Series Name: WPS

Paper No.: 232

Issue Date: 10/6/23

More than particulate matters: Pollution and productivity in Indian call centers

Deepshikha Batheja, Sarojini Hirshleifer, and Jamie T. Mullins

More than particulate matters: Pollution and productivity in Indian call centers*

Deepshikha Batheja[†], Sarojini Hirshleifer[‡], Jamie T. Mullins[§]

September 28, 2023

Abstract

We measure the impact of air pollution on daily labor productivity in call centers in five Indian cities. A one standard deviation increase in $PM_{2.5}$, a pollutant that has been widely studied in the literature, decreases intensive margin productivity by 0.15σ. The equivalent impacts for carbon monoxide and ozone, however, are 0.14σ and $.09\sigma$, respectively. Furthermore, carbon monoxide is responsible for more than half of the total productivity lost from pollution in our sample. We consider the potential productivity impacts of a national policy in India that targets $PM_{2.5}$ alongside a counterfactual policy that targets a broader range of pollutants.

Keywords: environment, pollution, productivity, labor, personnel economics, India, development, health
JEL Codes: Q52, M54, I15, O15

*We are grateful to the CEOs and center heads of partner call centers, Anil Sinha, Rajeev Jha and Sujit Sharma of Call-2-Connect and Kapil Sharma, Lakhan Joshi and Soham Ghosh of Five Splash for sharing their data and offering their support. We appreciate the help of the Ministry of Electronics and IT: Government of India, Software Technology Parks of India, and Bihar Industries Association for introducing and recommending this project to call centers.

[†]One Health Trust (email: deepshikha@onehealthtrust.org)
[‡]University of California, Riverside and CEGA. (email: sarojini.hirshleifer@ucr.edu)
[§]University of Massachusetts, Amherst (email: jmullins@umass.edu)

1 Introduction

Ambient air pollution has serious and well-documented impacts on human health. Even short-term pollution exposure has been linked to significantly reduced respiratory, cardiac and neurological function (Aguilar-Gomez et al., 2022). Thus, understanding the impact of pollution on economic outcomes such as productivity has become an important question for research and policy. Measuring the full effect of pollution on productivity is particularly relevant in developing countries such as India, where low productivity is likely a significant barrier to economic growth and high pollution levels are a major policy concern (Hsieh and Klenow, 2009; Damania et al., 2023).

This paper examines both the individual and combined effects of three components of air pollution: two key gaseous compounds, carbon monoxide (CO) and ozone (O_3), as well as particulate matter of less than 2.5 micrometers ($PM_{2.5}$). Thus, it more fully captures the impacts of air pollution on productivity, rather than relying on particulate matter as a proxy for all ambient air quality. Our analysis relies on daily pollution and productivity data from five cities across India, with the worker-level productivity data coming from two call center companies. This study has broad relevance as call centers are a major source of private sector employment in India, employing 4.47 million people (Ministry of Electronics and IT, 2021).

This setting is conducive to identifying the causal impacts of ambient air pollutants on productivity for a few reasons. First, call centers are not a major contributor to pollution, thus, we do not expect changing productivity in call centers to drive changes in pollution. Second, we are exploiting the variation in daily data. This data, combined with date and worker fixed effects, allows us to abstract away from the selection effects of any sectoral shifts driven by pollution. Third, it is unlikely that productivity and pollution are moving jointly due to some third factor. In particular, the call centers in our study handle calls to and from wide geographic areas, so changes in productivity are unlikely to be systematically driven by changes in the behavior of people on the other end of the phone lines from the call center employees. Finally, we find significant isolated variation in the levels of the three pollutants, allowing us to separately identify their effects.

Our first results examine the impact of the three pollutants on the extensive margin of productivity, which we define as being available for work. Thus, we include two extensive margin measures: being at work on a given day and net login time per day. The impact of pollution on the extensive margin is theoretically ambiguous. On the one hand, there could be a reduction in days worked or hours worked due to pollution-related health effects. On the other hand, there could be an increase in the time spent at work on more polluted days if workers value leisure more on days with better air quality. Given this ambiguity, it is not surprising that we do not find meaningful or significant effects on the extensive productivity margin.

We also examine the impact of the pollutants on the intensive margin of productivity and find substantial and significant results.[1] Our analysis focuses on an intensive margin productivity index which includes two measures of efficiency of calls handled. In our main results, which examine the linear average effect of each pollutant separately, we find that a one standard deviation increase in $PM_{2.5}$ induces a 0.145σ reduction in intensive margin productivity.[2] Notably, the impact of CO on pollution is similar at -0.142σ. In addition, the impact of O_3 is -0.086σ. These three results are each statistically significant at the 1% level, but the coefficients on the three pollutants are not statistically different from one another. For a one standard deviation increase in $PM_{2.5}$, CO and O_3, these effects translate to declines in productivity of 11.8%, 10.8% and 6.0% respectively.

We then measure the combined effects of the three pollutants as well as the relative importance of each pollutant in contributing to lower worker productivity. The combined effect of these three components of air pollution on productivity is -0.38σ. Furthermore, since the share of days with high CO levels is relatively larger than for the other pollutants in our sample, CO is responsible for 58% of the pollution-driven reductions in productivity. Thus, is the single most important pollutant we examine. In contrast, $PM_{2.5}$ and O_3 each contribute 21% of the productivity lost from pollution.

We extend our analysis to a stylized illustration of the potential impact on productivity of policies that only target $PM_{2.5}$ compared to those that target multiple pollutants. Specifically, we consider the potential impact on worker-level call center productivity of a recent policy, India's National Clean Air Programme (NCAP), that sets $PM_{2.5}$ reduction targets compared to a hypothetical policy that directly targets all three pollutants with a similar reduction target. Accounting for observed within-sample correlations in pollutants, we find a policy such as NCAP is likely to reduce the overall productivity loss from pollution by 12% relative to a setting without NCAP. In contrast, a hypothetical policy that targets all three pollutants would instead reduce the productivity loss from pollution by 24%. Whether policies controlling for all three pollutants are cost-effective, is a topic for future research.

This study is the first to examine the effect of multiple pollutants on labor productivity in an office setting or a developing country. The growing literature on pollution and labor productivity has focused on the impacts of single pollutants, and generally found negative effects.[3] Specifically,

[1]The lack of impact on the extensive margin allows the interpretation of the intensive margin effects to be relatively straightforward.

[2]We also consider flexible functional forms and interactions of these pollutants but find no evidence of non-linear or interactive effects.

[3]This work is also related to a broader literature on pollution and performance, which includes studies on the impact of pollution on academic (Lavy, Ebenstein and Roth, 2014; Zhang, Chen and Zhang, 2018) as well as athletic performance (Beavan et al., 2023; Cusick, Rowland and DeFelice, 2023; Mullins, 2018). Studies on academic/cognitive and athletic performance which have considered the impacts of multiple pollutants have generally found negative effects.

most papers have focused on the impact of $PM_{2.5}$ in physically demanding factory work (Adhvaryu, Kala and Nyshadham, 2022; He, Liu and Salvo, 2019). Other studies focus on the productivity impacts of O_3 in outdoor occupations (Graff Zivin and Neidell, 2012; Wang, Lin and Qiu, 2022). One study that, like this one, examines the productivity of office workers is Chang et al. (2019), which measures the impact of an Air Pollution Index (API) in call centers in China. The API index in their setting is defined primarily by PM_{10} levels, thus our study builds on that work by examining multiple pollutants and considering levels of fine particulates ($PM_{2.5}$). Although it is difficult to compare across study settings directly, we also find larger effects of particulate matter.

