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

Matthew S. Brooks†   Faraz Usmani‡

October 30, 2025

Please click here for the latest version of this paper.

Abstract

Air pollution harms labor productivity, yet little is known about whether workers adapt to chronic exposure. We address this question using performance data from India's premier cricket league, which provides exogenous variation in both acute pollution exposure and long-term exposure histories. We find that a 10 microgram per cubic meter increase in PM2.5 (equivalent to 50% of a standard deviation in our sample) reduces productivity by about 1 percent, with effects concentrated at extreme pollution levels that far exceed WHO guidelines. However, workers appear to adapt: performance among those with the highest levels of chronic exposure shows dramatically smaller responses to acute pollution episodes, with the most chronically exposed experiencing approximately 40 percent lower productivity losses than those with median exposure histories. Our findings suggest that standard estimates from low-pollution environments poorly capture the dynamics between short- and long-term exposure in high-pollution settings, with important implications for environmental regulation in developing economies where chronic exposure is widespread.

*We are grateful for feedback from colleagues at the 22nd Occasional Workshop in Environmental & Resource Economics, the Indian Statistical Institute in Delhi, University of Massachusetts - Amherst, the 2025 Association of Environmental and Resource Economists Summer Conference, the 25th Annual University of Colorado Boulder Environmental and Resource Economics Workshop and the Natural Resource Economics and Policy Lab at UC Davis. Any remaining errors are our own.

†Department of Agricultural and Resource Economics; University of California, Davis; One Shields Avenue, Davis, CA 95616; msbrooks@ucdavis.edu.

‡Global Energy Alliance for People and Planet; faraz.usmani@energyalliance.org.

1 Introduction

Over 7.3 billion people—94% of the global population—are exposed to unsafe annual levels of fine particulate matter (PM2.5) concentrations (Rentschler and Leonova, 2023). In addition to harming human health, PM2.5 exposure also impairs labor productivity, putting a drag on economic growth—especially in emerging economies where PM2.5 exposure tends to be highest. Although PM2.5 exposure is a problem globally, the level of exposure varies dramatically across space, with over one-third of the global population exposed to hazardous levels of PM2.5. Although an extensive literature has studied the productivity impacts of PM2.5 exposure in moderate pollution environments such as the United States— where PM2.5 levels hovered below 10 $\mu g\, m^{-3}$ in the past decade—we know comparatively little about the productivity impacts of PM2.5 at high-levels of exposure, which are representative for a third of the global population.

Understanding the effects of PM2.5 exposure at high-levels is complicated for two reasons. First, there may be non-linearities: a marginal increase in PM2.5 exposure at high-levels may be result in different productivity impacts than the same increase at lower levels. Second, there may be heterogeneity in the dose-response: the impact of increases in PM2.5 may be different for people who are chronically exposed to high-levels of PM2.5 than for those who are used to cleaner air. While several studies have begun to quantify non-linearity in the response to PM2.5 exposure (Hoffmann and Rud, 2024), the question of how accumulated exposure to PM2.5 affects the effect of current exposure is largely unexplored.

To fill this gap, this paper answers the question: "does the marginal effect of contemporaneous PM2.5 exposure on labor productivity vary by past exposure to PM2.5?" While existing research estimates average treatment effects of pollution on productivity, we investigate whether individuals accustomed to high pollution levels respond differently to acute pollution episodes than those typically exposed to cleaner air. We address this question using variation in match-day pollution levels and the assignment of cricket players to teams across stadiums with varying baseline pollution levels in the Indian Premier League (IPL).

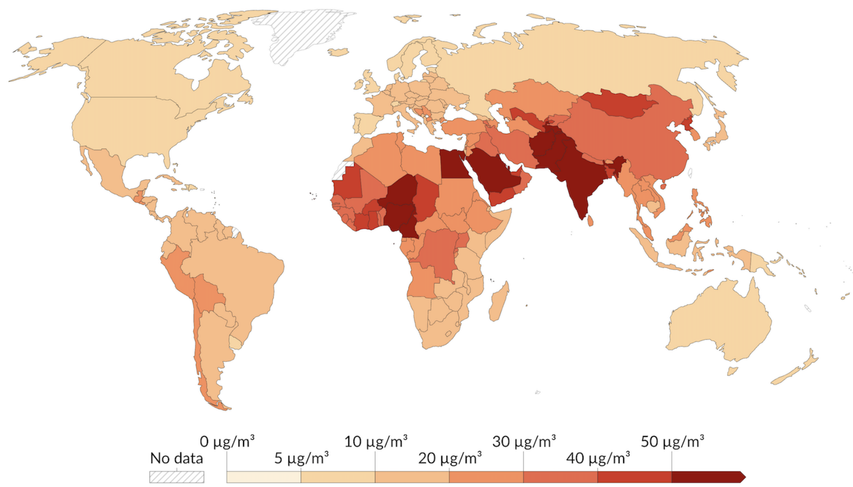
Our empirical strategy exploits the IPL's scheduling system, which exogenously assigns players to compete across different cities while maintaining fixed team bases. This design provides quasi-experimental variation in both short-term exposure (match-day pollution at various venues) and long-term exposure (a weighted average of baseline pollution at players' home stadiums and their various match locations). We combine ball-by-ball performance data from 773 IPL matches spanning 2008–2022 with a novel machine learning-based pollution dataset that provides daily PM2.5 estimates for all stadium locations (Wang et al., 2024). Our setting provides unusually rich variation: PM2.5 concentrations during matches average 42 $\mu g\, m^{-3}$ and reach as high as 160 $\mu g\, m^{-3}$—over ten times the World Health Organization (2021) daily safe limit and representative of conditions faced by billions of workers in developing countries.

We find that pollution has economically significant and heterogeneous effects on worker productivity that depend on both task characteristics and workers' adaptation history. A 10 $\mu g\, m^{-3}$ increase in PM2.5 increases bowlers' probability of conceding a run by 0.27–0.41 percentage points, with these non-linear effects concentrated in the highest pollution quintile. Importantly, we document evidence of adaptation: players with higher long-term pollution exposure show smaller responses to acute pollution episodes. The differential impact on bowlers versus batters likely reflects bowlers' greater physical exertion and longer field exposure.

Our setting addresses important gaps in the existing literature. Research on air pollution and productivity exhibits a striking geographic bias that limits our understanding of pollution impacts where they are more highest. A recent review by Aguilar-Gomez et al. (2022) finds that among ten major studies on pollution's physical productivity effects, half examine the U.S. or Europe (Archsmith et al., 2018; Chang et al., 2016; Graff Zivin and Neidell, 2012; Mullins, 2018), four study China (Chang et al., 2019; He et al., 2019; Guo and Fu, 2019; Kahn and Li, 2020), and only one examines India (Adhvaryu et al., 2022). This concentration of research in relatively low-pollution settings may miss important non-linearities in dose–

response relationships at the higher pollution levels that characterize much of the developing world (Figure 1).

Figure 1: Average Annual PM2.5 Concentrations in 2019



Notes. This figure displays annual average PM2.5 concentrations (population-weighted) in 2019. Source: World Health Organization - Global Health Observatory (2025).

Moreover, existing studies primarily focus on short-term pollution shocks while largely ignoring potential adaptation mechanisms in chronically polluted environments. For example, of the ten studies cited above, only one examines pollution exposure beyond a week (He et al., 2019). This gap is particularly problematic because adaptation responses may be crucial for understanding productivity impacts in persistently polluted environments typical of developing countries.

We make three main contributions to the literature on the productivity impacts of air pollution. First, we provide among the first estimates of pollution's productivity impacts in a genuinely high-pollution setting with reliable measurement. Previous studies of India have been constrained by sparse ground monitoring networks—during our study period, only 5 of 20 IPL stadiums had nearby air quality monitors. We overcome this limitation using a novel machine learning-based dataset that assimilates ground monitors and satellite data to predict air pollution levels across all of India (Wang et al., 2024). This India-specific model substantially outperforms commonly used satellite proxies like MODIS AOD, which predict

PM2.5 poorly in India's highly variable, extreme pollution environment.

Second, we document significant non-linearities in pollution's dose–response relationship. Understanding these non-linearities is particularly important because linear extrapolations from low-pollution settings may mischaracterize impacts in high-pollution environments typical of developing countries.Although non-linearities in the response to environmental conditions such as temperature have been well-documented (CITE), non-linearities in response to pollution exposure are just beginning to be explored, and a better understanding is needed of how these non-linearities manifest across contexts involving differing degrees of mental and physical performance.

Third, we generate novel evidence on how long-term pollution exposure mediates short-term responses—a relationship largely unexplored in the productivity literature. This gap stems partly from methodological limitations: most studies examine short-term pollution shocks for workers in fixed locations, typically using wind patterns or thermal inversions as instruments to address endogeneity between pollution exposure and location choices (Chung et al., 2025; Hansen-Lewis, 2024; He et al., 2019; Hill et al., 2024; Merfeld, 2023). Our approach differs fundamentally by exploiting exogenous variation in both dimensions of exposure. This is a major methodological advance relative to studies that must treat average exposure levels to environmental conditions (e.g., temperature) as unidentified, since they are correlated with a host of unobservable place-based characteristics. Here, we are able to calculate plausibly exogenous levels of average exposure for cricket players who are moving through locations with varying levels of pollution due to the scheduling rules of the league.

As a corallary, we are the first paper to our knowledge to identify the trade-off between the performance and health effects of accumulated exposure to air pollution. Although accumulated past exposure to PM2.5 has been linked to a number of detrimental health outcomes (CITE), we document with this paper a potential benefit of past exposure: it softens the blow of acute pollution episodes. This tension is particularly relevant for high-intensity activities, where there are substantial economic returns even to small gains in performance.

Top professional cricket players, who are paid in the millions of U.S. dollars (MoneyBall, 2025) fall into this category, but are not alone. Another example of a lucrative profession where minute changes in performance have outsized effects are stock-traders, who also have been shown to be impacted by PM2.5, thereby affecting same-day stock returns (Heyes et al., 2016).

Our study also builds on epidemiological methods for studying long-term pollution exposure (Schwartz, 2000; Wei et al., 2021; Zanobetti et al., 2002) by extending this framework from health outcomes such as mortality and cardiovascular disease to worker productivity, where elasticities may differ substantially.

Our findings have important implications for environmental policy in developing countries. The non-linear dose–response relationship—where workers in higher PM2.5 concentrations experience larger marginal effects than those in lower concentrations—suggests that pollution control efforts targeting extreme episodes may yield particularly large productivity benefits, even in chronically polluted environmental. However, our evidence for adaptation implies that workers in persistently polluted areas may develop some resilience to acute shocks. Moreover, the asymmetric impacts we document across different physical tasks suggest that air quality regulations may need to account for varying occupational exposures and adaptation possibilities. Understanding these patterns is crucial for both targeting environmental regulations and designing workplace health policies that account for task-specific vulnerabilities in polluted environments.

This paper proceeds as follows. Section 2 describes our cricket-based empirical setting for measuring labor productivity impacts. Section 3 presents our conceptual framework for non-linear and adaptation effects. Section 5 outlines our empirical strategy. Section 4 describes the data on cricket performance, air quality, and weather. Section 6 presents our findings.

2 Institutional detail

Cricket provides an advantageous setting for studying the impact of air quality on physical labor productivity because players perform specialized roles, resulting in differing susceptibility to PM2.5 exposure throughout the game. A cricket match thus generates rich data on observable outputs that are affected by air pollution. At its core, cricket is a bat-and-ball sport in which two teams alternate between batting and bowling, with the objective of scoring more runs than the opposition. However, cricket is distinctive among team sports in that it features players in highly specialized roles who engage in what is essentially a series of repeated individual contests within a team framework (Bartlett, 2003). Specifically, the fundamental unit of play is the interaction between a bowler (analogous to a pitcher in baseball) and batter (analagous to a baseball batter), with each delivery of a ball representing a discrete episode of measurable performance. Bowlers are responsible for delivering the ball toward the batter's wicket, employing various techniques to make run-scoring difficult, while batters attempt to score runs through offensive batting.[1]

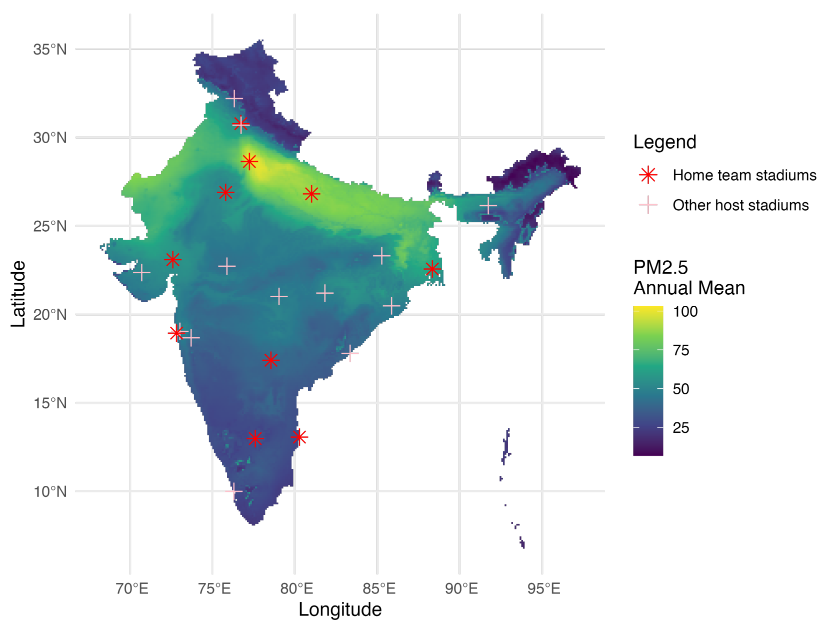
The Indian Premier League (IPL), founded in 2008, has emerged as the world's premier Twenty20 cricket competition and provides our empirical setting.[2] Cricket players in the IPL are high-salaried workers, with the top performing players regularly earning in the range of \$1 to \$2 million USD in the span of a two month season (MoneyBall, 2025). The league itself was valued at over \$10 billion USD in 2022 (Times, 2022). The productivity effects of air pollution on cricket players are thus economically important in their own right. In addition, the clean causal identification and rich performance data that the IPL yields make it an ideal setting to understand some of the nuanced effects of pollution on performance, which can then be explored in further studies on other populations of workers.

The IPL's structure and characteristics create two distinct sources of variation that make it particularly valuable for studying air quality impacts on worker productivity. First, the

[1]Fielders, the third type of player, support the bowling side by catching batted balls and preventing runs.
[2]Twenty20, indicating two innings with 20 overs each, is one style of cricket that is designed to be played over an approximately 3 hour period, in contrast to the traditional style in which games can take days.

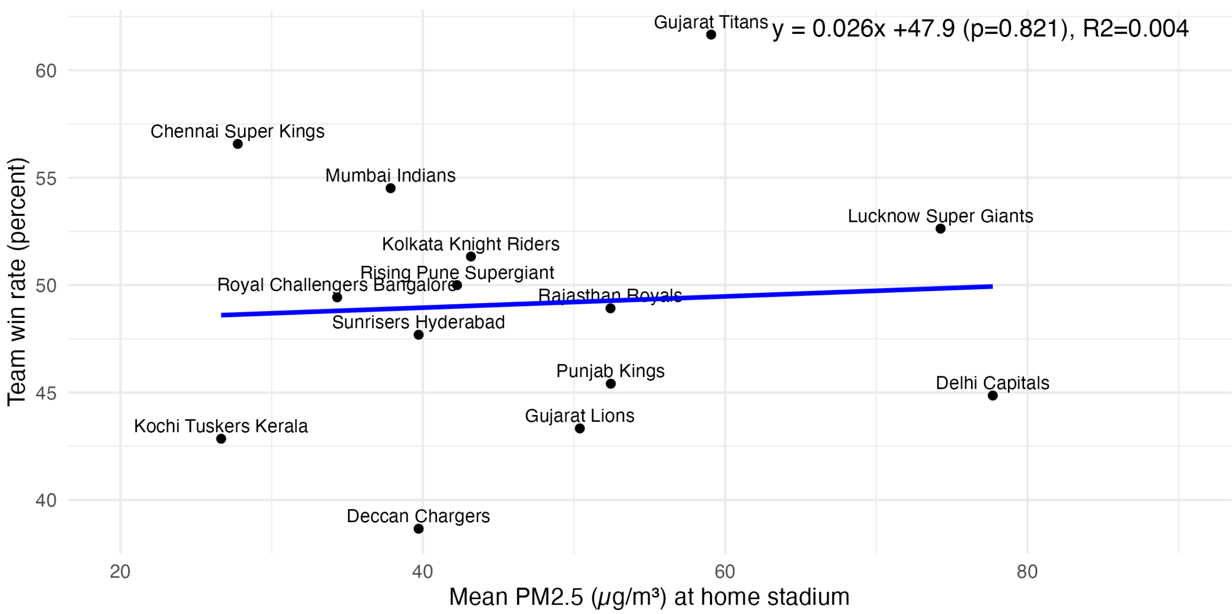
IPL operates as a franchise system with ten teams based in different cities: Chennai, Delhi, Ahmedabad, Kolkata, Lucknow, Mumbai, Mullanpur, Jaipur, Bengalaru, and Hyderabad (Figure 2). These cities span India's diverse geography, from coastal regions to inland valleys, and vary substantially in their air quality due to differences in vehicle traffic, industrial activity, and natural features. This geographic dispersion creates variation in pollution exposure across team training locations, as indicated in the wide dispersion of pollution in Figure 2. Crucially for our analysis, this variation in PM2.5 exposure is not correlated with team quality as indicated by the proportion of matches that teams win (Figure 3). This lack of relationship makes sense given the fact that each cricket team is provided with an equal amount of funding with which to buy players in each season's auction for players. This rule in the tournament prevents teams from purchasing systematically better players due to wealth.

