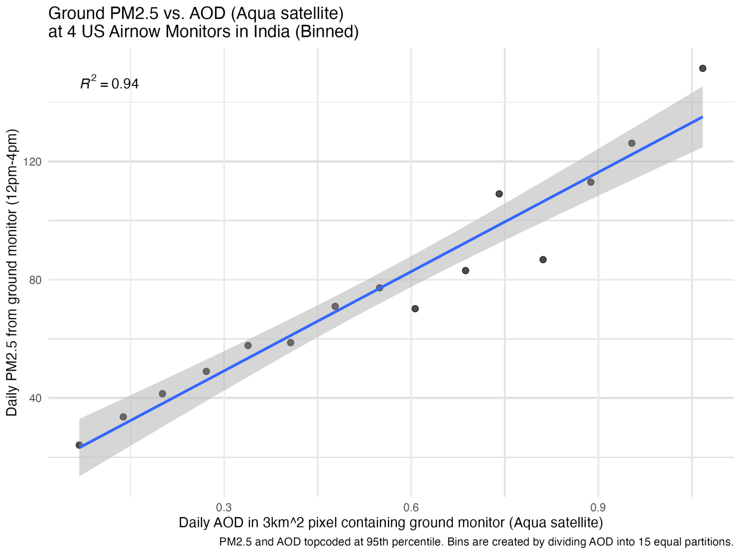


Figure 29: Ground PM2.5 vs. AOD

And even better when taking annual averages (R2 on par with G-C et al.)

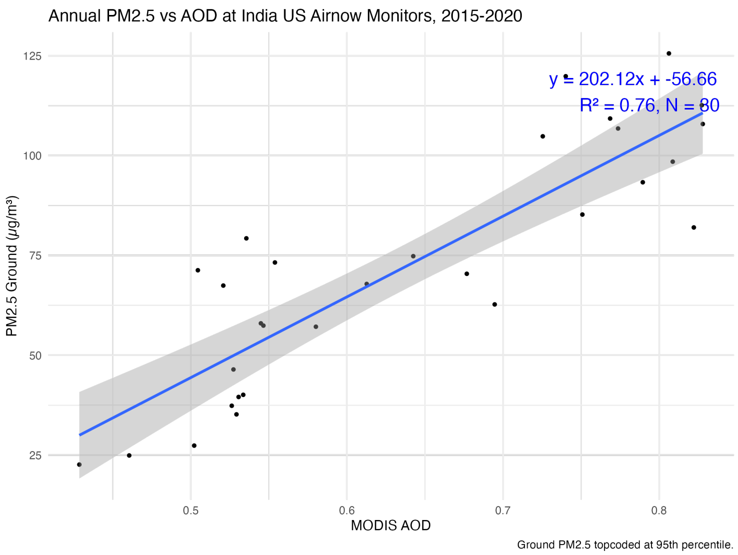
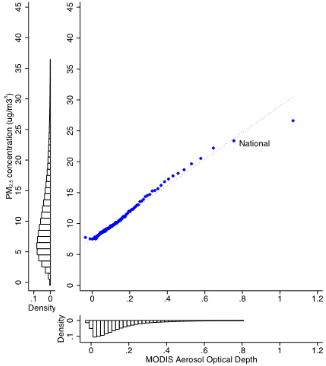


Figure 30: Ground PM2.5 vs. AOD

Compare to Zou paper

Figure D.6: PM2.5 and Aerosol Correlation, 2001-2013



Notes: This figure presents the correlation between monitor PM2.5 readings and the satellite aerosol optical depth measure, defined as the aerosol level within the 10km×10km area where the monitor lives in. Dots show average PM2.5 within 100 equally-sized aerosol bin. Histograms show density of the raw PM2.5 and aerosol data.

Figure 31: Ground PM2.5 vs. AOD

How I made the plots

1. I downloaded the .tif rasters of daily AOD for the Aqua and Terra satellites from Gendron-Carrier et al. 2022.
2. I used ground monitor readings from the US Airnow network in India, which has a high quality monitor at each US consulate and embassy in India. These monitors report measurements of PM2.5 each hour (there is some missingness but it is complete for the most part). The reason I chose to use these monitors is that these are what I have been using for my analysis so far since I am fairly confident about the quality of the data and they are also close to the cricket stadiums ($<$ 5km in most cases). The downside is that there are only 5 of them, whereas there are 24 cricket stadiums in India where I am planning to do my analysis. Hence the need for the AOD data. (In the plot, there are only 4 locations; that is because in the fifth location there were not enough days that had a reading of AOD and a ground reading).
3. I select the value of the AOD raster in the 3km x 3km pixel that contains the US Airnow monitor.

How I made the plots, cont.

1. I looked up what time of day the Aqua and Terra satellites pass over India. It is approximately 10:30am and 1:30pm respectively.
2. I created a ground monitor reading for Aqua and another for Terra using the 4 hours surrounding the time the respective satellite took a measurement over the region.
3. For both AOD and ground PM2.5, I topcode values at the 95th percentile by location.
4. I scatter (and regress) the ground monitor reading for the time period for the satellite on the satellite AOD reading. The R2s were higher in the regressions with readings from the Aqua satellite, so I am reporting those. This would likely make more sense for my context since cricket games are typically in the afternoon.
5. The plot disaggregated by monitor shows scatters of daily PM2.5 readings on daily readings from Aqua. The R2s range from 0.18 to 0.4.