Few studies have examined the impact of multiple pollutants in a workplace setting, and those that have focused on physical or outdoor work. Chang et al. (2016) examine the impact of $PM_{2.5}$ and O_3 on pear-packers in the U.S. and only find effects for $PM_{2.5}$. Only Archsmith, Heyes and Saberian (2018), who study umpires' decisions in American baseball, also measures the impacts of the three pollutants examined in this study. They find impacts for $PM_{2.5}$ and CO, but not O_3. This study is therefore the first to document the impacts of CO and O_3 in an indoor workplace. It also uniquely considers the combined effects of multiple pollutants in a workplace setting that is representative of a broad section of the economy in the world's most populous country. India is an emerging economy in which pollution is high, and office-based service work is seen as critical to future economic development.

2 Context

2.1 Pollution and health

This study measures the impact of air pollution on labor productivity with the health impacts of those pollutants being the likely mediator. Fine particulate matter, $PM_{2.5}$, is known to penetrate indoors (Thatcher and Layton, 1995) and its harmful health effects are well-established (Brunekreef and Holgate, 2002). In particular, $PM_{2.5}$ can cause serious health issues by impairing cardiovascular and lung functioning (Liu et al., 2017; Pope III and Dockery, 2006), or cause daily allergies resulting in nose and throat irritation and mild headaches (Bernstein et al., 2008; Ghio, Kim and Devlin, 2000). Thus, $PM_{2.5}$ can potentially hamper an individual's productivity indirectly through changes in health, and may also be able to do so directly through reductions in cognitive performance (Sakhvidi et al., 2022; Ebenstein, Lavy and Roth, 2016; Ye et al., 2023).

Carbon monoxide is an odorless gas, which also has established negative health effects (Wang et al., 2019; Bell et al., 2009; Liu et al., 2018). The main sources of CO pollution include automotive fumes and industrial combustion emissions. When inhaled, it reduces oxygen flow within the body (Raphael et al., 1989). The immediate symptoms of inhaling CO include headache, dizziness, confusion, and disorientation, while in the longer term (2 to 28 days), it can lead to a rise in hypertension, lethal arrythmia, electrocardiographic changes and neuropsychiatric impairment (Raub et al., 2000). A U.S. based study documenting the pollution impacts of living in proximity to airports found that the CO from airplane fumes was associated with a rise in hospitalization rates and costs (Schlenker and Walker, 2016).

Exposure to O_3 in both the short- and long-term can also impact human health. In the short-term, it can cause breathing difficulties, including shortness of breath and pain when taking a deep breath, irritation to the eye and nose, coughing and sore or scratchy throat and inflammation of airways (Zhang, Wei and Fang, 2019). Long-term exposure to O_3 increases the risk of lung infections and aggravates lung diseases such as asthma, emphysema, and chronic bronchitis (McDonnell et al., 1999; Zhang, Wei and Fang, 2019).

Ambient, outdoor O_3 readily penetrates buildings (Salonen, Salthammer and Morawska, 2018; Ma et al., 2022), but indoor exposure to O_3 was previously considered less impactful because O_3 breaks down relatively quickly in indoor environments (Chang et al., 2016). We include O_3 levels in our analyses, however, since recent studies suggest that substantial shares, 25% to 60%, of daily ozone intake still occur indoors (Nazaroff and Weschler, 2022; Weschler, 2006) and workers are exposed to outdoor conditions during commuting.

2.2 Work in call centers

Call centers are part of the business process outsourcing (BPO) industry, which is a major source of employment in India (Ministry of Electronics and IT, 2021). The two call center companies in this study organize their work around individual contracts to provide voice support for companies that need such services. Each of these contracts leads to the establishment of a *process* or group of employees engaged in handling the same type of calls. Processes are typically categorized as inbound or outbound. Inbound processes receive calls for customer service, from industries such as food delivery, retail and telecommunication. In contrast, outbound processes require agents to make calls, to sell products or to conduct surveys. Our study includes ten processes (five inbound and five outbound), spread across five Indian cities (states): Noida (Uttar Pradesh), Mumbai (Maharashtra), Patna (Bihar), Hubli (Karnataka), and Udaipur (Rajasthan).

The processes rely on entry-level employees, known as agents, who handle the voice support. Agents are organized into teams of 20 to 25, each of which is managed by a team leader. Agents are expected to work six days a week with staggered schedules. Call center work is a common first job for young people, and there is significant churn in the workforce (Jensen, 2012).

3 Identification Strategy

Our identification strategy relies on the increasingly common approach in this literature of exploiting relatively high-frequency variation in pollution and productivity.[4] The high-frequency aspect of our data is most important in allowing us to isolate the impact of pollution on productivity from selection or composition effects in the workforce. In particular, daily variation in pollution allows us to abstract away from seasonal or long-term shifts in the composition of the workforce of the type that is the focus of another thread of research.[5]

Identification strategies that use high-frequency pollution measurement also rely on the assumption that there is no other causal relationship that primarily explains the co-movement of daily pollution and productivity levels. In particular, we do not expect that productivity would determine pollution levels, since the call center industry is not a significant contributor to pollution.[6] We also do not expect that pollution is indirectly affecting productivity through a third factor. Perhaps the most likely potential third factor is that pollution could reduce incoming calls to the call center if customers are affected by pollution. In this case, however, both inbound and outbound calls are targeting a wide geographic area in our sample. Thus, customers and employees are not in general exposed to the same pollution levels.

Finally, we note that this study is designed to identify the total impact of daily pollutant levels on productivity, that is, the joint effect of exposure to pollution during a worker's commute and the effect of indoor air pollution during the day. This effect is important in measuring the overall harms of pollution. Furthermore, any policy efforts to reduce ambient air pollution can address both these potential channels by addressing outdoor air pollution.[7]

[4] Adhvaryu, Kala and Nyshadham (2022) use hourly variation in pollution, Archsmith, Heyes and Saberian (2018) use three-hour variation for CO and a 12-hour/daily variation for O_3 and $PM_{2.5}$, Graff Zivin and Neidell (2012) and Chang et al. (2019) use daily variation in pollution. In the literature on athletic performance, Lichter, Pestel and Sommer (2017) use hourly variation in PM_{10} and O_3 and, Beavan et al. (2023) consider daily variation in PM_{10}, O_3 and NO_2 to study the performance of professional soccer players. Similarly, for academic performance related outcomes, La Nauze and Severnini (2021) use daily variations in PM2.5 and, Zhang, Chen and Zhang (2018) use daily variations in PM_{10}, SO_2 and NO_2 to study cognitive skills in adults.

[5] See for example, Khanna et al. (2021), Chen et al. (2013) and He, Xie and Zhang (2020).