Figure 2: Geographic distribution of cricket stadiums and PM2.5 in the IPL



Notes. This figure shows annual mean PM2.5 across India in 2019 as estimated by Wang et al. (2024) and the locations of cricket stadiums used in the Indian Premier League (IPL) from 2008–2022. Red asterisks indicate home stadiums of the ten IPL franchise teams: Chennai Super Kings (Chennai), Delhi Capitals (New Delhi), Gujarat Titans (Ahmedabad), Kolkata Knight Riders (Kolkata), Lucknow Super Giants (Lucknow), Mumbai Indians (Mumbai), Punjab Kings (Mullanpur), Rajasthan Royals (Jaipur), Royal Challengers Bengaluru (Bengaluru), and Sunrisers Hyderabad (Hyderabad). Pink crosses indicate other stadiums that have hosted IPL matches during this period but are not permanent home venues for any current IPL team.

Figure 3: Team quality is not correlated with long-term PM2.5 exposure



Notes. This figure shows the win rate (number of matches won out of all matches played) for IPL teams as plotted against the average PM2.5 at the team's home stadium. Average PM2.5 is defined as the mean level of PM2.5 at the team's home stadium during March, April, and May (the months when IPL matches typically occur) for the study period, 2008-2022. The lack of significant relationship between win rate and PM2.5 is robust to alternative definitions of long-term PM2.5, such as including non-IPL months, or including PM2.5 levels before the IPL franchise began. Note that the Pune Warriors, who played only 2011-2013 and had the lowest win-rate (26.7%, more than 2 standard deviations below the mean) are excluded from this graph. The lack of significant relationship also holds when including them.

Second, the distinct physical demands across playing positions in Twenty20 cricket create variation in how air quality might affect different types of players. Fast bowlers face the most intense physiological requirements (Noakes and Durandt, 2000). During delivery, a bowler accelerates through their run-up reaching speeds of approximately 22 kilometers per hour, plants their front foot with ground reaction forces between 2.4 and 5.8 times their body weight, and decelerates rapidly after ball release (Bartlett, 2003). In Twenty20 cricket, fast bowlers perform approximately 23 sprints per hour of play, with significantly less recovery time between high-intensity efforts compared to other positions (Petersen et al., 2010). During bowling spells, heart rates can reach 180–190 beats per minute, and in hot conditions, bowlers have recorded sweat rates up to 1.5 liters per hour, comparable to marathon runners (Noakes and Durandt, 2000).

For batters, the physical requirements combine intermittent high-intensity running with periods of technical performance (Noakes and Durandt, 2000), surrounded by periods of rest. In Twenty20 cricket, batters perform approximately 15 sprints per hour (Petersen et al., 2010). The activity pattern, however, is highly intermittent: mean heart rates during a day's cricket rarely exceed 128 beats per minute for batters. Over a complete Twenty20 innings, batters cover approximately 3.5 kilometers in total distance—a relatively small amount considering that innings stretch on for about 1.5 hours. Only about 20 percent of this distance covered in high-intensity running (Petersen et al., 2010). Unlike fast bowlers who run before delivering each ball, a batter's main strenuous task is not running, but the technical demands of facing deliveries from bowlers that can exceed 140 kilometers per hour (Bartlett, 2003). These factors indicate that batters' requirements are less physically demanding than bowler requirements, suggesting that bowlers may be more affected by air pollution than are batters.

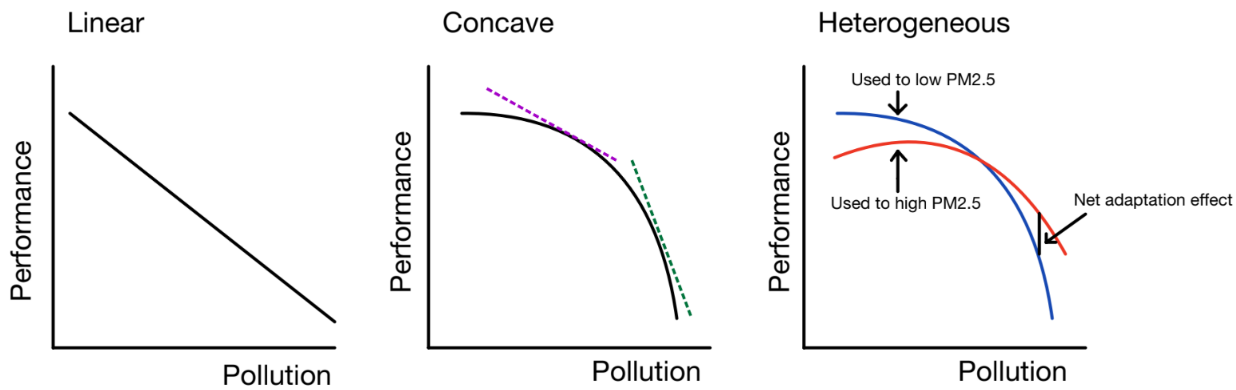
Together, these features of professional cricket in India create a uniquely suitable setting for studying air quality's effects on worker productivity across multiple temporal dimensions. Match schedules in the IPL are predetermined and beyond teams' control, creating plausibly exogenous variation in short-term exposure through day-of-match pollution levels. The IPL's geographic dispersion provides variation in medium-term exposure through differences in baseline air quality across team locations, while the international composition of teams introduces variation in longer-term exposure through players' diverse origins. These temporal variations in pollution exposure, combined with the distinctly different physical demands placed on bowlers versus batters, allow us to examine how both acute and chronic exposure to air pollution affects workers performing different types of physical tasks. Moreover, the discrete nature of cricket's bowler–batter interactions generates rich data on individual performance, enabling precise measurement of productivity effects.

3 Conceptual framework

While the negative effect of air pollution on performance is well documented (Aguilar-Gomez et al., 2022), the relationship between pollution and productivity is likely to be more nuanced than a simple linear decline. Our conceptual framework highlights three channels through which air pollution may affect worker performance. First, the negative impacts of air pollution on productivity may be non-linear, with potentially larger marginal effects at higher concentrations. Second, workers may adapt to chronic exposure through both behavioral responses (e.g., using protective equipment such as masks, or respiratory therapies such as inhalers or supplemental oxygen) and physiological changes, as suggested by medical research showing respiratory system adaptation to repeated pollution exposure (Dimeo et al., 1981; Hackney et al., 1977; Hamade and Tankersley, 2009). Third, pollution's impacts likely vary by task type, depending on the interplay of specific physical and cognitive requirements.

First, the marginal effect of pollution on performance may vary with the level of pollution itself. As depicted in the left panel of Figure 4, a linear damage function implies constant marginal effects across all pollution levels. However, the relationship could be concave (middle panel), where the drop in productivity from an increase in pollution is larger at

Figure 4: Dose response of performance to pollution



higher pollution levels. Alternatively, the relationship could be convex, though the key insight remains: there is little theoretical justification for assuming constant marginal effects across all pollution levels.

Second, the effect of pollution may vary systematically across individuals based on their typical exposure levels. The right panel of Figure 4 illustrates how performance might differ between individuals accustomed to high versus low pollution levels.[3] The figure suggests a potential crossing of damage functions: individuals adapted to high pollution might perform better than their low-pollution counterparts when pollution is high, but worse when pollution is low. Third, the impact of pollution on performance may vary by the type of task performed, with tasks that require certain types of effort or prolonged exposure being more vulnerable to pollution's effects.

There may be negative effects to performance resulting from cumulative physiological damage from long-term pollution exposure, though this need not be permanent as recovery may occur during periods of lower exposure. There also may be positive effects from acclimatization along several dimensions: behavioral (e.g., learning when and how to avoid pollution by wearing masks, using inhalers), physiological (e.g., changes in lung function) (Hackney et al., 1977; Hamade and Tankersley, 2009), and psychological (e.g., developing coping strategies for pollution-related discomfort).

[3]Similarly, we could categorize damage functions into different tasks, with different damage functions for each type of task.

We define a production function describing a cricket bowler's performance as $F(P, a)$ where P is pollution and a are adaptation measures that can occur in the long-run (either positive or negative).[4]

Equation 1 decomposes the total effect of pollution on performance into its constituent parts:

$$\underbrace{\frac{\mathrm{d}F}{\mathrm{d}p}}_{\text{Long-run effect}} = \underbrace{\frac{\partial F}{\partial p}}_{\text{Short-run effect}} + \underbrace{\frac{\partial F}{\partial a}\frac{\mathrm{d}a}{\mathrm{d}p}}_{\text{Adaptation effect}} \tag{1}$$

Standard reduced-form estimates capture only the short-run effect $\dfrac{\partial F}{\partial p}$. By interacting current pollution with a longer-term average exposure in our empirical specifications, we can begin to disentangle the short- and long-run effects. In the next section, we outline the physiological basis for why the short-run effect may be mediated by the long-run effect.

3.1 Physiological basis for adaptation

While the notion of adaptation is not new to the economics literature (Dell et al., 2014; Burke and Emerick, 2016; Moore, 2017; Burke et al., 2024) it is typically conceived as a behavioral response to long-run changes, consistent with an optimization framework. The adaptation under study here is similar in that it involves a response to a changing environment; however, the optimizing agent we focus on is not an individual person, but rather biochemical processes in cells.

Using randomized controlled trials exposing mice to varying levels of air pollution over time, the environmental toxicology literature has established that mice become resistant to the acute damages of pollution with repeated exposure (West et al., 2003; Kültz et al., 2015; Lee et al., 2018). Cells in the respiratory tract and lungs—i.e., the cells hit hardest by air pollution—can mitigate the damage from air pollution using an antioxidant called glutathione that resides in cells within the lining of the walls of the lungs and respiratory

[4] This builds on the conceptual framework in Hagerty (2022) which is applied to long-term changes in agricultural practices.

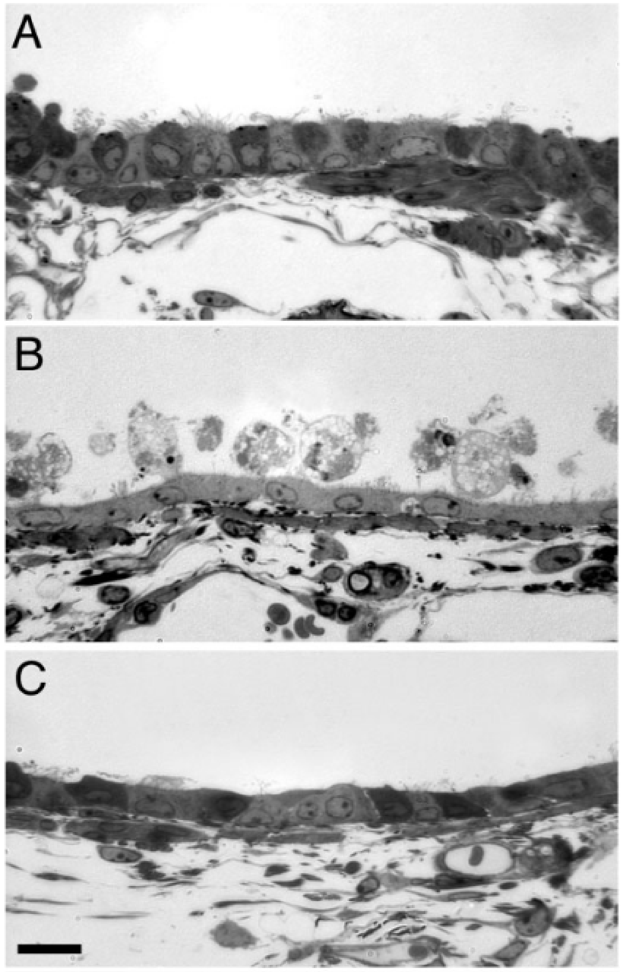
tract (i.e., the epithilial lining). Glutathione is also present in humans and is understood to perform a similar function in humans as in mice (Ketterer et al., 1983), an insight that serves as the basis for an extensive literature investigating the effects of air pollution on mice to learn about implications for humans.

When air pollution enters the body through respiration, it brings with it a variety of harmful chemicals that the body must process in order to expel. The first step of this process, known as Phase I metabolism, involves converting the molecule into a more polar version of itself—often a reactive electrophile—which can result in it becoming more toxic. This is a preliminary step which is necessary for Phase II, in which the body makes the molecule more water soluble, and thus easier to excrete. Glutathione in the lungs (in both its substrate and enzyme form), performs this second step by attaching itself to the molecules produced from Phase I (a process known as conjugation), neutralizing the reactive electrophiles, thereby completing Phase II metabolism.

At low levels of pollution, there is enough glutathione present at baseline to neutralize the harmful effects of air pollution. As air pollution increases, however, the body becomes unable to produce enough glutathione at a sufficient rate to have enough to conjugate incoming electrophiles from air pollution. The rate limiting step in the chemical reaction for the body to produce glutathione is an enzyme called γ-glutamylcysteine synthetase (γ-GCS), which itself must be produced by cells in the lungs. The key adaptation that the environmental toxicology literature has identified is that, when exposed to air pollution repeatedly, cells augment their ability to produce γ-GCS, enabling them to produce glutathione more quickly in response to incoming air pollution.

While the body does not become completely immune to air pollution at all levels, augmenting the ability to produce γ-GCS enables it to endure higher levels of air pollution than it would otherwise while suffering minimal damage from an acute episode of exposure. To visualize the effects of this adaptation response, Figure 5 shows magnified images of cells from the lungs of mice after a single day of air pollution exposure (panel B) and 7 days of

Figure 5: Cellular adaptation to repeated exposure to air pollution



Notes. 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. As for the cells in Panel B, their cells are imaged 24 hours after their last exposure.

exposure (panel C), relative to a control group of mice who were exposed to clean air during the same period (panel A). The key insight is that the mice who were exposed over 7 days have an intact cellular structure that is more similar to the control cells than to the ones that experienced only a single episode of exposure.

4 Data

We combine three data sources to study the effect of PM2.5 on cricket player performance: detailed ball-by-ball performance data from 773 IPL matches in India[5] spanning 2008–2022; high-resolution daily air pollution measures from both ground monitors and satellite-based estimates; and comprehensive weather controls. Our ball-by-ball data capture granular measures of player performance through specific bowler-batter interactions, while our pollution data allow us to precisely characterize both acute and chronic exposure patterns across India's diverse geography. The combination of these three datasets yields a rich analytical sample that allows us to examine the relationship between air quality and athletic performance while controlling for potentially confounding weather conditions. This section describes each data source and discusses features relevant for our empirical analysis.

4.1 Cricket performance

We obtain ball-by-ball cricket performance data for the IPL from Cricsheet, an open-source repository of cricket match data. Our sample covers IPL matches from the league's inaugural 2008 season through the 2022 season, comprising over 180,000 deliveries (analogous to pitches in baseball) across 773 matches matches.[6] For each delivery, we observe detailed information including the identities of both the bowler and batter, the outcome of the delivery (such as the number of runs scored or wickets taken), and the precise match situation (such as the over number and the score).

We leverage the data's ball-by-ball granularity by defining our main outcome as a binary indicator for whether any runs were conceded on a given delivery. This approach has three advantages over aggregate performance measures (such as total runs scored). First,

[5]Several seasons (and parts of seasons) of the IPL were played outside India due to security and public health concerns. We exclude these games from our analysis to focus on productivity impacts within the high-pollution setting of India.

[6]Cricsheet data are available at www.cricsheet.org. Several seasons (and parts of seasons) of the IPL were played outside India due to security and public health concerns. We exclude these games from our analysis to focus on productivity impacts within the high-pollution setting of India.

it preserves the rich variation in the data, allowing us to control for both time-invariant characteristics of players and evolving match conditions. Second, it provides a player-level measure of performance based on specific bowler–batter interactions, rather than team-level outcomes that depend on many additional factors.

Third, by focusing on the binary outcome of conceding any runs rather than the number of runs conceded, we avoid potential complications from cricket's non-linear scoring system where certain outcomes (e.g., boundaries worth four or six runs) occur discretely based on specific field events rather than incremental performance differences. As a robustness test, we conduct analyses accounting for the full range of run outcomes (0, 1, 2, 3, 4, 5, or 6) using an ordered logit model.

Table 1: Summary Statistics from IPL Matches (2008–2022)

Statistic	**Mean**	**Median**	**Note**
Matches in India	773	–	–
Runs scored per match	298	302	Range: 51 to 448
Deliveries per match	237.5	244	Range: 51 to 263
Deliveries per match resulting in ≥ 1 run	142	146	60% of deliveries
Bowlers in sample	445	–	–
Strikers in sample	575	–	–
Players in sample	619	–	Bowlers may play as strikers
Bowlers per team per game	5.9	6	Range: 2 to 9
Deliveries per bowler per match	20.1	24	Range: 1 to 34

Our analytic sample contains 773 games in the IPL from 2008-2022 that took place in India.[7] Table 1 describes summary statistics from these games. Bowlers bowl a maximum of four overs per game, where an over consists of six legal deliveries,[8] resulting in a theoretical maximum of 24 legal deliveries per bowler in a game. Bowlers typically meet this threshold: the median number of deliveries per game for a bowler is 24. A match has two innings, where each inning consists of up to 20 overs. This means that a match may have up to $2 \times 20 \times 6 = 240$ legal deliveries which is approximately what we observe: the mean number

[7] Some games took place outside India due to security concerns; we exclude these from our sample.

[8] A delivery can be classified as illegal several reasons, but they boil down to giving the bowler an unfair advantage against the batter (e.g., bowling the ball too far away from the batter, known as a "wide.")

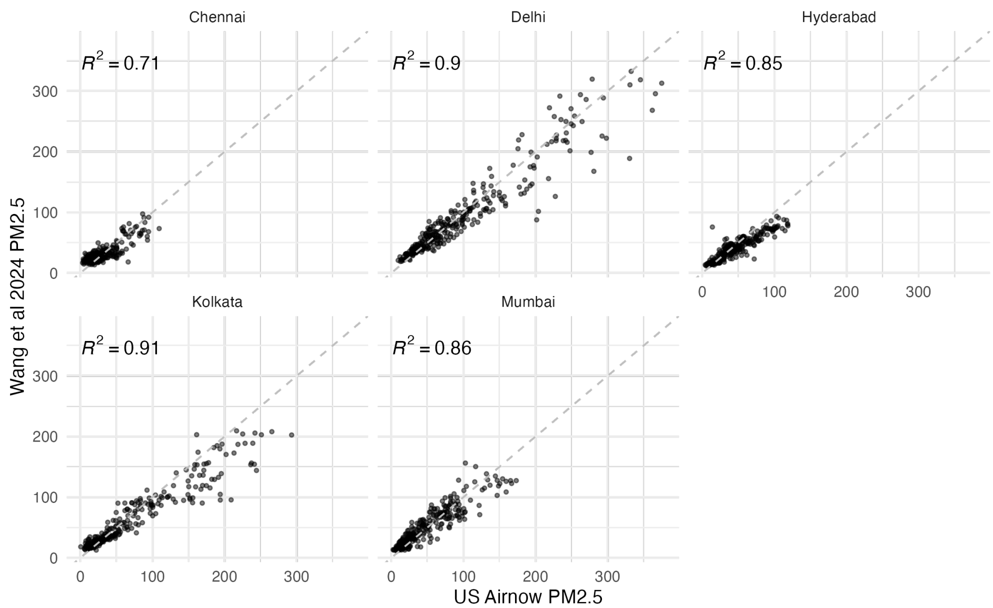
of deliveries per match is 237.5 (median 244). That the number of deliveries per bowler exceeds 24 in some cases and that the number of delivieries per match exceeds 240 in some cases is a result of a handful of illegal deliveries in each match.

4.2 Air quality

Our primary source of air pollution data comes from Wang et al. (2024), who provide daily estimates of ground-level fine particulate matter (PM2.5) concentrations at a 10km × 10km resolution across India for the period 1980–2022. These estimates are derived from a machine learning model that combines satellite data, meteorological information, and ground monitor readings. The high spatial resolution of this dataset allows us to precisely match daily air quality conditions to each of the 24 stadiums where IPL matches were played during our study period, including both the ten permanent home venues and fourteen additional stadiums that periodically hosted matches (Figure 2). We validate the Wang et al. (2024) data against the U.S. AirNow network in Figure 6, which demonstrates a strong correlation between the Wang et al. (2024) dataset and ground readings. By contrast, Figure A.1 reports the results of an analogous exercise with MODIS AOD, and shows a much weaker correlation. We therefore choose Wang et al. (2024) as our source of data for PM2.5 concentrations.

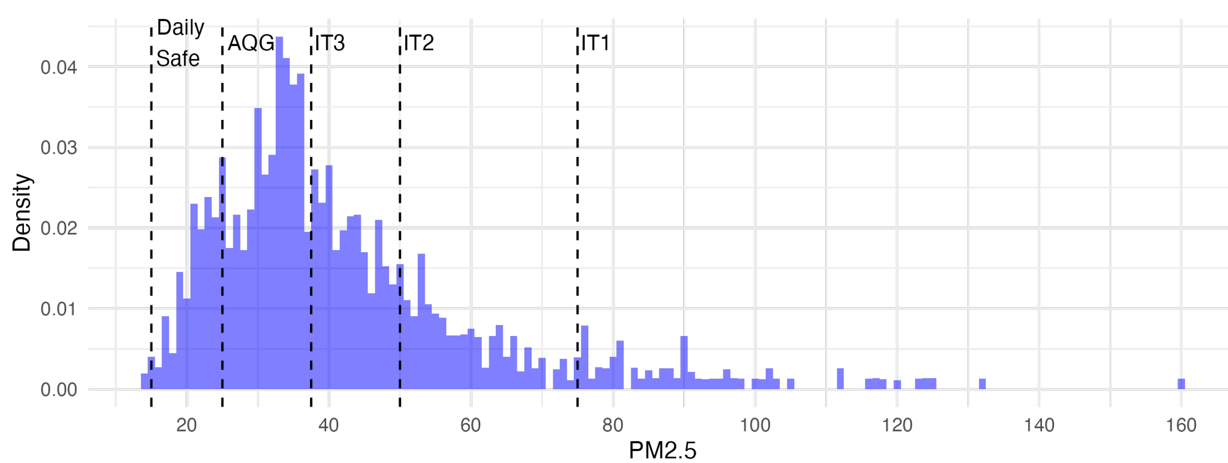
Leveraging the Wang et al. (2024) data to gain complete coverage of all IPL games during our study period, we emphasize that the distribution of PM2.5 in IPL games is dramatically higher than PM2.5 in high-income countries, with a mean PM2.5 concentration of 42 $\mu g\,m^{-3}$ on game day (median 37 $\mu g\,m^{-3}$), and a maximum of 160 $\mu g\,m^{-3}$—over 10 times the WHO safe daily limit (Figure 7, World Health Organization (2021)). These PM2.5 concentrations are orders of magnitude higher than annual average PM2.5 concentrations in, for exmaple, the U.S., which ranged from 7-9 $\mu g\,m^{-3}$ in 2013-2023.

Figure 6: U.S. Airnow vs. Wang et al. (2024) Daily PM2.5



Notes. This figure displays the correlation between Wang et al. (2024) estimates of PM2.5 and ground observations of PM2.5 at five cities with a U.S. AirNow ground monitor, a network of high-quality monitors that are calibrated to EPA standards. Each dot represents a daily mean value of PM2.5. The 45 degree line is shown for reference. Importantly, the U.S. AirNow network is not included as training data for the Wang et al. (2024) model.

Figure 7: IPL Game PM2.5 Distribution with WHO Thresholds Indicated



Notes. This figure shows the distribution of PM2.5 at IPL games in the period 2008-2022 with WHO thresholds for daily PM2.5 exposure indicated. WHO Interim Threshold 1 (IT1) is 75 $\mu g\, m^{-3}$, WHO Interim Threshold 2 (IT2) is 50 $\mu g\, m^{-3}$, WHO Interim Threshold 3 (IT3) is 37.5 $\mu g\, m^{-3}$, WHO Air Quality Guideline (AQG) is 25 $\mu g\, m^{-3}$, and The WHO Daily Safe exposure threshold (Daily Safe) is 15 $\mu g\, m^{-3}$. These thresholds are based on the levels of PM2.5 associated with higher short-term mortality risk (World Health Organization, 2021).

4.3 Weather

Particulate matter is not the only meteorological factor that may affect performance in cricket; the weather may do so as well. PM2.5 concentrations themselves are a product of many meteorological factors, including temperature, humidity,[9] and wind, and thus have substantial correlations (sometimes negative, sometimes positive) with each of these. We therefore include temperature, temperature-squared (to capture the non-linear effect of temperature on performance), humidity, precipitation, solar radiation, and wind speed in our regression specifications. We obtain weather variables from ERA5-Land, a state-of-the-art reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts, accessed through Google Earth Engine (Muñoz Sabater, 2019). The ERA5-Land data provide global coverage at approximately 11km $\times$ 11km spatial resolution, enabling precise matching to match locations. These weather controls are crucial for our analysis as they may independently affect both cricket performance and pollution levels.

5 Empirical approach

We examine how air pollution affects cricket performance, with particular attention to how players' typical exposure levels mediate these effects. Our main analysis centers on ball-by-ball outcomes, using a binary indicator for whether a run is scored on each delivery as our primary outcome variable. This choice of dependent variable is motivated by three considerations. First, conditional on scoring, the distribution of runs is heavily concentrated at one run, with over 60 percent of scoring deliveries resulting in a single run (Figure A.2). Second, certain run values (particularly fours and sixes) are achieved through qualitatively different actions—hitting the ball to the boundary of the stadium rather than running between wickets (analogous to bases in baseball)—making a continuous measure potentially misleading. Third,

[9] ERA5-Land contains dewpoint temperature rather than relative humidity. We calculate relative humidity from temperature and dewpoint temperature using the meteorological approximation outlined by Lawrence (2005).

the binary outcome provides a clear interpretation: we can characterize the marginal effect of pollution as the change in probability of a bowler conceding (or batter scoring) a run, where a bowler conceding a run indicates poor performance while the opposite is the case for a batter.[10]

5.1 Main specification

We examine the relationship between match-day PM2.5 levels and run-scoring probability with our baseline specification:

$$R_{ij\ell t} = \beta \text{PM2.5}_{\ell d} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \Lambda_i + \Delta_j + \varepsilon_{ij\ell t} \tag{2}$$

where $R_{ij\ell t}$ is an indicator for whether a run is scored on a delivery from bowler j to batter i from ball-of-match t at location (stadium) ℓ. Our coefficient of interest is β, which captures the effect of match-day PM2.5 levels (measured in 10 $\mu\text{g}\,\text{m}^{-3}$) on run-scoring probability.[11] Our preferred specification includes bowler fixed effects (ψ_j), batter fixed effects (ϕ_i), stadium-by-year fixed effects ($\delta_{\ell y}$), innings fixed effects (θ_n) (there are two innings per game), over fixed effects (η_o) (there are 20 overs per inning), dummy variables for whether the stadium is the home stadium for the batter (Λ_i) or bowler (Δ_j),[12] and a vector of weather controls for the day d and location ℓ of the match ($\mathbf{X}_{\ell d}$) including a linear and quadratic term for temperature, and linear terms for relative humidity, atmospheric pressure, precipitation, solar radiation, and wind speed.

Identification in Equation 2 comes from the fact that cricket matches are scheduled well in advance (before reliable air pollution forecasts are available) and cannot be rescheduled

[10] There is a slight wrinkle for bowlers in that they also seek to get the batter out, which may justify allowing a run in rare cases. We consider this a minor consideration and abstract away from it in our analysis.

[11] The standard measure of PM2.5 is 1 $\mu\text{g}\,\text{m}^{-3}$. When measured in these units, the effects we observe are very small in magnitude. Given the distribution of PM2.5 in our sample (see Figure 7), measuring PM2.5 in 10 $\mu\text{g}\,\text{m}^{-3}$ is appropriate. The effect sizes we observe are similar to those of other studies that measure effects in 10 $\mu\text{g}\,\text{m}^{-3}$ (Adhvaryu et al., 2022).

[12] Note that some matches are played at stadiums that are a home stadium for neither time, which prevents these two terms from being collinear.

due to pollution levels, making match-day PM2.5 exposure plausibly exogenous to player characteristics and performance potential. The validity of our estimates of the causal effect of air pollution on performance rests on the assumption that, conditional on our included fixed effects and controls, match-day pollution levels are uncorrelated with unobserved determinants of cricket performance.

All specifications cluster standard errors two-way at the match and bowler level to account for potential serial correlation in performance for bowlers and the fact that pollution exposure is observed for each match. Additionally, bowler and batter fixed effects control for any time-invariant player-specific factors, stadium-by-year fixed effects account for venue-specific temporal trends that might correlate with both pollution levels and cricket performance, and innings and over fixed effects control for stage-of-game effects that could influence playing styles and strategies. The dummy variables for the home stadium of the bowler and batter account for home field advantage.

5.2 Non-linear effects

To examine potential non-linearities in the pollution-performance relationship, we also estimate the following specification replacing the continuous PM2.5 measure with indicators for PM2.5 quantiles:

$$R_{ij\ell t} = \sum_{k=2}^{5} \beta_k Q_k \left(\text{PM2.5}_{\ell d}\right) + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \Lambda_i + \Delta_j + \varepsilon_{ij\ell t} \tag{3}$$

where $Q_k(\text{PM2.5}_{\ell d})$ represents indicators for the second through fifth quintiles of match-day PM2.5, with the lowest quintile serving as the reference category.

We explore spline regression with varying numbers and placements of knots to identify non-linearities in the response of performance to pollution. The spline allows us to construct the dose-response function of performance to pollution as a series of piece-wise functions across varying levels of pollution, while allowing the slope (and in polynomials of degree

$p > 1$, the curvature) of the function to vary. This estimation procedure allows us to directly answer whether the dose-response function exhibits non-linearities. To do this, we estimate

$$R_{ij\ell t} = \sum_{j=0}^{p} \beta_j (\text{PM2.5}_{\ell d}^{j}) + \sum_{k=1}^{N} \beta_{p+k} (\text{PM2.5}_{\ell d} - \tau_k)^p \mathbf{1}\{\text{PM2.5}_{\ell d} \geq \tau_k\} + \mathbf{X}_{\ell d}' \phi + \psi_j + \varepsilon_{ij\ell t} \quad (4)$$

where we vary the number of knots N and the degree of the polynomial p. In the spline specification, we omit fixed effects other than bowler fixed effects to ease computation.

5.3 Defining past exposure to PM2.5

Our approach to exploring heterogeneity in the short-run effect of pollution by exposure to pollution in the long-run—suggesting a form of adaptation—connects to the climate economics literature which examines how long-term temperature averages in a specific region mediate short-term response to temperature changes (Dell et al., 2014; Mérel and Gammans, 2021; Mérel et al., 2024). However, unlike the climate literature which focuses almost exclusively on adaptation within spatial units (e.g. agricultural fields, Census tracts, or nations), we shift the unit of analysis to the person.

We build on the person-based approach to measuring pollution exposure histories in environmental toxicology (Dodson et al., 2007; Garcia et al., 2013). This approach, to our knowledge, is new to the economics literature and offers several advantages, along with a few challenges. The advantage of measuring person-specific adaptation is that we can track exposure to pollution for the same individual over time. This is a more precise way to measure pollution exposure than assuming, for example, that all individuals with an address in a district live and work in that district and therefore are exposed continuously to the pollution levels there. People move fairly regularly, especially lower-income migratory workers, who may work in one place during one season of the year and reside in another during other parts of the year. This movement may also occur within a day: for example, sleeping in a high-pollution exposed area but working in a low-exposure area, or vice versa.