[6] Furthermore, in industries that are major contributors to pollution, we typically expect to find that increased productivity would cause increased pollution, which is the opposite of what we find here.

[7] Of course, if harmful effects are driven by indoor pollution, employers can potentially mitigate directly with air purification. We will examine the role of indoor air pollution in future research.

3.1 Data and variable construction

All outcome measures in this study rely on data that is collected automatically by technology-based monitoring systems in the call centers.[8] We use data from 2,777 workers, for a total of 138,337 observations. Our extensive margin of productivity includes two approaches to measuring that employees are available for work: attendance and net login time. Attendance is simply an indicator of whether an employee comes to work on a given day. Net login time captures the amount of time that an employee is logged into the computer system and available to work. Thus, this measure captures the time spent actually working and excludes the time spent on breaks.[9] For the intensive margin of productivity, we consider two measures of the intensity of time spent conditional on being at work: calls per shift and calls per hour. These are distinct measures that account for the intensive margin in two different ways. Calls per shift is the number of calls made in a day, irrespective of the time at work, while calls per hour is based on calls conditional on total time at work. It is important to measure both as they allow us to understand the nature of intensive margin effects more precisely in the event that we also observe extensive margin effects. We index both the intensive and extensive margin measures in order to have a single, standardized outcome measure for each.

For air pollution data in the main analysis, we rely on measures maintained by the World Air Quality Index Project WAQI (2021).[10] The data itself comes from monitors managed by the Central Pollution Control Board of India (CPCB). Specifically, we use daily data from the closest monitor to each of the five call-center offices in the productivity data.

We report our results using two types of units for the pollutants in our sample. First, we benchmark the pollutants using within-sample standard deviations of those pollutants. Since the reference point for pollution should be zero (rather than the potentially harmful mean level of pollution in a given environment), we simply divide the pollutants by their standard deviations. We also report our results using concentrations, which are internationally comparable and objective measures.

[8] Our data begins in 2018 and ends before the COVID-19 pandemic begins in early 2020. For more details on this data, see Section SA1.2.

[9] Hence, it is zero if the employee does not come to work. We note that Chang et al. (2019) have two measures of time spent at work, all time spent at work and net login time. They assign the first of these to the extensive margin and the second to the intensive margin. In our setting, it is more natural to categorize time spent at work as an extensive margin measure and intensity of work measures, conditional on time spent, as informative regarding the intensive margin.

[10] See Section SA1.1 for further details on the pollution and weather data. Our results are robust to alternative pollution measures and sources as indicated in Section SA2.

3.2 Distributions of observed pollutants

In order to better understand the implications of our results, we first examine the distribution of the three pollutants in our sample.[11] Pollution levels are exceptionally high compared to the global average. For example, the mean concentration of $PM_{2.5}$ across all of the city-worker-day observations in our sample is 66.54 $\mu g/m^3$ (s.d. 68.28 $\mu g/m^3$). These concentrations of $PM_{2.5}$ are at least 10 times the WHO guideline for safe exposure, which is 5 $\mu g/m^3$ average per year (WHO, 2021). $PM_{2.5}$ levels in our sample are also far above the global average for cities, which was 27 $\mu g/m^3$ in 2018 (WHO, 2022). They are fairly representative of the yearly average for all cities in India, however, which had a mean of 61.41$\mu g/m^3$ in 2018.[12] In addition, such concentrations are also highly relevant for hundreds of millions of people living in cities across Sub-Saharan Africa and South Asia, in particular. For example, in 2018, the average $PM_{2.5}$ level in cities in Nigeria was 55.13 $\mu g/m^3$, while for cities in Pakistan, it was 59.51 $\mu g/m^3$ (WHO, 2022).

Unlike $PM_{2.5}$, CO and O_3 concentration averages in our sample do not exceed the WHO recommended guidelines.[13] Specifically, the mean CO concentration observed in our sample is 1.05 mg/m^3 (s.d. 0.62 mg/m^3), and the 99th percentile level is 3.04 mg/m^3. In contrast, the WHO daily average guideline for the 99th percentile of daily averages is 4 mg/m^3.[14] The average O_3 level in our sample is 51.57 $\mu g/m^3$ (s.d. 30.77 $\mu g/m^3$), which does not exceed the WHO guideline of an average of 60 $\mu g/m^3$ during peak season.[15] The 90th percentile date-worker observation in our sample, however, is 107.8 $\mu g/m^3$, which does exceed the WHO day-level maximum guideline of 100 $\mu g/m^3$. Thus, for at least 10% of the worker-date observations in our sample, O_3 is above the recommended levels. In general, however, we will be assessing the impacts of CO and O_3 at levels below those at which the WHO has thus far recognized as important for public health (WHO, 2021).

Our identification strategy relies on short-term variation in pollution measures. Thus, we examine the extent of the granular variation in our sample both within and across individual pollutants. One concern that may arise is whether these three pollutants are highly correlated in the short term, and thus move in lock-step. This would make it challenging to separately identify the effects of these three pollutants. We find evidence, however, of rich variation across the levels of the different pollutants even within a single day. In addition, we examine variation within each pollutant from

[11] See Table SA1 for summary statistics and Figure SA1 for histograms of the three pollutants.

[12] For more on the representativeness of our data to India, see Section SA1.3.

[13] Despite issuing these guidelines, WHO does not track CO and O_3, thus it is more difficult to benchmark these measures in our sample to global averages.

[14] Note that CO measures are in milligrams per cubic meter rather than $\mu g/m^3$.

[15] The WHO defines the peak season as the six months of the year during which O_3 is at its highest levels.

one day to the next, and find it to be meaningful as well.[16]

3.3 Estimation

Our estimation strategy relies on measuring the impact of daily variation in city-specific pollution on worker-level outcomes. Thus, our main estimating equation,

$$Outcome_{it} = \beta_0 + \beta_1 \sum_k \beta_k \; Pollutant_{ct} + X_t + \alpha_i + \delta_t + \epsilon_{it} \tag{1}$$

includes controls as well as both worker-level and date fixed effects.[17] The controls include atmospheric variables that can influence pollution, including temperature, precipitation, dew point and cloud cover. The worker fixed effects in our model account for any differences in workers that could be correlated with different pollution levels. Furthermore, the date fixed effects account for differences in productivity across days that could also be correlated with pollution. For example, if weekends have lower pollution, but workers tend to be less productive on weekends, the date fixed effects would account for such patterns. Since these are date rather than day-of-the-week fixed effects, they also account for a similar pattern for any holidays in our sample. Of course, that particular pattern would actually work against the identification of significant effects. We also account for autocorrelation in the error term over time (within worker) and on a given day across workers using two-way clustering by worker and date.[18]

Our outcome measures include standardized indices for the intensive and extensive margin as well as their unadjusted components. Although three of the four component measures of productivity take on only positive values (net login time, calls per hour, and calls per shift), taking logs of these measures does not generate the preferred specification in this setting. First, the residuals of the regressions of the unadjusted measures are largely normally distributed.[19] It is the distribution of the residuals that is relevant to implementing a linear model, as opposed to the distribution of the

[16] See histograms of within-day pairwise variation across the individual pollutants (Figure SA4) and within-pollutant variation across days (Figure SA5).