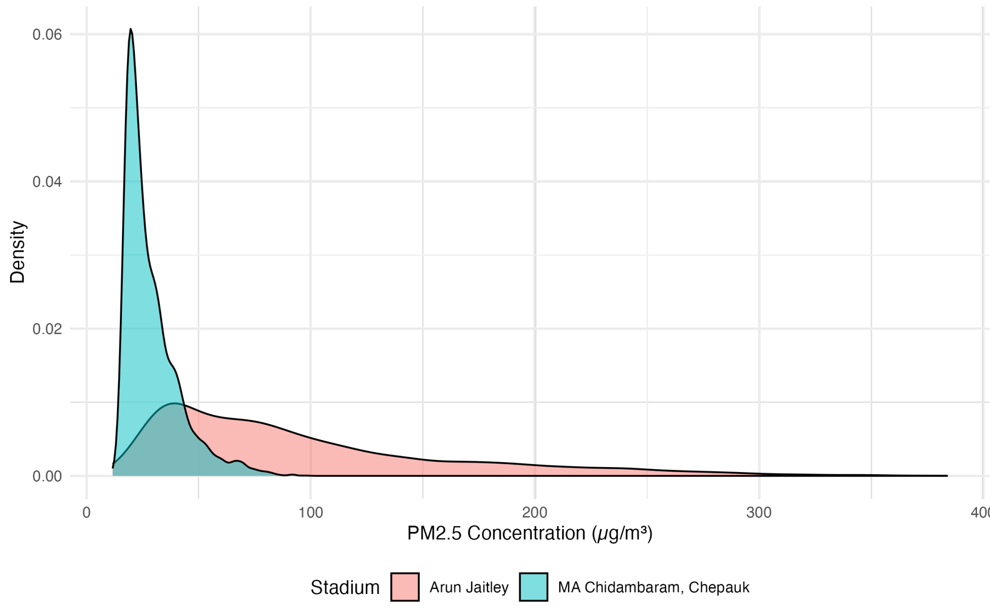
The challenges of the person-based approach is developing a conceptually sound measure of long-term exposure to pollution. Typically, the long-term exposure level for a unit is considered to be fixed. However, in the IPL context, players may switch teams across seasons. Among players who played in more than one season in our sample, the average number of times each player switched teams is 1.4, with a range of 2 to 12 times. The traditional space-based approach to long-term exposure measures would suggest assigning a long-term PM2.5 level to a player based on the location of their home-stadium, but this is problematic when their home stadium changes across years. We therefore explore alternative definitions of long-term exposure, discuss the advantages and drawbacks of each, and make a recommendation for how to measure this consistently in similar settings.

Following the climate econometrics literature (Dell et al., 2014), we conceptualize each team's home stadium as having its own pollution "climate," where daily pollution levels are realizations from this distribution. When teams play a game at a stadium in a rival's city, they experience a new pollution climate where daily pollution realizations are not different realizations from the same distribution but realizations from a different distribution. This framework is consistent with the observed data indicating diverging average pollution levels across stadiums (Figure 2). For example, in Figure 8 we see that although there is a common support of the probability mass at low pollution levels, the Arun Jaitley stadium Delhi regularly experiences pollution levels that the stadium in Chepauk, Chennai never experiences.

5.3.1 Medium-term exposure

We ground our definition of long-term pollution exposure in the environmental toxicology literature which finds that mice exhibit tolerance to air pollution when exposed to it repeatedly for seven days (West et al., 2003). Although the study found tolerance effects after seven days,

Figure 8: Probability Density Functions of PM2.5 in Two Example Stadiums



Notes. This figure displays the probability density function (PDF) of PM2.5 concentration in two example stadiums with divergent distributions of PM2.5.

this does not mean that this is the only time window in which these effects may occur[13]—we therefore flexibly adjust our time window from 1 to 90 days prior to the match.

We focus on three measure of past exposure to PM2.5:

1. Mean PM2.5 over past X days

$$PM2.5_{J(j)d} = \sum_{d=1}^{X} \overline{PM2.5_{J(j)d}} \tag{5}$$

2. Number of days in past X days where PM2.5 was above Z threshold

   E.g. WHO thresholds i.e. $Z \in \{25, 37.5, 50, 75\}$

$$PM2.5_{J(j)d} = \sum_{d=1}^{X} \mathbf{1}(PM2.5_{J(j)d} > Z) \tag{6}$$

[13] In personal communication with Laura Van Winkle, one of the authors of West et al. (2003), she suggested that tolerance effects may be even more pronounced after a longer period of exposure, but that the logistics of running a study with mice beyond seven days is complicated which forced them to limit the time frame of their study.

3. Degree day analogous measure (number of units above Z threshold for each day above Z threshold in past X days)

$$PM2.5_{J(j)d} = \sum_{d=1}^{X} [\mathbf{1}(PM2.5_{J(j)d} > Z)](PM2.5_{dj} - Z) \tag{7}$$

Since each of these measures extends into the past by at most three months and there is a nine-month gap between seasons of the IPL, these past-exposure measures are assigned at the level of the team, not that of the individual player. The team-level of assignment is reflected in the subscript $J(j)$, where J is a function that maps bowler j to team J.

To construct past exposure levels for each of these measures, we adopt a set of assumptions for assigning a team's exposure to pollution within a season for days on which a team does not play a match.

Rules for assigning PM2.5 exposure to teams

1. If the team plays a match on a given day, we assign the PM2.5 for the 10x10km grid containing the stadium for the match to the team.

2. If the team does not play a match on a given day, we implement the following set rules:

   (a) If the team plays a match on the next day, we assume the team travels to the new match location on that day, so we assign it the PM2.5 for the location of the next match's stadium.

   (b) If the team does not play a match on the next day, but did play a match on the previous day, we assume the team is still in the location of the previous match, so we assign them the PM2.5 for the location of the previous match's stadium.

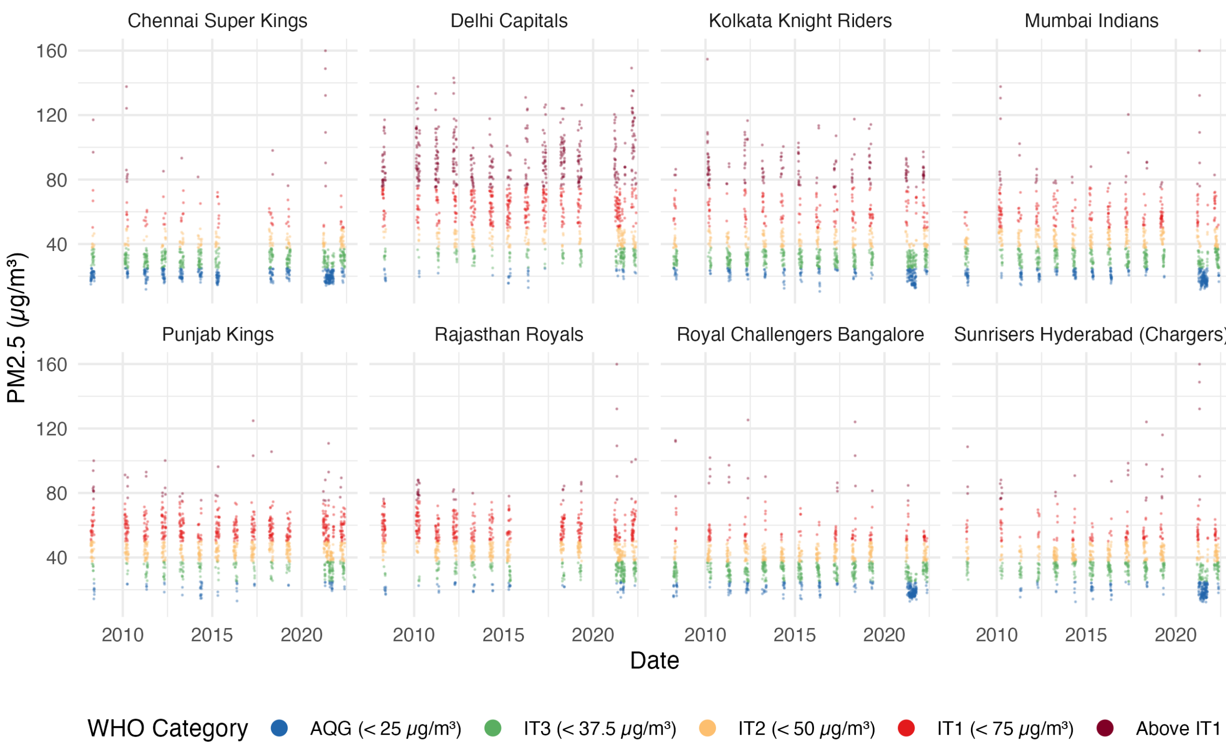
   (c) If neither of these are the case (i.e., the team did not play a match the day before and will not play one the day after), we assign them the PM2.5 for their home stadium.

Figure 9 shows the distribution of PM2.5 exposure histories for each of the eight teams with the highest numbers of matches in the IPL as of 2022.[14] In Figure 10, we compare the PM2.5 estimates assigned to teams (based on their travel itinerary, as described above) relative to the PM2.5 at their home stadium. The red line shows the level of PM2.5 exposure at the team's home stadium where the team spends a disproportionate amount, though not all, of their time. The blue line shows the PM2.5 exposure for each team taking into account their away games and travel schedule. When the two lines overlap, as they do most of the time, since teams spend most of their time in their home stadium, the line is purple. However, there are some teams that have extended periods of exposure that differ from the exposure levels at their home stadium. This tends to be the case for a team based in a high pollution location that travels to matches in lower pollution settings (e.g., Delhi Capitals), or a team with low pollution at its home stadium that then travels to higher pollution areas (e.g., Chennai Super Kings). That the red and blue lines are not consistently overlapping underscores the importance of carefully accounting for the actual levels of exposure each individual team faces.

We note that the measures in Equations 5, 6, 7 are piece-wise linear functions of one another and we therefore expect them to yield somewhat similar results. However, each measure presumes a conceptually distinct hypothesis as to how past exposure to PM2.5 may affect the body. Use of the measure in Equation 5 would suggest that the mean level of exposure in the past, say, 30 days is meaningful. Measuring the mean, however, may gloss over substantial variation in intensity of exposure over time. For example, in the hypothetical extreme case of someone who is exposed alternately to a day with perfectly clean air (i.e., 0 $\mu g\, m^{-3}$) and then a day with some of the most polluted air, say, 80 $\mu g\, m^{-3}$, the measure in Equation 5 would be 40 $\mu g\, m^{-3}$. This would be the same as that measure for someone who is exposed to days of exactly 40 $\mu g\, m^{-3}$ each day, even though the experience of being exposed to extreme pollution and then clean air alternately may be quite different than being exposed

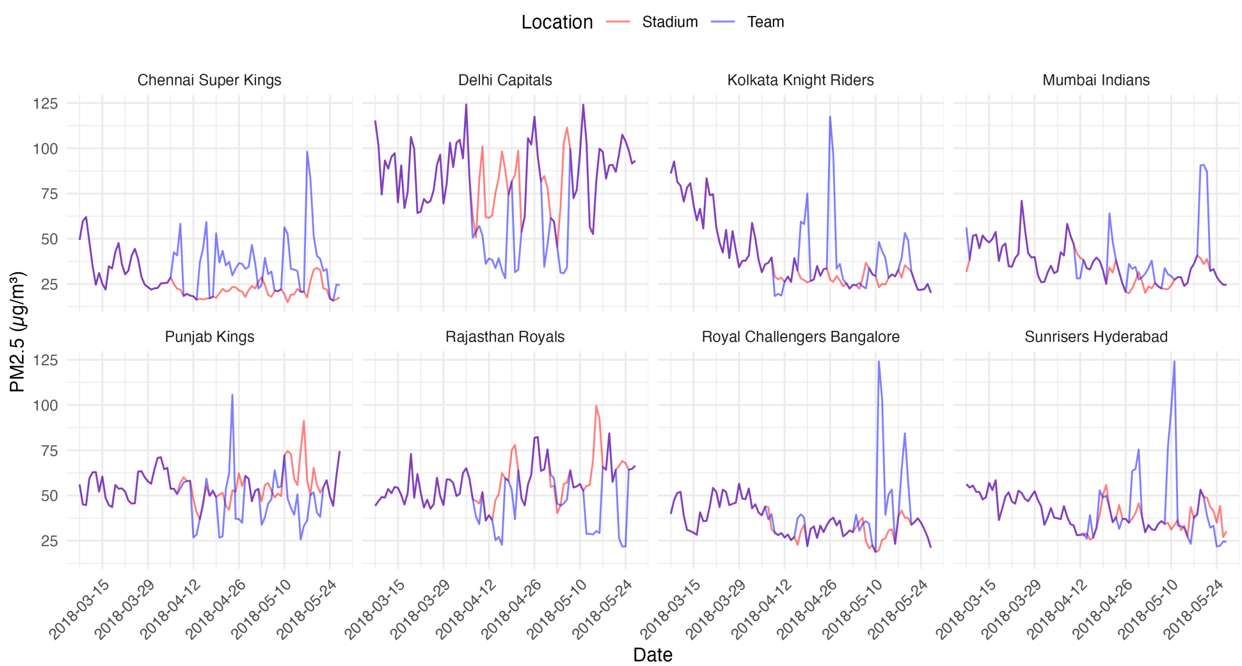
[14]We combine the PM2.5 exposure histories for teams that changed names across seasons but retained the same home city and stadium; the graph shows the most recent name of each team.

Figure 9: Team PM2.5 Exposure Histories in IPL, 2008-2022



Notes. This figure displays the PM2.5 exposure history of each IPL team in the years 2008-2022. Teams are assumed to be exposed to the level of PM2.5 in their home stadium, unless they are playing a match at an away stadium, in which case they are exposed to the PM2.5 at that location. The IPL takes place in three months each year and we assume teams are present and training in their home stadiums 30 days before the start date of each season. We do not interpolate PM2.5 exposure between seasons, leading to blank spaces between the clusters of dots for each season.

Figure 10: Team PM2.5 Exposure vs. Home Stadium PM2.5 Exposure in 2018



Notes. This figure compares the PM2.5 estimates assigned to teams (based on their travel itinerary, as in Section 5.3.1) relative to the PM2.5 at their home stadium. The red line shows the level of PM2.5 exposure at the team's home stadium whereas the blue line shows the PM2.5 exposure for each team taking into account their away games and travel schedule.

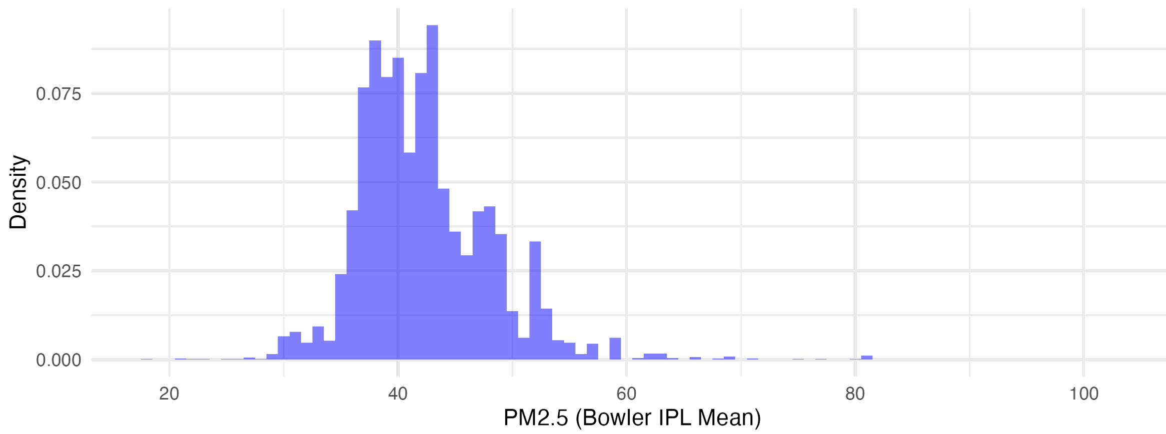
to moderate pollution consistently.

The measure in Equation 6 corrects for this coarseness to some extent by counting how many "bad" days of pollution there were for a team in a given time window, where "bad" is a threshold that can be adjusted. The underlying idea that there is a certain threshold beyond which exposure to PM2.5 is harmful and below which it is not. We use the thresholds that WHO sets (World Health Organization, 2021) for the Air Quality Guideline (25 $\mu g\, m^{-3}$, Interim Target 3 (37.5 $\mu g\, m^{-3}$), Interim Target 2 (50 $\mu g\, m^{-3}$) and Interim Target 1 (75 $\mu g\, m^{-3}$), and explore robustness of the measure to each threshold. Returning to the example above, using the threshold of 50 $\mu g\, m^{-3}$ would mean that the measure in Equation 6 would be 15 in the alternating case and 0 in the consistent case.