[17] Two-way fixed effect models have come under significant criticism in recent years in the context of difference-in-difference framework (Roth et al., 2022). We do not think our setting, however, directly maps into those frameworks, since we rely on different identifying assumptions, and thus do not rely on the parallel trends assumption. Furthermore, this literature does not in general allow for a continuous treatment measure. An exception is Callaway, Goodman-Bacon and Sant'Anna (2021), but that model does assume a pre-period in which to test parallel trends, and that set-up is not relevant to our context.

[18] Our approach to clustering aligns with Chang et al. (2019), who have similarly structured data.

[19] See Figure SA6. Although there are a few outliers, the mass of the distribution is centered. Thus, this is more appropriately addressed by winsorizing rather than taking logs. We confirm our results are robust to winsorizing in Section SA2.

dependent variable. Second, we cannot reject linearity of the relationship between pollution and productivity.[20] Thus, it is preferred to analyze the unadjusted outcome in this case as opposed to a non-linear transformation of that outcome which imposes additional functional form assumptions.

4 Results

4.1 Mean impact of the pollutants

First, we examine the impact of pollution on the extensive margin index and rule out meaningful impacts (Table 1, Panel A). The measured coefficients for each of the three pollutants are close to zero, not statistically significant, and relatively precisely estimated. Specifically, focusing on the estimates in which the pollutants are measured in standard deviations, the coefficient on $PM_{2.5}$ is approximately 0.039σ with a standard error of 0.028σ. Meanwhile, the coefficients on CO and O_3 are only 0.004σ (s.e. 0.021) and 0.003σ (s.e. 0.017), respectively. The pattern of these findings also holds for both the sub-components of the extensive margin index: whether someone comes to work at all on a given day, and daily net login time. We thus do not find evidence that pollution affects the extensive margin of productivity in our context.

Next, we examine the impact of the three pollutants on the intensive margin of productivity and find meaningful and statistically significant reductions in productivity across all three pollutants (Table 1, Panel A). We first measure these impacts in terms of standard deviations of observed pollution concentrations in our sample. $PM_{2.5}$ and CO have similar average impacts on intensive margin productivity. A one standard deviation increase in $PM_{2.5}$ or CO reduces the intensive margin productivity index by 0.145σ or 0.142σ, respectively. O_3 has a smaller average impact of -0.086σ, but that estimate is not statistically significantly different from those for $PM_{2.5}$ and O_3. All three coefficients are significant at the 1% level, and the estimates are similar across the two sub-components of the index.

We also report these results using concentrations to understand the objective impact of the pollutants along an internationally comparable measure (Table 1, Panel B). The impact of an increase of a single $\mu g/m^3$ of $PM_{2.5}$ on the intensive margin of productivity is -0.00213σ, while the effects of a 1 mg/m^3 increase in CO and a 1 $\mu g/m^3$ increase in O_3 are -0.23014σ and -0.00281σ respectively. Although the effects of the three pollutants have similar magnitudes when they are measured in standard deviations, when measured in concentrations, the results for CO are two or-

[20] We present evidence for linearity in two ways. We plot a linear predictor along with semi-parametric distributional impacts in Section 4.3. We also plot the predictors against the residuals. That the residuals are centered around a flat line is consistent with the assumption of linearity (see Figure SA7).

ders of magnitude larger. This is because ambient concentrations of CO are measured in different units to account for the typically higher observed concentrations of CO.

See Section SA2 for robustness checks, including winsorizing, alternative pollution measures, and lagged pollution effects.

4.2 Main results in context

We examine an unadjusted component measure of the intensive margin productivity index, in order to simplify comparisons with the literature. Focusing on calls per shift, a one standard deviation increase in $PM_{2.5}$ decreases calls per shift by 12.95 calls or 11.8%.[21] In addition, a one standard deviation increase in CO or O_3 reduces calls per shift by 10.8% or 6.0% respectively. The paper most closely related to this one is Chang et al. (2019), and in fact, they use calls per shift as their main outcome. Still, it is difficult to directly compare our results to theirs, since they use an API-based measure of PM_{10} as their sole measure of pollution. They find that a 10-point increase in that API, which is 24.1% of a standard deviation in their setting, decreases calls per shift by 0.35%. Thus, a one standard deviation increase in pollution in their setting should reduce calls per shift by 1.45%. That said, the standard deviation of API in their setting is a smaller percentage of its mean compared to the most closely analogous measure of pollution in our setting, $PM_{2.5}$. Thus, these standard deviations are not directly comparable.

Nonetheless, it is likely that we observe larger effects in this study, even when only considering the effect of $PM_{2.5}$, rather than the combined effect of all three pollutants. This is perhaps not surprising since in contrast to Chang et al. (2019), we are able to measure the relevant form of particulate matter more directly through a concentration-based measure of $PM_{2.5}$. As they indicate, $PM_{2.5}$ is the form of particulate matter that is most likely to affect productivity in these types of workplaces, and thus in the context of this research question, PM_{10} is a weakened proxy measure for it. Furthermore, API is not an immutable, scientific measure such as concentration, but a subjective assessment of harm from a given concentration based on existing research (See Section SA1.1). In addition, it is possible that either workers or managers are less able to adapt to pollution in our setting relative to Chang et al. (2019)'s call center setting (Adhvaryu, Kala and Nyshadham, 2022).

Finally, we note that we find that CO has a meaningful impact on productivity despite the fact that CO levels in our setting are generally below the recommended WHO guidelines. These results, however, are generally aligned with other recent work on the effects of CO pollution in the U.S.,

[21] It is not possible to calculate these percentages for the index since the mean is zero. Since the relative magnitude of the results for calls per shift and calls per hour are similar, this analysis could be done for either.

although such studies benchmark to EPA rather than WHO guidelines. In particular, Schlenker and Walker (2016) and Archsmith, Heyes and Saberian (2018) identify harmful effects of ambient CO on overall health and the productivity of umpires, respectively, at levels well below relevant guidelines.

4.2.1 Validity of the intensive margin results

A potentially important consideration in examining the intensive margin effects in this setting is whether there may be selection into the intensive margin sample. Specifically, if employees who show up on high pollution days are either more or less productive than the employees who show up on low pollution days, that could have important implications in interpreting the results. This type of selection does not appear to be a significant issue in this setting. In particular, there are no impacts on the extensive margin here, which means that the same percentage of employees show up and spend the same amount of time working regardless of pollution levels. Thus, under the common assumption of monotonicity of selection, this would be sufficient to determine that we do not observe selection on this margin.[22]

4.3 Distributional impact of pollutants

To better understand the role of pollution in reducing productivity, we examine the distributional impact of the three pollutants on productivity. First, we examine impacts from binned measures of the pollutants denominated in within-sample standard deviations, relative to a reference bin, which includes pollution levels that are less than or equal to one standard deviation above *zero pollution levels*. This exercise allows us to make an initial assessment about whether the effects are concentrated in one part of the distribution and whether there are differences in the impacts across the three pollutants. We confirm that even when examining the entire distribution, we do not find any meaningful or consistent effects of pollution on the extensive margin (Figure 1, Panel A). Along the intensive margin, we find what appear to be increasing effects across the distribution for all three pollutants (Figure 1, Panel B). In addition, we do not find evidence that the impact of the three pollutants on productivity differs at any point across the distribution.[23]

We also examine the distributional effects of the concentration-based measures of pollution on the

[22] For example, the bounding approach proposed by Lee (2009), a widely-used selection correction, relies on this assumption. Since there are no meaningful differences in the extensive margin across the distribution of pollution here, the upper and lower Lee bounds would be approximately equivalent to each other and the treatment estimate of interest. In addition, our treatment variable is continuous, so it would not be straightforward to implement such bounds.