The measure in Equation 7 is a further refinement of that in Equation 6—it quantifies not only whether a day has "bad" pollution, but also "how bad" it is; in other words, by how much it exceeds a given threshold. This method is analogous to a cooling degree day in the climate literature. The underlying idea that it matters not only whether PM2.5 exceeds a threshold, but by how much it exceeds it. Returning to the first example and using a threshold of 50 $\mu g\, m^{-3}$, the team exposed to alternate 0 and 80 $\mu g\, m^{-3}$ days for 30 days would be assigned $30 \times 15 = 450$ for this measure, while the team exposed to 40 $\mu g\, m^{-3}$ consistently would be assigned 0. The advantage of Equation 7 relative to Equation 6 is that Equation 7 makes it possible to distinguish between a scenario where on the "high" days in the alternating scenario the high is, say, 55 $\mu g\, m^{-3}$ versus 80 $\mu g\, m^{-3}$. The former measure would be the same in both cases (since the "high" day is above 50 in both cases) while the latter would reflect this variation.

In addition to these three measures (and the varying thresholds of Z), we also vary the time window in which we look at past exposure from 1 day prior to the match to 90 days prior to it. For this pollution exposure history for games early in each season, we extend back only 30 days prior to the season start. This decision reflects the fact that teams tend to only start to train in their home stadiums several weeks before the season starts.

Figure 11: Histogram of Bowler PM2.5 Mean in IPL Games 2008-2022



Notes. This figure displays the distribution of bowler mean PM2.5 defined as the mean PM2.5 exposure across all games for a given bowler in the IPL 2008-2022 (i.e. PM2.5$_{j0}$ in Equation 11).

5.3.2 Long-term exposure

In addition to the measures of past exposure in Section 5.3.1 which vary for the same team (and individual) over time, we also test the robustness of our results for a measure that is fixed for each individual: the average of their exposure across all the games they play in the IPL. This measure is displayed in Figure 11. It consists not only of a player's home-stadium's average, but also includes exposures from matches played in away-stadiums, though is disproportionately weighted to the pollution levels at their home stadium (since teams play roughly half their games at home). It only tracks their pollution exposure during IPL matches, not the rest of the year. However, it is likely that pollution exposure during cricket matches themselves contributes to a bowler's ability to tolerate pollution while playing cricket, and this measure appropriately captures that type of exposure.

Table 2 presents six additional alternative definitions of long-term PM2.5. Our preferred definition for long-term exposure to PM2.5 is the mean exposure to PM2.5 across all IPL games (in all seasons) for a given bowler. This is reported in the first row and has as many distinct values as there are bowlers. The advantage of this measure is that it is bowler-specific, reflecting the fact that the bowler is the unit of observation that could potentially adapt over time. The downside of this measure is that it incorporates future data on PM2.5 exposure into past matches. However, if this results in classical measurement error, this noise in measurement would only attenuate results. In addition, we find that pollution levels are serially correlated within player.

In the second row, we calculate a similar measure, but instead of taking the mean across all games in the IPL (2008-2022) we split the panel in two (2008-2014 and 2015-2022) and take the mean exposure to PM2.5 across games in the first half of the panel. Since there are fewer distinct players in the first half of the IPL, the number of distinct values is substantially smaller. This measure addresses the temporal problem of the first measure since we can restrict estimating to the second half of the panel, using the first half as a baseline. A major drawback in doing so, however, is loss of statistical power from loss of sample size.

The remaining four rows of the table report means of PM2.5 in the location of the home-stadium for a given IPL team. There are 13 teams that have been in the IPL, so there are 13 distinct values.[15] In row 3, we calculate this mean over the 10 years prior to the IPL (1998-2007) to get a baseline exposure level. In row 4, we calculate the mean over the period of the IPL (2008-2022). In the next two alternative definitions, we account for the fact that IPL games occur only in three months of the year—March, April, and May—and construct measures analogous to those in rows 3 and 4. These measures all assign long-term pollution to a bowler on the basis of the team they play on, but since they can switch teams between seasons, these measures are not bowler-specific, but bowler-team specific. We include all alternative definition of long-term exposure as robustness checks.

[15] There are 10 teams in the IPL currently, but some have come and gone, leading to 13 distinct teams over the 2008-2022 period.

5.4 Hetereogeneity in the short-run effect of PM2.5 by past PM2.5 exposure

We hypothesize that short-run responses to PM2.5 levels on the day of the match vary by how much pollution a cricket player is exposed to in the past. To quantify this effect, we interact the realization of pollution in the day and location of the match, $\text{PM2.5}_{\ell d}$, with past exposure. The medium-term exposure levels defined in Section 5.3.1 vary at the team-match level whereas the long-term exposure levels defined in Section 5.3.2 vary at the bowler (or bowler-season) level. Each measure therefore requires a slightly different econometric specification to correctly identify the extent to which past exposure mediates the harms of present exposure.

Beginning with the medium-term measure in Equation 5, we estimate:

$$\begin{aligned} R_{ij\ell t} = &\beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{J(j)d} + \beta_3 \text{PM2.5}_{J(j)d} \\ &+ \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \Lambda_i + \Delta_j + \varepsilon_{ij\ell t} \end{aligned} \tag{8}$$

where $\text{PM2.5}_{J(j)d}$ is defined as the mean level of PM2.5 exposure in the past X days from day d for bowler j who is on team J, and other variables are as defined in Equation 2. We include the full interaction of present PM2.5 and past PM2.5 to disambiguate the level-effect of past PM2.5 itself from the modulating effect that past exposure may have on the harms of present exposure. Importantly, because this measure of past exposure is not fixed for an individual, it is not collinear with bowler fixed effects.

We estimate analogous equations for the measure in Equation 6

$$\begin{aligned} R_{ij\ell t} = &\beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \sum_{d=1}^{X} \mathbf{1}(PM2.5_{J(j)d} > Z) + \sum_{d=1}^{X} \mathbf{1}(PM2.5_{J(j)d} > Z) \\ &+ \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \Lambda_i + \Delta_j + \varepsilon_{ij\ell t} \end{aligned} \tag{9}$$

and the measure in Equation 7:

$$
\begin{aligned}
R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} &+ \beta_2 \text{PM2.5}_{\ell d} \times \sum_{d=1}^{X} [\mathbf{1}(PM2.5_{dj} > Z)](PM2.5_{dj} - Z) \\
&+ \sum_{d=1}^{X} [\mathbf{1}(PM2.5_{dj} > Z)](PM2.5_{dj} - Z) \\
&+ \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \Lambda_i + \Delta_j + \varepsilon_{ij\ell t}.
\end{aligned} \tag{10}
$$

For the measure of long-term exposure that is fixed for each bowler in Section 5.3.2, we estimate:

$$
R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_j + \phi_i + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \tag{11}
$$

where PM2.5_{j0} is long-term exposure to pollution for a given bowler, and other variables are as defined in Equation 2. Identification in Equation 11 comes from the fact that different teams' home stadiums experience varying pollution levels, generating quasi-random variation in players' typical exposure levels. Unlike in the specifications for past PM2.5 exposure in the medium-term—which vary across time—in Equation 11, we omit the term for past exposure, PM2.5_{j0}, from the estimation since it is fixed within bowler and therefore would be collinear with bowler fixed effects.

5.5 Robustness: alternative measures of long-term PM2.5

Table 2: 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. [b] IPL 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.

We estimate Equation 11 for the long-term definitions of PM2.5 in rows 1 and 2 of Table 2. For the definitions in rows 3-6, we alter the specification to include bowling team-by-year fixed effects instead of bowler fixed effects, since the assignment of long-term PM2.5 to each bowler changes depending on what team they are on in each year.[16] We estimate the slightly modified specification:

$$R_{ij\ell t} = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_{yJ(j,y)} + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \tag{12}$$

where $\psi_{jJ(j,y)}$ represent bowling team-by-year fixed effects $(J(j,y))$ is a function that maps bowler j in year y to team J.

[16] Note that we include both bowling team-by-year fixed effects and stadium-by-year fixed effects since these represent two different things: the stadium is the stadium at which the game is played (which may or may not be a home stadium for either team), while the bowling team-by-year fixed effects control for the team's home stadium.

6 Results

6.1 Main finding: PM2.5 exposure and run-scoring probability

We begin by examining the baseline relationship between match-day air pollution and cricket performance. Column 1 of Table 3 shows that a one-unit increase in PM2.5 (measured in 10 micrograms per cubic meter) is associated with a 0.41 percentage point increase in the probability of a bowler conceding a run on a given delivery. When we include weather controls (column 2), the magnitude of the effect drops by 14 percentage points. We attribute this decline to two sources of omitted variable bias. First, PM2.5 and humidity are negatively correlated (with a Pearson correlation coefficient of -0.62), and run-scoring is negatively associated with humidity. Therefore, controlling for humidity controls takes away the part of the effect of PM2.5 on run-scoring that comes from the fact that higher PM2.5 days also tend to be lower humidity days. Second, PM2.5 and temperature are weakly positively correlated[17] (with a Pearson correlation coefficient of 0.22), and run-scoring is positively associated with temperature; controlling for temperature isolates the effect of PM2.5 independent of its association with higher temperatures. Once we include fixed effects for bowler, batter, stadium-by-year, innings, and over, however, we find that the magnitude of the effect is comparable to that in the naive regression. Since the outcome of run-scoring is the result of a complex array of factors, controlling for these fixed effects reduces residual variation in the outcome not attributable to our regressor of interest, PM2.5.

The consistently positive relationship between pollution and run-scoring probability reveals an important asymmetry in how air quality affects different aspects of cricket performance. Since each delivery represents an interaction between a bowler and batter, an increased probability of scoring implies that pollution disproportionately impairs bowling performance relative to batting performance. This asymmetry is intuitive given the physical demands of each role: bowlers must maintain a running approach before each delivery, requiring

[17] The association between temperature and PM2.5 varies depending on many meteorological factors and is sometimes negative; in IPL games in our sample, however, it is a positive correlation.

Table 3: PM2.5 exposure and run-scoring probability

	(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	0.0072
				(0.0064)	(0.0069)	(0.0060)
Q3 (Match-day PM2.5)				0.00015	-0.00065	0.0099
				(0.0065)	(0.0073)	(0.0069)
Q4 (Match-day PM2.5)				0.0086	0.0047	0.013
				(0.0076)	(0.0096)	(0.0086)
Q5 (Match-day PM2.5)				0.023***	0.017	0.027***
				(0.0069)	(0.010)	(0.0099)
Weather controls		✓	✓		✓	✓
Bowler FE			✓			✓
Batter FE			✓			✓
Stadium-by-year FE			✓			✓
Home stadium FE			✓			✓
Innings FE			✓			✓
Over FE			✓			✓
N	183,572	183,572	183,556	183,572	183,572	183,556
R^2	0.00028	0.00039	0.052	0.00032	0.00043	0.052

Notes. The outcome variable is a binary indicator equal to 1 if at least one run is scored on a delivery and 0 otherwise. PM2.5 is measured in 10 $\mu g/m^3$. Quantiles 1 through 5 of PM2.5 are separated at 27, 34, 41, and 53 $\mu g/m^3$, respectively. Columns (1–3) present results from regressions of this run indicator on match-day PM2.5 levels (measured in 10 $\mu g/m^3$ per cubic meter). Columns (4–6) present results from regressions using PM2.5 quintile indicators, where the lowest quintile (Q1) is the omitted category. Columns (3) and (6) include fixed effects for individual batters, bowlers, stadium-by-year, home stadium, innings, and over. Columns (2), (3), (5) and (6) include controls for weather conditions including temperature, temperature squared, relative humidity, precipitation, solar radiation, and wind speed. Following Correia (2015), 16 singleton observations are dropped in columns (3) and (6). Standard errors (in parentheses) are two-way clustered at the match and bowler levels. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

sustained cardiovascular exertion, while batters engage in more intermittent activity. This pattern suggests that the cardiovascular stress from pollution exposure is a limiting factor for activities requiring continuous aerobic effort.

6.2 Robustness of main result to PM2.5 and weather interactions

Our identifying assumption is that contemporaneous exposure to PM2.5 is as-good-as randomly assigned with respect to innate performance potential of players —in other words, it is not the case that high-performing players are systematically more (or less) exposed to PM2.5 than low-performing players. While the IPL scheduling rules strongly support this assumption, we also need to rule out sources of omitted variable bias. Specifically, since PM2.5 levels are in part a function of meteorological conditions, and weather conditions such as temperature, humidity, and wind all may affect player performance, it is essential to flexibly control for these factors. Our baseline specification in Equation 2 accounts for the linear effect of each of these weather conditions on performance but does not account for the role of non-nonlinearities as well as potential interactions between each of these aspects of weather and PM2.5 exposure.

To isolate the direct effect of PM2.5 on performance, we perform Post-Double Selection (PDS) Lasso following the method outlined in Belloni et al. (2014). We first regress the outcome of interest (a binary variable of whether one or more runs was scored from a delivery) on a high-dimensional set of controls with bowler fixed effects and heteroskedasticity robust standard errors clustered at the match level. The set of controls includes the the dummy variables for the home stadium for the bowler and the home stadium for the batter (2 variables), linear and quadratic terms of all six weather variables (12 variables), as well as the interactions between each of the linear weather terms with each other (6 choose 2 results in 15 variables), and with PM2.5 (6 variables), resulting in a regression of the outcome on 35 variables. Notably, in this first step, the regression does not include the PM2.5 term—our main regressor of interest—itself. The PDS Lasso method selects which of these 35 controls are most predictive of the outcome. Empirically, we find that none of these 35 controls are selected. This result indicates that none of these factors are substantially predictive of the outcome of run-scoring after accounting for the penalty of including additional variables. This result is consistent whether we include the bowler fixed effects or not and whether we cluster standard errors by match, strongly suggesting that these weather conditions (and their interactions with each other and PM2.5) are not strongly affecting bowler performance.

The second step of the PDS Lasso procedure is to regress the regressor of interest (here, PM2.5) on the high-dimensional set of controls to identify which controls strongly predict the regressor of interest. Empirically, we find that two terms are selected: (1) the interaction of temperature with relative humidity and (2) the interaction of relative humidity with wind speed. Each of these makes sense given the role of temperature, humidity, and wind in shaping how particulate matter moves throughout space. Notably, our weather variables are sourced from ERA-5, which itself is one of the inputs in the machine learning model that Wang et al. (2024) uses to predict PM2.5, so it is reassuring that some of these variables are selected.

The final step of PDS Lasso is to perform OLS of the outcome on the regressor of interest and the union of the set of variables selected in the first step (in this case, the empty set) and the variables selected in the second step. Because we selected the variables that contribute to predicting the outcome (run-scoring) and those that contribute to predicting the regressor of interest (PM2.5), this OLS regression identifies the effect of PM2.5 on run-scoring. Reassuringly, we find a coefficient of 0.0035 ($p < .01$), which is very close to the coefficient of 0.0041 ($p < 0.05$) from the baseline specification in Equation 2 reported in Column (3) of Table 3.

As a further robustness check to compare the magnitude of the effect from PM2.5 to that of temperature, we repeat the PDS Lasso procedure described with temperature as the regressor of interest. We find that in the OLS regression with PDS-selected variables, the coefficient on temperature (0.0014, $p = 0.102$) is about one-third of the magnitude of that of

PM2.5, is not statistically significant, and has a 95% confidence interval that spans on either side of zero. This finding suggests that, at least in this empirical setting, PM2.5 has a bigger impact on performance than does heat.[18]

6.3 Non-linearities

To investigate potential non-linearities in this relationship, columns 4–6 of Table 3 replace the continuous PM2.5 measure with indicators for PM2.5 quintiles. The results suggest that the impact of pollution is concentrated in the highest quintile of exposure. In our preferred specification with full controls (column 6), exposure to PM2.5 levels in the top quintile is associated with a 2.8 percentage point increase in the probability of scoring a run, relative to the lowest quintile, an effect that is nearly seven times the magnitude of the baseline effect of 0.0041 percentage points. The effects for the second through fourth quintiles are smaller in magnitude and generally not statistically significant, suggesting a non-linear dose-response relationship where the most severe pollution episodes have disproportionate effects on performance. This is a surprising finding considering that even lower quintiles of PM2.5 in our sample are well-above WHO thresholds: the first quantile ends at 27 $\mu g\, m^{-3}$, nearly twice the WHO recommended safe daily exposure level of 15 $\mu g\, m^{-3}$. The second quantile begins at 34, the third at 41, and the fifth at 53 $\mu g\, m^{-3}$. See Figure 7 for the full distribution of PM2.5 in our data relative to WHO thresholds.