[23] Although the standard errors in the graph overlap, it does not necessarily imply that we cannot reject that coefficients are the same. Instead, we conduct pairwise F-tests across the coefficients on the pollutants within each bin and do not reject the null hypothesis that these coefficients are the same (not reported).

intensive margin productivity index in order to further comparisons across settings and examine the linearity of response curves (Figure 2). Here, the first non-omitted bin for each pollutant includes the median and the remaining bins are designed to particularly examine the upper half of each distribution. We do not reject that the effects are linear for any of the three pollutants, even though the highest bin for each pollutant was chosen to isolate relatively extreme days and thus is likely to identify non-linearities.[24] If the effects are in fact linear, the bin in which the effects become significant is simply driven by statistical power.

4.4 Overall impacts and policy application

Next, we measure the combined effects of the three pollutants in our sample. The overall average reduction in productivity attributable to pollution exposure is -0.38σ (Figure 3). This is estimated by relying on the combinations of pollutants and their frequency as observed in our data. This total productivity loss due to pollution is largely driven by carbon monoxide at 58% of the total, with the remaining damages allocated equally to $PM_{2.5}$ and O_3 (21% each). Although the magnitude of the effects on productivity from $PM_{2.5}$ and CO are similar across their respective distributions, days with relatively high levels of CO are more common than days with relatively high levels of $PM_{2.5}$ in our data.[25] This suggests that policies that aim to reduce CO directly could have meaningful impacts.

Thus, we consider the implications of our results for policies and regulations designed to reduce the harm from air pollution. In particular, some major policy initiatives focus explicitly on $PM_{2.5}$, while our findings indicate the importance of CO and O_3 in reducing productivity. An example of such an initiative that is highly relevant to our context is India's National Clean Air Programme (NCAP). Launched in 2019, it identified more than 100 cities in India as being in "non-attainment," and forced each to undertake a series of actions to quantify and ameliorate local air quality, and specifically reduce $PM_{2.5}$ levels by 20-30% by 2024 relative to 2017 levels (Ganguly, Selvaraj and Guttikunda, 2020).

We consider the implications of NCAP for reducing the impact of pollution on productivity relative to a similar policy that aims to reduce all three pollutants by 20% (Figure 3).[26] We conduct this simple exercise by estimating a counterfactual level of productivity lost to pollution in a setting

[24]There is less than 8% of the sample in that bin for each of the pollutants.

[25]See Section SA1.3 for a discussion of the representativeness of our data.

[26]All pollutant reductions are simulated as equal percentage reductions for every day in the sample. Given that our estimates suggest linearity, alternative approaches are unlikely to substantially change the main spirit of our results, and reduction approaches that focused specifically on high-pollution days would lead to even larger advantages for programs that targeted CO and O_3 in addition to $PM_{2.5}$, as these pollutants have more relatively high-pollution days for which percentage reductions would lead to larger reductions in absolute levels.

in which the NCAP had been implemented before our study period began, and $PM_{2.5}$ was 20% lower each day of our sample. As a benchmark, we consider the counterfactual setting in which reductions in pollution driven by the NCAP only affect $PM_{2.5}$, and the policy has no spillover effects. In that estimate, the negative impact of pollution on productivity is modestly reduced to -0.36σ from a baseline level -0.38σ. Of course, reducing $PM_{2.5}$ is likely to lead to some spillover reductions in CO and O_3. Thus, we next estimate those potential spillovers using the correlation in the three pollutants observed within our data.[27] After accounting for potential beneficial spillovers to CO and O_3 using this approach, we estimate that NCAP could reduce the damages to production by a total of -0.34σ, a 12% reduction relative to the baseline without NCAP.

If NCAP instead required that $PM_{2.5}$, CO, and O_3 were all reduced by 20%, then the impact of pollution on productivity would be reduced to -0.29σ, a 24% reduction relative to the baseline. Whether this additional reduction would be cost-effective is beyond the scope of this paper. That calculation would depend on any additional abatement or enforcement costs from adding these two pollutants to the policy compared to the returns to increasing productivity estimated here as well as any additional health benefits that are specific to reductions in CO and O_3.

5 Conclusion

This paper contributes to the emerging and important literature on the impact of pollution on productivity. We exploit daily variation in pollution and productivity across five Indian cities to estimate the individual and combined effects of $PM_{2.5}$, CO and O_3 on both the extensive and intensive margins of productivity. We find no effects on the extensive margin and substantial effects on the intensive margin of productivity. Specifically, a one standard deviation increase in $PM_{2.5}$, CO and O_3 leads to significant reductions in intensive margin productivity of 0.145σ, 0.142σ and 0.086σ, respectively. Our analysis of the combined effects of the pollutants finds that CO is responsible for 58%, and $PM_{2.5}$ and O_3 each for 21%, of the pollution-driven reductions in productivity in our sample. Since the call center industry is one of the largest private employers in India and contributes significantly to the national income, these productivity effects are likely to be highly relevant to the overall economy and the well-being of workers.

Our findings suggest that the overall impact of pollution on productivity is substantially larger than previously shown in the literature and is driven by multiple pollutants. These results have broad policy implications for the design and implementation of environmental protection laws, particularly in developing countries. Existing environmental policies and programs may be focus-

[27] Of course these correlations across pollutants are not known to be determined by changes in pollution policy, so it is not entirely clear if they would be replicated under NCAP.

ing disproportionately on particulate matter, and underweighting or ignoring the damaging effects of other pollutants. This study indicates that strengthening local regulations on multi-pollutants, particularly carbon monoxide, should be carefully considered.

References

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2022. "Management and shocks to worker productivity." *Journal of Political Economy*, 130(1): 1–47.

Aguilar-Gomez, Sandra, Holt Dwyer, Joshua Graff Zivin, and Matthew Neidell. 2022. "This is air: The "nonhealth" effects of air pollution." *Annual Review of Resource Economics*, 14: 403–425.

Archsmith, James, Anthony Heyes, and Soodeh 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(4): 827–863.

Batheja, Deepshikha. 2019. "Gender peer effects in the workplace: A field experiment in Indian call centers." *Working Paper*.

Beavan, Adam, Sascha Härtel, Jan Spielmann, and Michael Koehle. 2023. "Air pollution, a worthy opponent? How pollution levels impair athlete performance across physical, technical, and cognitive domains." *Science of The Total Environment*, 900: 165707.