This non-linear response pattern is illustrated graphically in the spline regression specifications in Figures 12, and 13. Figure 12 displays the result of estimating Equation 4 with a linear spline ($p = 1$) with 4 knots. In Figure 12, the knots are placed at WHO thresholds for PM2.5 concentrations, which are based on the health literature about expected effects. In Figure 12, the interpretation of the y-axis is the change in run scoring probability relative to the mean run scoring probability for a given bowler. We hypothesize that we may see differences in marginal effects when crossing these thresholds, which does appear to be the

[18]This finding is notable given that the matches are played in generally warm conditions, ranging from 71 to 99 degrees Fahrenheit, with a mean of 85 degrees Fahrenheit.

Table 4: Evidence of Adaptation to Air Pollution Climates

	(1)	(2)	(3)	(4)	(5)	(6)
		$\mathbb{1}$ (At least one run scored)				
Match PM2.5	0.0041***	0.0027*	0.0043***	0.0042**	0.0041**	0.013**
	(0.0010)	(0.0014)	(0.0016)	(0.0016)	(0.0017)	(0.0052)
Match PM2.5 X Bowler PM2.5						-0.0019*
						(0.0011)
Weather controls		✓	✓	✓	✓	✓
Bowler FE					✓	✓
Striker FE					✓	✓
Stadium-by-year FE			✓	✓	✓	✓
Over FE			✓	✓	✓	✓
Match innings FE				✓	✓	✓
N	183,572	183,572	183,558	183,558	183,556	183,556
R^2	0.00028	0.00039	0.046	0.046	0.052	0.052

Notes. The outcome variable is a binary indicator equal to 1 if at least one run is scored on a delivery and 0 otherwise. Bowler PM2.5 is defined as the mean PM2.5 in all of bowler's games in the IPL. Standard errors are clustered two-way at the match and bowler level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

case: the slope is positive—indicating a deterioration of performance—as PM2.5 increases from low levels to midrange levels, then becomes negative around WHO Interim Target 2 of 50 $\mu g\,m^{-3}$. It then becomes more steeply positive beyond WHO Interim Target 3 of $\mu g\,m^{-3}$. However, these changes in slopes are not statistically significant.

Figure 13 estimates Equation 4 with $p = 3$ and $k = 5$, but sets the knots dynamically based on quintiles of the PM2.5 distribution in our data. We see a similar pattern of non-linearity, reinforcing the finding that higher levels of pollution have a more extreme marginal effect of pollution on performance.

6.4 Acclimation to air pollution exposure

6.4.1 Acclimation to medium-term exposure

We adopt a data-driven approach to measuring acclimation to medium-term exposure. We explore the sensitivity of performance to past exposure with varying measures and time windows and see which most substantially mediates the harms of contemporaneous exposure to air pollution. We find that the measures of medium-term exposure in Equations 5, 6, and 7 consistently agree with one another in finding that the marginal impact of contemporaneous

Figure 12: Effect of PM2.5 on Run Probability (Linear Spline, WHO Thresholds)

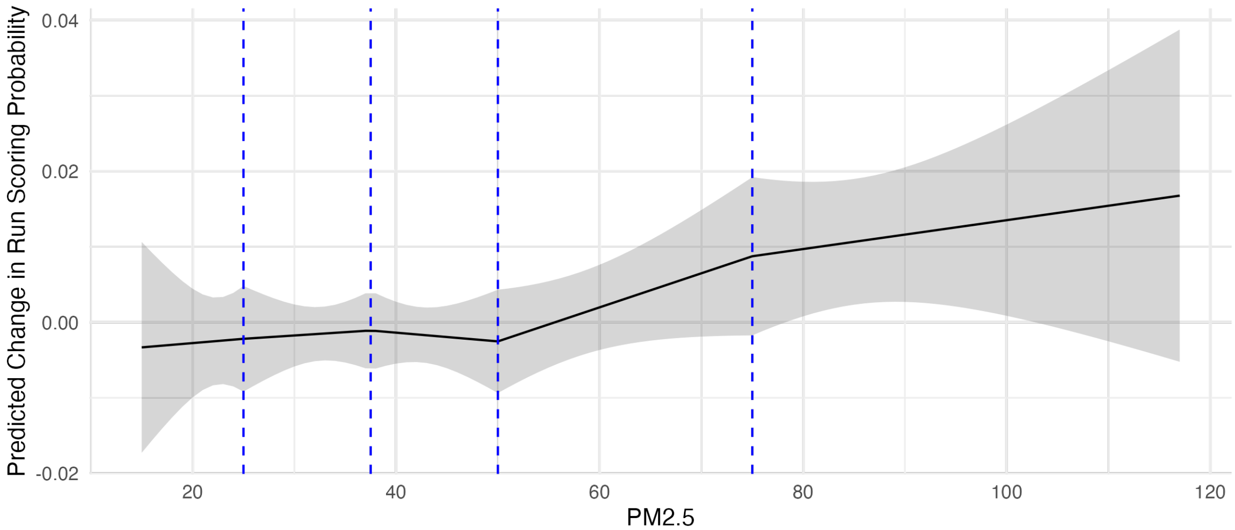
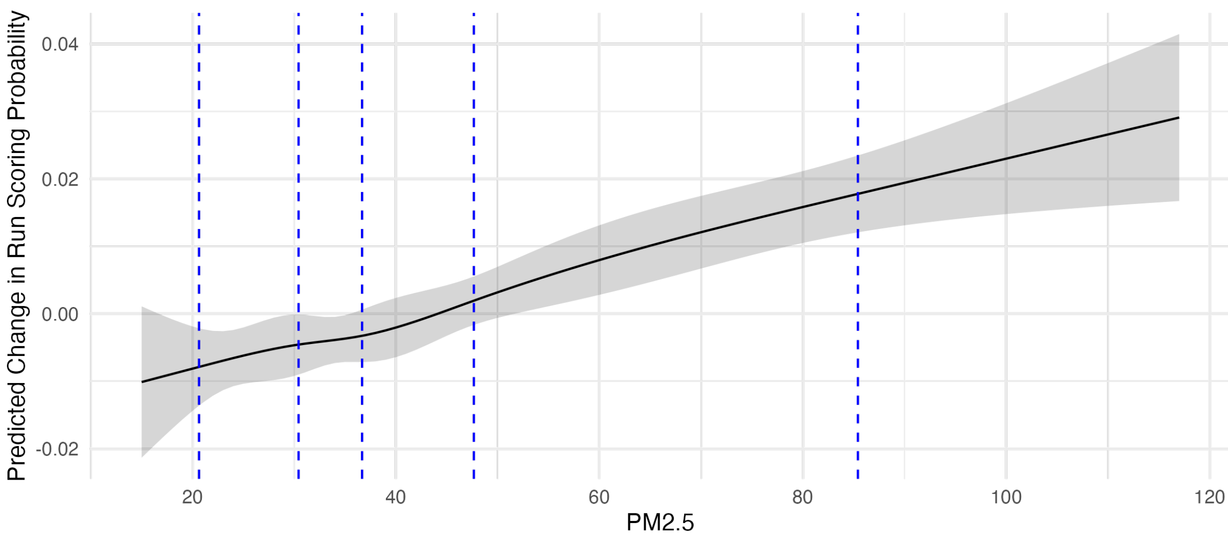


Figure 13: Effect of PM2.5 on Run Probability (Restricted Cubic Spline, Quantile Knots)



Notes. Figures 12 and 13 display the predicted change in run scoring probability in a given batter-bowler interaction as a function of PM2.5 concentration on match-day. All variables are demeaned within bowler. The interpretation of the y-axis is the change in run scoring probability relative to the mean run scoring probability for a given bowler. In Figure 12, the regression specification is a linear spline with 4 knots. The knots are specified at WHO target air pollution levels: Air Quality Guideline, 25 $\mu g\, m^{-3}$, Interim Target 3, 37.5 $\mu g\, m^{-3}$, Interim Target 2, 50 $\mu g\, m^{-3}$, and Interim Target 1, 75 $\mu g\, m^{-3}$. In Figure 13, the regression specification is a restricted cubic spline with 4 knots. The knots are specified at the 20th, 40th, 60th, and 80th quantiles of the PM2.5 distribution: 21, 30, 37, 48, and 85 $\mu g\, m^{-3}$, respectively. Knot locations in both figures are indicated with dashed verticle lines.

Table 5: Evidence of Acclimation to Air Pollution

	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$ (At least one run scored)					
Contemporaneous PM2.5	0.0041***	0.0027*	0.0042***	0.0042**	0.0041**	0.0066*
	(0.0010)	(0.0014)	(0.0016)	(0.0016)	(0.0017)	(0.0034)
Past PM2.5						0.0061*
						(0.0034)
Contemporaneous PM2.5 X Past PM2.5						-0.00055
						(0.00063)
Weather controls		✓	✓	✓	✓	✓
Home stadium dummies		✓	✓	✓	✓	✓
Stadium-by-year FE			✓	✓	✓	✓
Over FE			✓	✓	✓	✓
Innings FE				✓	✓	✓
Bowler FE					✓	✓
Batter FE					✓	✓
N	183,572	183,572	183,558	183,558	183,556	183,556
R^2	0.00028	0.00041	0.046	0.046	0.052	0.052

Notes. The outcome variable is a binary indicator equal to 1 if at least one run is scored on a delivery and 0 otherwise. Past PM2.5 is defined as the mean PM2.5 exposure for the bowler in the past 30 days. Standard errors are clustered two-way at the match and bowler level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

PM2.5 exposure is smaller when one is exposed to more PM2.5 in the past. However, each measure offers a different perspective on the time frame over which this acclimation effect materializes, and the magnitude of the acclimation effect.

We define the "acclimation effect" as the difference in the marginal effect of PM2.5 for those with past exposure at the 95$^{\text{th}}$ percentile relative to those with median levels of past exposure:

$$AcclimationEffect = \frac{ME_{p50} - ME_{p95}}{ME_{p50}} \tag{13}$$

where ME_{pX} represents the marginal effect of an increase of PM2.5 by 10 $\mu g\, m^{-3}$ on run-scoring probability for bowlers at the X^{th} percentile of past exposure to PM2.5.[19]

For example, using the measure in Equation 5 and a window of 30 days of past exposure, we find that the marginal effect of a 10 $\mu g\, m^{-3}$ PM2.5 on run-scoring is 0.0042 percentage points for a bowler with median past exposure to PM2.5 in the past 30 days (exposed to 44.0

[19]For the measure in Equation 6, we slightly modify this definition to be $\frac{ME_{p50} - ME_{max}}{ME_{p50}}$ to account for the comparatively lower variation in number of days above a particular threshold. The logic of comparing the marginal effect of contemporaneous PM2.5 for someone exposed at a "high" level relative to the typical level remains the same.

$\mu g\,m^{-3}$ of PM2.5 per day on average). However, for a bowler at the 95th percentile of past PM2.5 exposure in the same 30 day period (exposed to 70.6 $\mu g\,m^{-3}$ of PM2.5 per day on average), the marginal effect is 0.0027 percentage points, 34.8% smaller. Figure 14 shows how the marginal effects vary as more past days of exposure are included in the measure. Figure 15 plots the ratio in Equation 13 as a function of the number of past days of exposure included. Empirically, we find that the acclimation effect is most pronounced when we include the previous 24 days of exposure in a player's mean exposure level as defined in Equation 5.

The three measures of past PM2.5 in the medium-term largely agree on the time window over which this "acclimation effect" is largest in magnitude. For the measure in Equation 5, it is largest at 24 days (41.1%); for the measure in Equation 6 with $Z = 50$ it is largest at 32 days (54.0%), and for the measure in Equation 7 with $Z = 50$ it is largest at 24 days (40.3%). Table 6 summarizes these results. The only discrepancies from the time-window of approximately 30 days occur when using the measure in Equation 6 for lower values of Z. This result stems from the fact that there is comparatively less variation in the number of days above lower thresholds, since X% of days in our sample are above 25 $\mu g\,m^{-3}$ and Y% are above 37.5 $\mu g\,m^{-3}$. It is reassuring to find that the acclimation effect reaches its peak at approximately 30 days when Z is a level that results in more variation within the sample. It is notable that the acclimation effect is approximately twice the magnitude when counting days above 75 $\mu g\,m^{-3}$ in Equation 6. This result suggests that past-exposure to the most extreme days provides more of an insulating effect against present exposure, consistent with the mechanism of glutathione production augmentation outlined in Section 3.1.

Since the measures largely agree on the time frame over which past exposure has the most pronounced effect on mitigating the harms from present exposure, we choose the measure in Equation 5 as our main measure of past exposure to PM2.5 in the medium term since it is conceptually simplest, noting that the results are robust to the choice of alternate measures. Table 6 presents results from estimating this equation for a time-window of 30 days in column (6). We find that an increase in contemporaneous PM2.5 by 10 $\mu g\,m^{-3}$ is associated with a

Figure 14: Marginal Effect of PM2.5 on Run Probability for Varying Exposure Windows

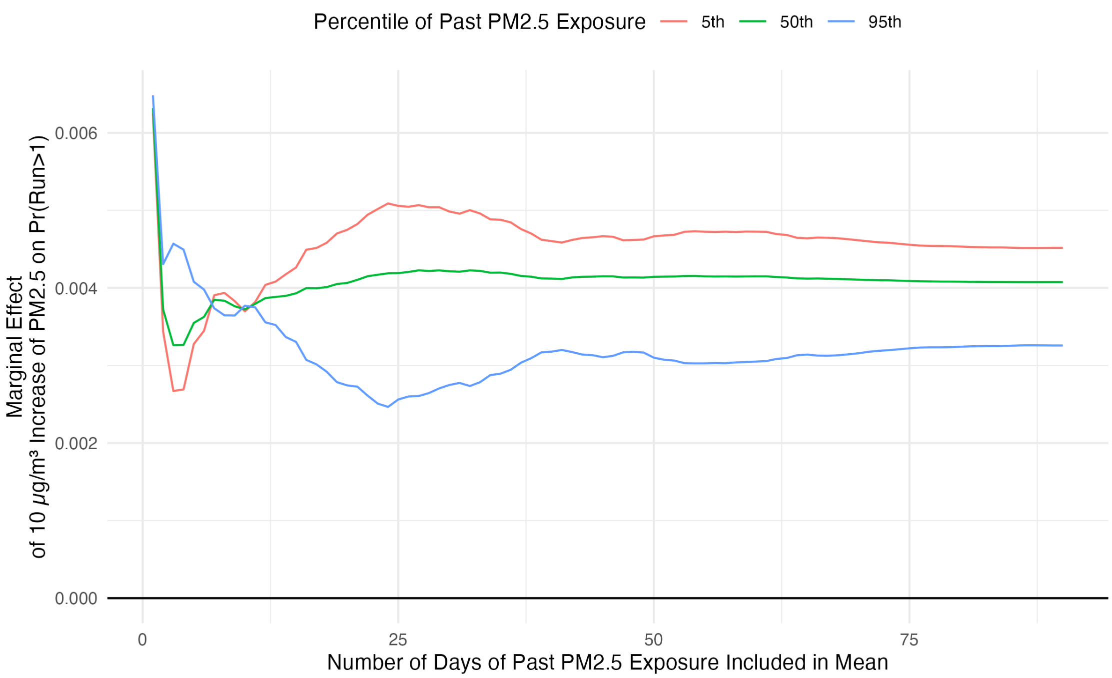
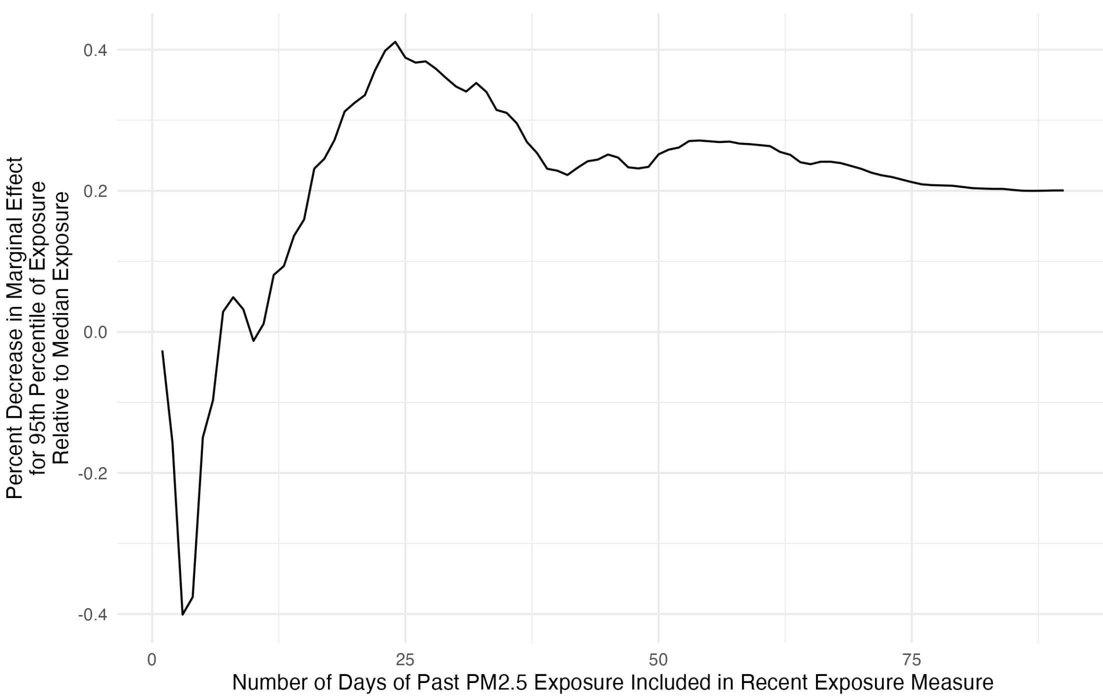


Figure 15: Magnitude of Adaptation Effect for Varying Exposure WIndows



Notes. Figure 14 show the marginal effect of PM2.5 exposure for the 5th, median, and 95th percentile of past exposure as estimated in Equation 8. Figure 15 displays the measure in Equation 13 as a function of the number of days of past exposure included. It reaches its maximum when the blue and green lines (representing the 95th and 50th percentile of past exposure) in Figure 14 are farthest apart.