Bell, Michelle L, Roger D Peng, Francesca Dominici, and Jonathan M Samet. 2009. "Emergency hospital admissions for cardiovascular diseases and ambient levels of carbon monoxide: results for 126 United States urban counties, 1999–2005." *Circulation*, 120(11): 949–955.

Bernstein, Jonathan A, Neil Alexis, Hyacinth Bacchus, I Leonard Bernstein, Pat Fritz, Elliot Horner, Ning Li, Stephany Mason, Andre Nel, John Oullette, et al. 2008. "The health effects of nonindustrial indoor air pollution." *Journal of Allergy and Clinical Immunology*, 121(3): 585–591.

Brunekreef, Bert, and Stephen T Holgate. 2002. "Air pollution and health." *The Lancet*, 360(9341): 1233–1242.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with a continuous treatment." *arXiv preprint arXiv:2107.02637*.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. "Particulate pollution and the productivity of pear packers." *American Economic Journal: Economic Policy*, 8(3): 141–69.

Chang, Tom Y, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2019. "The effect of pollution on worker productivity: evidence from call center workers in China." *American Economic Journal: Applied Economics*, 11(1): 151–72.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy." *Proceedings of the National Academy of Sciences*, 110(32): 12936–12941.

CPCB, India. n.d.. "Central Control Room for Air Quality Management - All India." Dataset accessed February 2023. https://airquality.cpcb.gov.in/ccr//caaqm-dashboard-all/caaqm-landing.

Cusick, Marika, Sebastian T Rowland, and Nicholas DeFelice. 2023. "Impact of air pollution on running performance." *Scientific Reports*, 13(1): 1832.

Damania, Richard, Esteban Balseca, Charlotte De Fontaubert, Joshua Gill, Jun Rentschler, Jason Russ, and Esha Zaveri. 2023. *Detox Development: Repurposing Environmentally Harmful Subsidies.* Washington, D.C.:World Bank.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution." *American Economic Journal: Applied Economics*, 8(4): 36–65.

EPA, US. 2018. "Technical assistance document for the reporting of daily air quality–The Air Quality Index (AQI)." *Environmental Protection Agency*.

Ganguly, Tanushree, Kurinji L Selvaraj, and Sarath K Guttikunda. 2020. "National Clean Air Programme (NCAP) for Indian cities: review and outlook of clean air action plans." *Atmospheric Environment: X*, 8: 100096.

Ghio, Andrew J, Chong Kim, and Robert B Devlin. 2000. "Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers." *American Journal of Respiratory and Critical Care Medicine*, 162(3): 981–988.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The impact of pollution on worker productivity." *American Economic Review*, 102(7): 3652–73.

He, Guojun, Yang Xie, and Bing Zhang. 2020. "Expressways, GDP, and the environment: The case of China." *Journal of Development Economics*, 145: 102485.

He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. "Severe air pollution and labor productivity: Evidence from industrial towns in China." *American Economic Journal: Applied Economics*, 11(1): 173–201.

Hersbach, H., Bell B. Berrisford P. Biavati G. Horányi A. Muñoz Sabater J. Nicolas J. Peubey C. Radu R. Rozum I. Schepers D. Simmons A. Soci C. Dee D. Thépaut J-N. 2023. "ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)."

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Jensen, Robert. 2012. "Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India." *The Quarterly Journal of Economics*, 127(2): 753–792.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. 2021. "The productivity consequences of pollution-induced migration in China." National Bureau of Economic Research Working Paper 28401.

La Nauze, Andrea, and Edson R Severnini. 2021. "Air pollution and adult cognition: evidence from brain training." National Bureau of Economic Research Working Paper 28785.

Lavy, Victor, Avraham Ebenstein, and Sefi Roth. 2014. "The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation." National Bureau of Economic Research Working Paper 20648.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071–1102.

Lichter, Andreas, Nico Pestel, and Eric Sommer. 2017. "Productivity effects of air pollution: Evidence from professional soccer." *Labour Economics*, 48: 54–66.

Liu, Cong, Peng Yin, Renjie Chen, Xia Meng, Lijun Wang, Yue Niu, Zhijing Lin, Yunning Liu, Jiangmei Liu, Jinlei Qi, et al. 2018. "Ambient carbon monoxide and cardiovascular mortality: a nationwide time-series analysis in 272 cities in China." *The Lancet Planetary Health*, 2(1): e12–e18.

Liu, Qian, Cheng Xu, Guixiang Ji, Hui Liu, Wentao Shao, Chunlan Zhang, Aihua Gu, and Peng Zhao. 2017. "Effect of exposure to ambient PM2. 5 pollution on the risk of respiratory tract diseases: a meta-analysis of cohort studies." *Journal of Biomedical Research*, 31(2): 130.

Ma, Nan, Max Hakkarainen, Miaomiao Hou, Dorit Aviv, and William W Braham. 2022. "Impacts of building envelope design on indoor ozone exposures and health risks in urban environments." *Indoor and Built Environment*, 1420326X221112010.

McDonnell, William F, David E Abbey, Naomi Nishino, and Michael D Lebowitz. 1999. "Long-term ambient ozone concentration and the incidence of asthma in nonsmoking adults: the AHSMOG Study." *Environmental Research*, 80(2): 110–121.

Ministry of Electronics and IT. 2021. "*Government of India*." Dataset accessed February 2023. https://www.meity.gov.in/content/employment.

Mullins, Jamie T. 2018. "Ambient air pollution and human performance: Contemporaneous and acclimatization effects of ozone exposure on athletic performance." *Health Economics*, 27(8): 1189–1200.

Muñoz Sabater, J. 2019. "ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)."

Nazaroff, William W, and Charles J Weschler. 2022. "Indoor ozone: Concentrations and influencing factors." *Indoor Air*, 32(1): e12942.

Pope III, C Arden, and Douglas W Dockery. 2006. "Health effects of fine particulate air pollution: lines that connect." *Journal of the Air and waste management association*, 56(6): 709–742.

Raphael, Jean-Claude, David Elkharrat, Marie-Claude Jars-Guincestre, Claude Chastang, Jean-Baptiste Vercken, Valérie Chasles, and Philippe Gajdos. 1989. "Trial of normobaric and hyperbaric oxygen for acute carbon monoxide intoxication." *The Lancet*, 334(8660): 414–419.

Raub, James A, Monique Mathieu-Nolf, Neil B Hampson, and Stephen R Thom. 2000. "Carbon monoxide poisoning—a public health perspective." *Toxicology*, 145(1): 1–14.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *arXiv preprint arXiv:2201.01194*.

Sakhvidi, Mohammad Javad Zare, Jun Yang, Emeline Lequy, Jie Chen, Kees de Hoogh, Noémie Letellier, Marion Mortamais, Anna Ozguler, Danielle Vienneau, Marie Zins, et al. 2022. "Outdoor air pollution exposure and cognitive performance: findings from the enrolment phase of the CONSTANCES cohort." *The Lancet Planetary Health*, 6(3): e219–e229.

Salonen, Heidi, Tunga Salthammer, and Lidia Morawska. 2018. "Human exposure to ozone in school and office indoor environments." *Environment International*, 119: 503–514.

Schlenker, Wolfram, and W Reed Walker. 2016. "Airports, air pollution, and contemporaneous health." *The Review of Economic Studies*, 83(2): 768–809.

Thatcher, Tracy L, and David W Layton. 1995. "Deposition, resuspension, and penetration of particles within a residence." *Atmospheric Environment*, 29(13): 1487–1497.

Wang, Chunchao, Qianqian Lin, and Yun Qiu. 2022. "Productivity loss amid invisible pollution." *Journal of Environmental Economics and Management*, 112: 102638.

Wang, Yu, Chengye Yao, Chengzhong Xu, Xinying Zeng, Maigeng Zhou, Yun Lin, Pei Zhang, and Peng Yin. 2019. "Carbon monoxide and risk of outpatient visits due to cause-specific diseases: a time-series study in Yichang, China." *Environmental Health*, 18(1): 1–11.

WAQI. 2021. "*Air Quality Historical Data*." Dataset accessed January 2023. https://aqicn.org/data-platform/register/.

Weschler, Charles J. 2006. "Ozone's impact on public health: contributions from indoor exposures to ozone and products of ozone-initiated chemistry." *Environmental Health Perspectives*, 114(10): 1489–1496.

WHO. 2021. *WHO global air quality guidelines: particulate matter (PM2. 5 and PM10), ozone,* *nitrogen dioxide, sulfur dioxide and carbon monoxide.* World Health Organization.

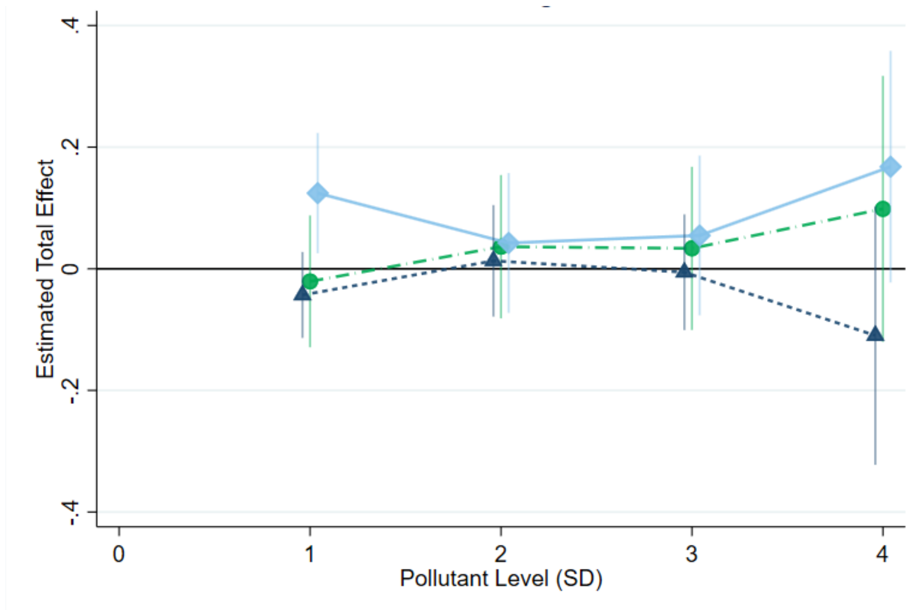
WHO. 2022. "Global Health Observatory." Dataset accessed August 2023 https://www.who.int/data/gho/data/indicators/indicator-details/GHO/concentrations-of-fine-particulate-matter-(pm2-5).

Ye, Mengwen, Jiacheng Yang, Jiaxin Li, Yifan Wang, Weiqiang Chen, Li Zhu, TianYao Wang, Jun Liu, Daoying Geng, and Zekuan Yu. 2023. "Progress in Mechanisms, Pathways and Cohort Studies About the Effects of PM2. 5 Exposure on the Central Nervous System." *Reviews of Environmental Contamination and Toxicology*, 261(1): 7.

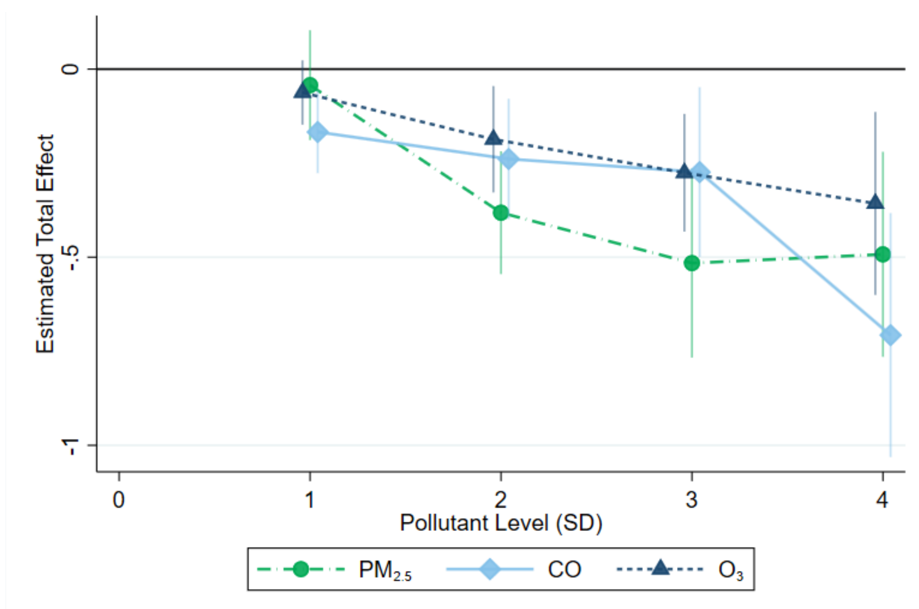
Zhang, Junfeng, Yongjie Wei, and Zhangfu Fang. 2019. "Ozone pollution: a major health hazard worldwide." *Frontiers in Immunology*, 10: 2518.

Zhang, Xin, Xi Chen, and Xiaobo Zhang. 2018. "The impact of exposure to air pollution on cognitive performance." *Proceedings of the National Academy of Sciences*, 115(37): 9193–9197.