Table 6: Medium-term acclimation effects: magnitude and time window across measures of past $PM_{2.5}$

Measure	Z **value**	**Days (largest effect)**	**Magnitude (%)**
Equation 5	–	24	41.1
Equation 6	25	90	55.0
	37.5	81	68.6
	50	32	54.0
	75	24	81.5
Equation 7	25	24	40.5
	37.5	24	42.7
	50	24	40.3
	75	24	44.3

0.66 percentage point increase in the probability of a bowler conceding a run. Similarly, an increase by 10 $\mu g\,m^{-3}$ in past PM2.5 exposure, as measured by the bowler's mean exposure in the past 30 days, is also associated with an increase in run-concession probability by a similar magnitude of 0.61 percentage points. We cannot reject the null hypothesis that these two coefficients are not statistically different from one another (p¡0.81). However, it is important to note that an increase in the 30-day mean exposure by 10 $\mu g\,m^{-3}$ entails increasing exposure on each of the 30 days by 10 $\mu g\,m^{-3}$ on average, a substantially higher dose of PM2.5 exposure than increasing PM2.5 by 10 $\mu g\,m^{-3}$ on the day of the match. This suggests that contemporaneous exposure to PM2.5 is more harmful to present performance than is past exposure. Past exposure itself also harms performance, but it lessense the detrimental impact of present day increases, consistent with the finding in the environmental toxicology literature that the lungs of mice who are exposed consistently to polluted air become "refractory to further injury" (West et al., 2003).

We plot the marginal effect of an increase in contemporaneous PM2.5 exposure on run-scoring as a function of past accumulated exposure to PM2.5 (defined as the 30-day mean) in Figure 16, noting that the effect is more muted with higher levels of past exposure.

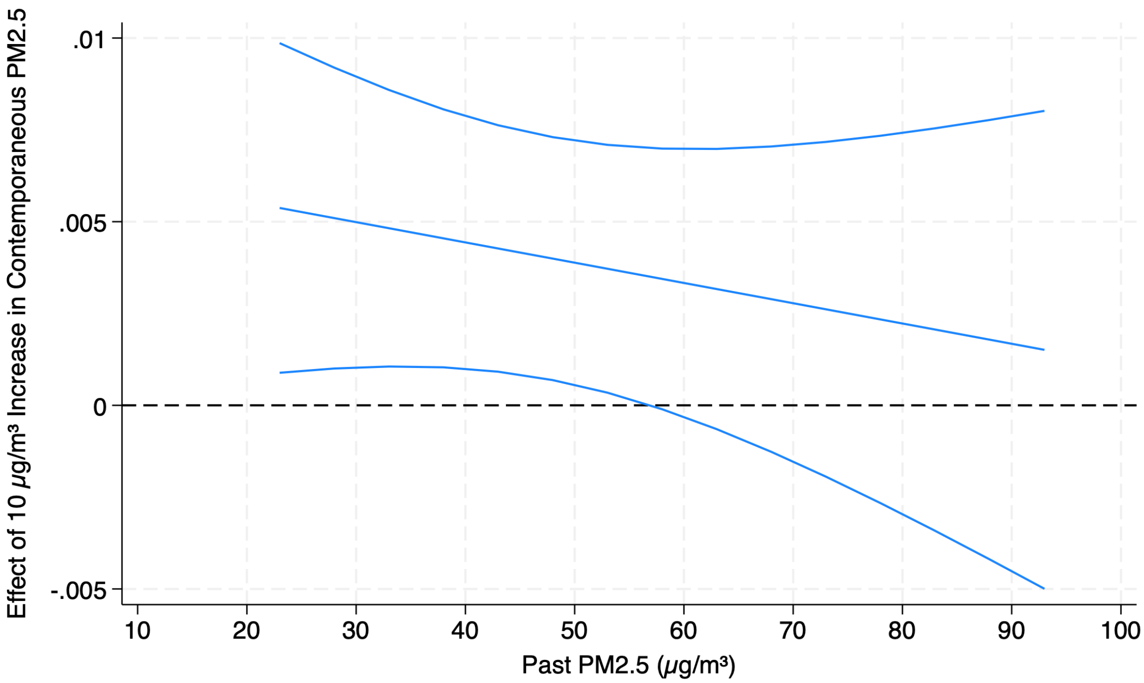
6.4.2 Acclimation to long-term exposure

In Table 5, we see evidence suggesting that cricket players have a heterogeneous response to short-term pollution shocks depending on their long-term exposure to pollution, defined as their mean exposure across all games in the IPL. While a 10 $\mu g\,m^{-3}$ increase in PM2.5 is associated with a 0.41 percentage point increase in conceding a run in the specification with weather controls and fixed effects in column (5), this effect is mitigated when bowlers have a higher long-term mean. The marginal effect of match-day PM2.5 in Equation 11 is $\beta_1 + \beta_2$. When long-run PM2.5 for a bowler is zero, the marginal effect of PM2.5 on the likelihood of conceding a run is a 1.3 percentage point increase. However, as long-run bowler PM2.5 increases, the marginal effect decreases in magnitude, since the sign of β_2 is negative. For example, at the mean level of long-run bowler PM2.5 of 42.62 $\mu g\,m^{-3}$, the marginal effect of PM2.5 on run-scoring probability is $0.013 + (-0.0019) \times 42.62/10 = 0.0049$[20], or a 0.49 percentage point increase in the likelihood of conceding a run. Figure 17 plots how this effect differs for the range of long-run PM2.5 describing nearly all bowlers in our sample. With high levels of long-run bowler PM2.5 close to 60 $\mu g\,m^{-3}$, the marginal effect of PM2.5 during game day attenuates close to zero, though is still somewhat harmful.

Table 7 presents the results of Equations 11 and 12 for the alternative long-term PM2.5 measures in Table 2. Although the point estimates on the interaction terms are no longer statistically significant, we find that the negative sign on the interaction remains consistent, suggesting that long-term exposure to PM2.5 mitigates the impact of PM2.5 on game-day. The smaller magnitude and lack of statistical significance in these results suggests that the way we measure long-term PM2.5 exposure matters. A person-specific measure results in a larger estimate than a place-based estimate.

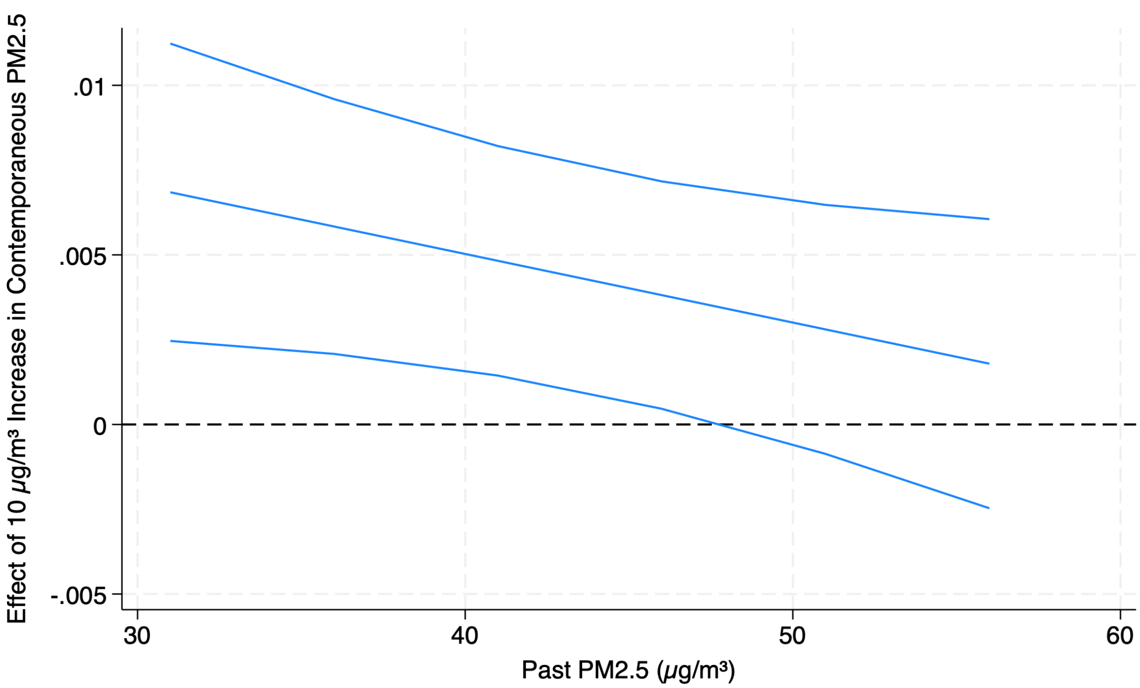
[20]Note that PM2.5 is in units of 10 $\mu g\,m^{-3}$ in all regression specifications.

Figure 16: Marginal Effect of Contemporaneous PM2.5 on Pr(Run) by Bowler 30-day mean PM2.5



Notes. This figure plots the marginal effect (with 95% confidence intervals) of Contemporaneous PM2.5 on the probability of run-scoring by varying levels of past PM2.5 exposure, defined as a bowler's 30-day mean-level of exposure, as estimated in Equation 8.

Figure 17: Marginal Effect of Contemporaneous PM2.5 on Pr(Run) by Bowler Long-term PM2.5



Notes. This figure plots the marginal effect (with 95% confidence intervals) of PM2.5 on the likelihood of run-scoring in Equation 11 by varying levels of $PM2.5_{j0}$.

Table 7: Robustness of Long-term PM2.5 Measures

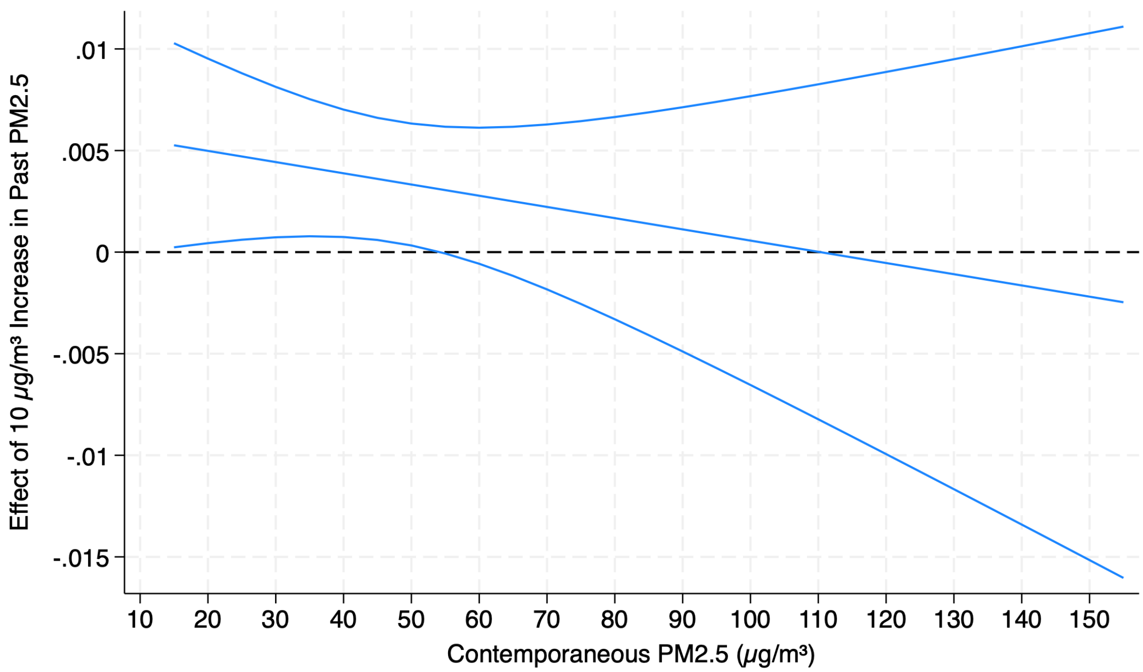
	(1)	(2)	(3)	(4)	(5)	(6)
	$\mathbb{1}$ (At least one run scored)					
Match PM2.5	0.0052	0.0067**	0.0062**	0.0067**	0.0058**	0.0071***
	(0.0060)	(0.0030)	(0.0025)	(0.0028)	(0.0024)	(0.0027)
PM2.5 X Bowler PM2.5 (IPL Seasons 2008-2014)	-0.0012					
	(0.0013)					
PM2.5 X Bowler Team Stadium PM2.5 (Preseason)		-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		✓	✓	✓	✓	✓
Striker FE	✓	✓	✓	✓	✓	✓
Stadium-by-year FE	✓	✓	✓	✓	✓	✓
Over FE	✓	✓	✓	✓	✓	✓
Innings FE	✓	✓	✓	✓	✓	✓
N	53,607	183,558	183,558	183,558	183,558	183,558
R^2	0.055	0.047	0.047	0.047	0.047	0.047

Notes. The outcome variable is a binary indicator equal to 1 if at least one run is scored on a delivery and 0 otherwise. Standard errors are clustered two-way at the match and bowler level. Column (1) restricts the sample to bowlers who appear in both the first and second half of the panel (first half: 2008-2014, second half: 2015-2022) and analyzing their performance in the second half using their mean PM2.5 exposure during IPL games in the first half. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.5 Tradeoff between exposure and performance

Our results consistently show that exposure to particulate matter—whether contemporaneously or accumulated past exposure–harms performance. However, since past exposure softens the blow of present increases (as indicated by the negative coefficient on the interaction term, β_2 in the Equations in Section 5.4, there is in principle a level of past exposure that would result in benefits to present performance. Figure 18 shows how the marginal effect of an increase in past PM2.5 exposure affects performance in the present as present exposure to PM2.5 varies. When present exposure to PM2.5 reaches extremely high-levels—approximately 110 $\mu g\, m^{-3}$—increases in past exposure to PM2.5 start to have a beneficial effect on performance (here, reducing the probability that a bowler concedes a run). However, fewer than 2% of deliveries take place in matches where PM2.5 exceeds this level. This sparsity of data is reflected in the widening of the 95% confidence intervals at the high-end of contemporaneous PM2.5—although we find suggestive evidence that the marginal effect of increases in past PM2.5 reduces the harm of present exposure to PM2.5, it is not conclusive that this acclimation effect ever reaches a large enough magnitude to make the impact of past PM2.5 on performance actually beneficial.

Figure 18: Marginal Effect of Past PM2.5 on Pr(Run) by Contemporaneous PM2.5



Notes. This figure plots the marginal effect (with 95% confidence intervals) of past PM2.5, defined as the 30-day mean as in Equation 5, on the probability of run-scoring for varying levels of contemporaneous PM2.5 as estimated in Equation 8.

7 Discussion and conclusion

These results paint a consistent picture of how air pollution affects labor productivity in a setting that combines aspects of physical and cognitive performance. First, pollution has significant effects on worker performance, with particularly strong impacts at high exposure levels. Second, these effects are asymmetric across player roles: the baseline results in Table 3 show that pollution increases run-scoring probability, suggesting bowlers are more affected than batters. The results in Table 5 show that bowlers are less-harmed by high pollution on match-day when they are used to higher levels of pollution in the long-run, all else equal. Together, these preliminary findings suggest that both the acute effects of pollution are mediated by past exposure. We emphasize, however, that long-term exposure to air pollution still may worsen productivity—as well as health–over the long-term, counteracting this positive side-effect in the short-term.

References

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.

Aguilar-Gomez, S., H. Dwyer, J. G. Zivin, and M. Neidell (2022): "This Is Air: The "Nonhealth" Effects of Air Pollution," *Annual Review of Resource Economics*, 14, 403–425, publisher: Annual Reviews.

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.

Bartlett, R. M. (2003): "The science and medicine of cricket: an overview and update," *Journal of Sports Sciences*, 21, 733–752, publisher: Routledge.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *The Review of Economic Studies*, 81, 608–650.

Burke, M. and K. Emerick (2016): "Adaptation to Climate Change: Evidence from US Agriculture," *American Economic Journal: Economic Policy*, 8, 106–140.

Burke, M., M. Zahid, M. C. M. Martins, C. W. Callahan, R. Lee, T. Avirmed, S. Heft-Neal, M. Kiang, S. M. Hsiang, and D. Lobell (2024): "Are We Adapting to Climate Change?" *NBER Working Paper*.

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.

Chang, T. Y., J. Graff Zivin, T. Gross, and M. Neidell (2019): "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China," *American Economic Journal: Applied Economics*, 11, 151–172.

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

Correia, S. (2015): "Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix," Working Paper.

Dell, M., B. F. Jones, and B. A. Olken (2014): "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 52, 740–798.

Dimeo, M. J., M. G. Glenn, M. J. Holtzman, J. R. Sheller, J. A. Nadel, and H. A. Boushey (1981): "Threshold Concentration of Ozone Causing an Increase in Bronchial Reactivity in Humans and Adaptation with Repeated Exposures," *American Review of Respiratory Disease*, 124, 245–248, publisher: American Thoracic Society - AJRCCM.

Dodson, R. E., E. A. Houseman, J. I. Levy, J. D. Spengler, J. P. Shine, and D. H. Bennett (2007): "Measured and modeled personal exposures to and risks from volatile organic compounds," *Environmental Science & Technology*, 41, 8498–8505.

Garcia, J., D. H. Bennett, D. Tancredi, M. B. Schenker, D. Mitchell, S. J. Reynolds, and F. M. Mitloehner (2013): "Occupational exposure to particulate matter and endotoxin for California dairy workers," *International Journal of Hygiene and Environmental Health*, 216, 56–62.

Graff Zivin, J. and M. Neidell (2012): "The Impact of Pollution on Worker Productivity," *American Economic Review*, 102, 3652–3673.

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.

Hackney, J. D., W. S. Linn, J. G. Mohler, and C. R. Collier (1977): "Adaptation to short-term respiratory effects of ozone in men exposed repeatedly," *Journal of Applied Physiology*, 43, 82–85, publisher: American Physiological Society.

Hagerty, N. (2022): "Adaptation to Surface Water Scarcity in Irrigated Agriculture," *Unpublished Manuscript*.

Hamade, A. K. and C. G. Tankersley (2009): "Interstrain variation in cardiac and respiratory adaptation to repeated ozone and particulate matter exposures," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 296, R1202–R1215, publisher: American Physiological Society.

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.

Heyes, A., M. Neidell, and S. Saberian (2016): "The Effect of Air Pollution on Investor Behavior: Evidence from the S&P 500," *NBER Working Paper*, w22753.

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.

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.

Ketterer, B., B. Coles, and D. J. Meyer (1983): "The role of glutathione in detoxication," *Environmental Health Perspectives*, 49, 59–69, publisher: Environmental Health Perspectives.

Kültz, D., J. Li, R. Sacchi, D. Morin, A. Buckpitt, and L. Van Winkle (2015): "Alterations in the proteome of the respiratory tract in response to single and multiple exposures to naphthalene," *PROTEOMICS*, 15, 2655–2668, _eprint: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/pdf/10.1002/pmic.201400445.

Lawrence, M. G. (2005): "The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications," *Bulletin of the American Meteorological Society*, 86, 225–234.

Lee, S.-H., S.-H. Hong, C.-H. Tang, Y. S. Ling, K.-H. Chen, H.-J. Liang, and C.-Y. Lin (2018): "Mass spectrometry-based lipidomics to explore the biochemical effects of naphthalene toxicity or tolerance in a mouse model," *PLOS ONE*, 13, e0204829, publisher: Public Library of Science.

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

MoneyBall (2025): "MoneyBall – The Wage Calculator of Indian Professional leagues," .

Moore, F. C. (2017): "Learning, adaptation, and weather in a changing climate," *Climate Change Economics*, 08, 1750010, publisher: World Scientific Publishing Co.

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).

Mérel, P. and M. Gammans (2021): "Climate Econometrics: Can the Panel Approach Account for Long-Run Adaptation?" *American Journal of Agricultural Economics*, 103, 1207–1238, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajae.12200.

Mérel, P., E. Paroissien, and M. Gammans (2024): "Sufficient statistics for climate change counterfactuals," *Journal of Environmental Economics and Management*, 124, 102940.

Noakes, T. and J. Durandt (2000): "Physiological requirements of cricket," *Journal of Sports Sciences*, 18, 919–929.

PETERSEN, C. J., D. PYNE, B. DAWSON, M. PORTUS, AND A. KELLETT (2010): "Movement patterns in cricket vary by both position and game format," *Journal of Sports Sciences*, 28, 45–52.

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

SCHWARTZ, J. (2000): "The Distributed Lag between Air Pollution and Daily Deaths," *Epidemiology*, 11, 320.

TIMES, T. E. (2022): "ipl valuation: IPL crosses $10 billion valuation to become a decacorn: D and P Advisory - The Economic Times," .

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.

WEI, Y., M. D. YAZDI, Q. DI, W. J. REQUIA, F. DOMINICI, A. ZANOBETTI, AND J. SCHWARTZ (2021): "Emulating causal dose-response relations between air pollutants and mortality in the Medicare population," *Environmental Health*, 20, 53.

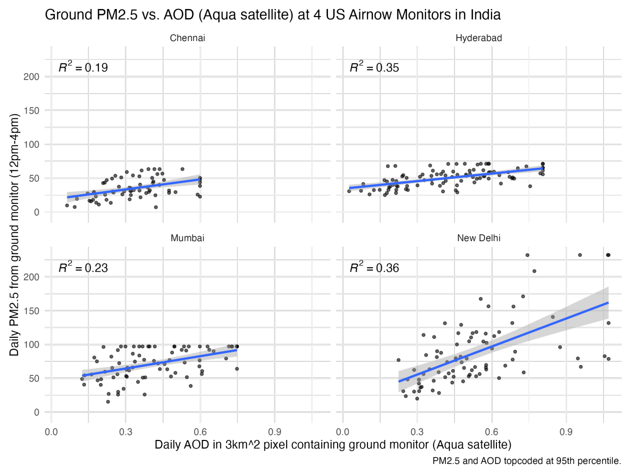
WEST, J. A. A., L. S. VAN WINKLE, D. MORIN, C. A. FLESCHNER, H. J. FORMAN, AND C. G. PLOPPER (2003): "Repeated inhalation exposures to the bioactivated cytotoxicant naphthalene (NA) produce airway-specific Clara cell tolerance in mice," *Toxicological Sciences: An Official Journal of the Society of Toxicology*, 75, 161–168.

WORLD HEALTH ORGANIZATION (2021): "WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide," Tech. rep.

ZANOBETTI, A., J. SCHWARTZ, E. SAMOLI, A. GRYPARIS, G. TOULOUMI, R. ATKINSON, A. LE TERTRE, J. BOBROS, M. CELKO, A. GOREN, B. FORSBERG, P. MICHELOZZI, D. RABCZENKO, E. ARANGUEZ RUIZ, AND K. KATSOUYANNI (2002): "The Temporal Pattern of Mortality Responses to Air Pollution: A Multicity Assessment of Mortality Displacement," *Epidemiology*, 13, 87.

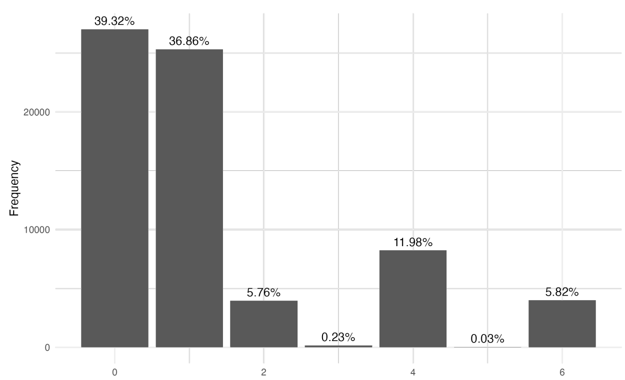
A Appendix: Tables and Figures

Figure A.1: U.S. Airnow vs. MODIS AOD



Notes. This figure displays the correlation between Aerosol Optical Depth from the MODIS Aqua satellite (which passes over India approximately once per day in the afternoon) with ground observations of PM2.5 at five cities with a U.S. AirNow ground monitor. Each dot represents a daily mean value of PM2.5. AOD is calculated as the mean AOD in the 3km x 3km pixel containing the ground monitor. Both PM2.5 and AOD are topcoded at the 95th percentile.

Figure A.2: Histogram of Runs



Notes. This figure shows the distribution of the outcome variable (runs scored) in our analysis.

B Appendix: Additional Robustness Checks

B.1 Robustness: incorporating variation in number of runs scored

In our main specifications, we define the outcome as a binary indicator for whether a bowler conceded a run. This definition abstracts away from variation in the number of runs scored, conditional on scoring any run. As a robustness check, we also estimate an ordered logit model which treats each number of runs scored as a discrete category with a higher number indicating better performance for the team opposing the bowler.

$$\log\left(\frac{P(R_{ij\ell t} \leq r)}{P(R_{ij\ell t} > r)}\right) = \beta_1 \text{PM2.5}_{\ell d} + \beta_2 \text{PM2.5}_{\ell d} \times \text{PM2.5}_{j0} + \mathbf{X}'_{\ell d}\phi + \psi_j + \delta_{\ell y} + \theta_n + \eta_o + \varepsilon_{ij\ell t} \quad (14)$$

where $R_{ij\ell t}$ is the number of runs scored ($\{0, 1, 2, 3, 4, 5, 6\}$). To ease computation and for graphically results, we also estimate a simpler version of this equation with only bowler fixed effects:

$$\log\left(\frac{P(R_{ij\ell t} \leq r)}{P(R_{ij\ell t} > r)}\right) = \beta_1 \text{PM2.5}_{\ell d} + \varepsilon_{ij\ell t}. \quad (15)$$

The results estimating Equation 14 are reported in Table A.1 and are qualitatively similar to those estimating analogous regressions in Table 5. Figures A.3 and A.4 report graphically the results of results estimating Equation 15.

Table A.1: Evidence of Adaptation to Air Pollution Climates (Ordered Logit)

	(1)	(2)	(3)	(4)	(5)	(6)
		$\mathbb{1}$ (At least one run scored)				
Match PM2.5	0.013***	0.013***	0.019***	0.021***	0.019***	0.075***
	(0.0027)	(0.0037)	(0.0054)	(0.0054)	(0.0055)	(0.020)
Match PM2.5 X Bowler PM2.5						-0.012***
						(0.0041)
Weather controls		✓	✓	✓	✓	✓
Stadium-by-year FE			✓	✓	✓	✓
Match innings FE			✓	✓	✓	✓
Over FE				✓	✓	✓
Bowler FE					✓	✓
N	183,572	183,572	183,572	183,572	183,572	183,572

Notes. The outcome variable is the number of runs scored (0 to 6). Regressions are estimated using ordered logit. Bowler PM2.5 is defined as the mean PM2.5 in all of bowler's games in the IPL. Standard errors clustered at the bowler level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

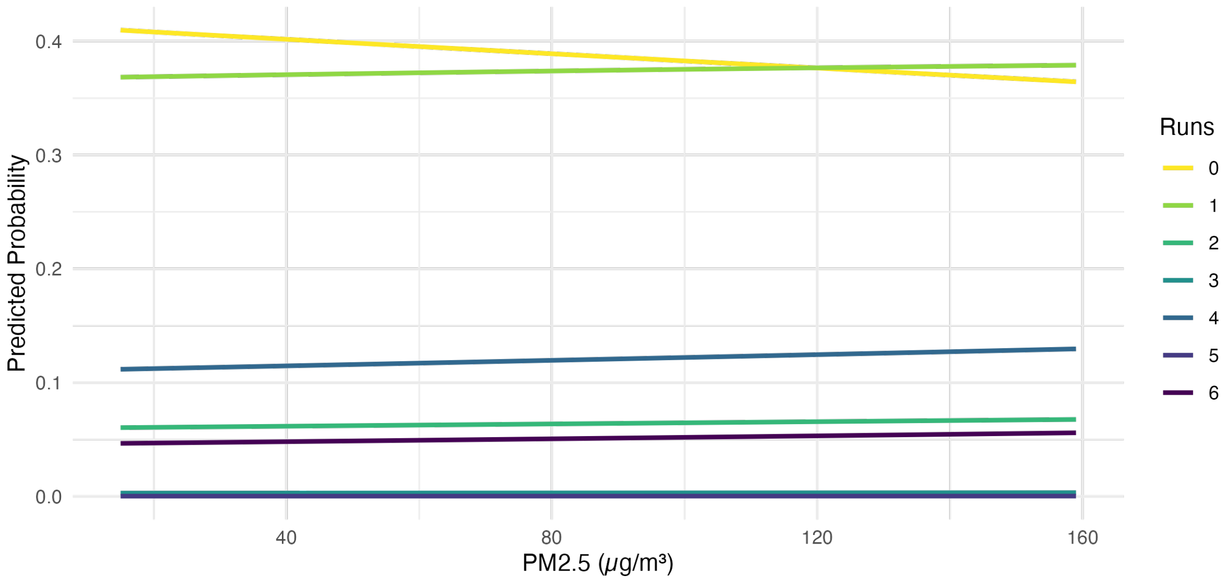


Figure A.3: Predicted Probability of Scoring 0, 1, 2, 3, 4, 5, or 6 runs by PM2.5

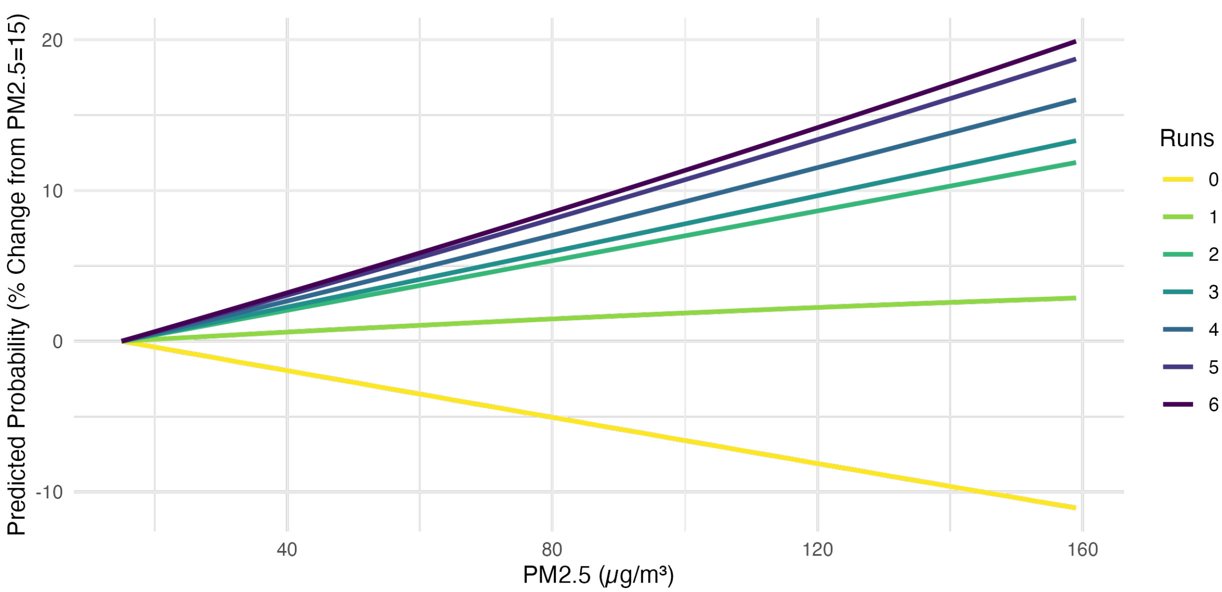


Figure A.4: Change in Predicted Probability of Scoring 0, 1, 2, 3, 4, 5, or 6 runs by PM2.5

Notes. These figures report graphically the results from Equation 15. Figure A.3 shows the predicted probability of scoring each number of runs given varying levels of game PM2.5. Figure A.4 presents the same results, but scaled relative to the predicted probabilities when PM2.5 is 15 $\mu g\, m^{-3}$.