Figure 1: Semi-parametric effects in standard deviations



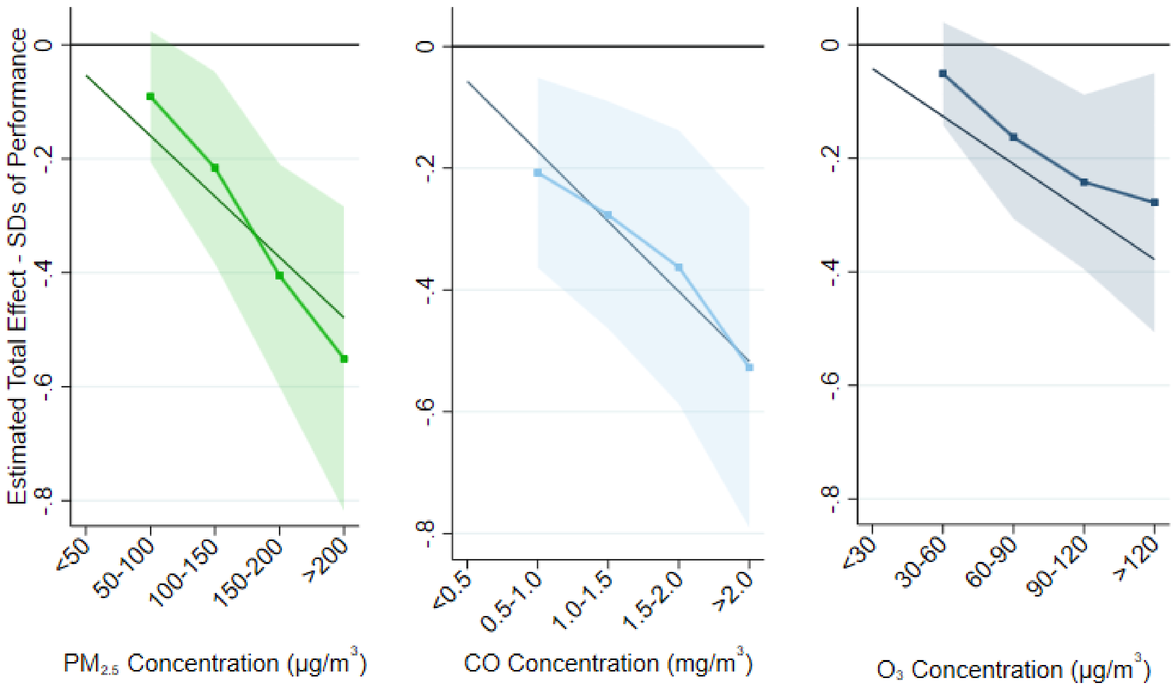
(a) Panel A: Extensive margin



(b) Panel B: Intensive margin

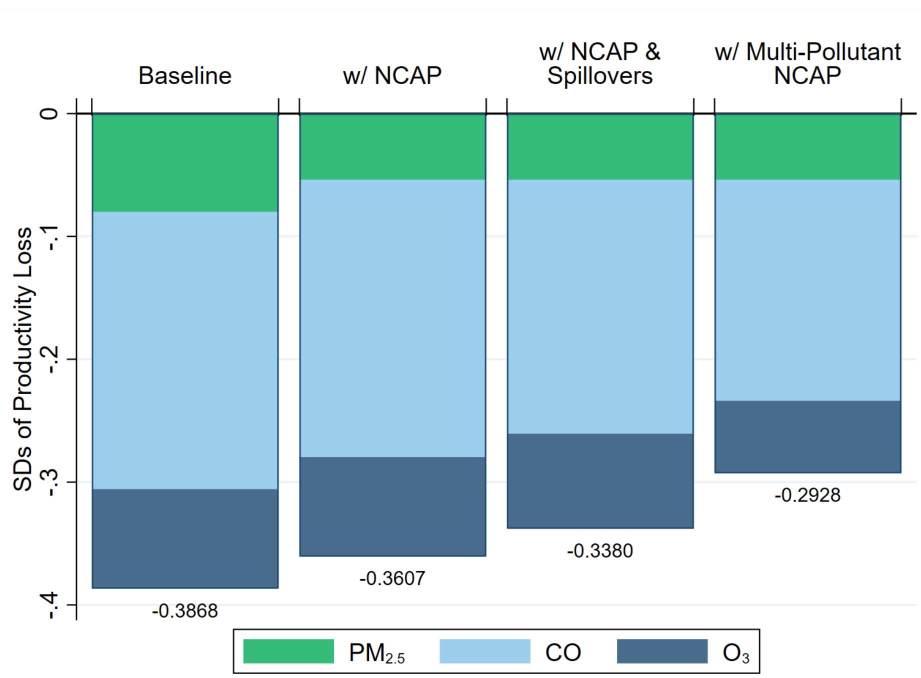
Notes: Estimates in each panel are from a single, separate regression (see Table SA7). Units of the outcome and pollution variables are standard deviations in the extensive margin sample. For the extensive margin index n=138,337. For the intensive margin index n=94,679. Dots indicate point estimates of the coefficient on the indicator for the daily average pollutant level falling in the indicated bin. Each of the bins has a width of one standard deviation for that pollutant. All estimates are based on specifications which also include quadratic controls for temperature, precipitation, humidity, and cloud cover, as well as worker and date fixed effects. Whiskers indicate 95% confidence intervals based on robust standard errors two-way clustered by worker and date.

Figure 2: Semi-parametric effects in concentrations on intensive margin index



Notes: Estimates in three plots are from a single regression, n=94,679 (see Table SA8). The outcome measure is the intensive margin productivity index. Dots indicate point estimates of indicator for the pollutant measure falling in the indicated bin. The lowest bin is omitted for each pollutant, so all estimates are relative to the lowest concentration category for the pollutant. Estimates are based on the main model which also includes quadratic controls for temperature, precipitation, humidity, and cloud cover, as well as worker and date fixed effects. Shading indicates 95% confidence intervals based on robust standard errors two-way clustered by worker and date. The additional line on each panel has the slope of the linear estimated effect from the main specification reported in Column 4 Panel B of Table 1 and passes through the point (0,0).

Figure 3: Damages by regulatory approach



Notes: Damage estimates are averages based on pollution coefficients from our main binned specification using the intensive margin productivity index as the outcome variable as reported in Figure 2. The baseline case represents the observed exposure levels in our data. "w/ NCAP" reduces observed exposure levels of $PM_{2.5}$ by 20% for every observation in our sample. "w/ NCAP & Spillovers" represents exposures equivalent to an across-the-board 20% reduction in $PM_{2.5}$, and reductions in CO and O_3 equal to the expected reductions in these pollutants given the 20% reduction in $PM_{2.5}$ based on the correlations of pollutant levels in our sample: 0.3612 for CO and 0.1478 for O_3. The "w/ Multi-Pollutant NCAP" is a scenario in which each observed level of all three pollutants is assumed to have been reduced by 20% relative to the observed level in our data. Damages are not reduced proportionally to the reduction in exposure levels because the binned specification allows for non-linear relationships between exposure levels and damages.

Table 1: Average effects of pollutants on productivity

	Extensive margin			Intensive margin		
	(1)	(2)	(3)	(4)	(5)	(6)
	EM index	At work	Net login time	IM index	Calls per shift	Calls per hour
Panel A: Pollutants in standard deviations						
$PM_{2.5}$ (SD)	0.039	0.019	8.559	-0.145	-12.950	-1.268
	(0.028)	(0.012)	(6.846)	(0.034)	(2.941)	(0.310)
CO (SD)	-0.004	-0.004	0.292	-0.142	-11.843	-1.323
	(0.021)	(0.009)	(5.148)	(0.036)	(3.007)	(0.340)
O_3 (SD)	0.003	-0.005	4.283	-0.086	-6.599	-0.872
	(0.017)	(0.008)	(4.245)	(0.029)	(2.489)	(0.266)
p-value: $\beta_{PM} = \beta_{CO}$	0.264	0.188	0.385	0.945	0.801	0.910
p-value: $\beta_{O3} = \beta_{CO}$	0.801	0.931	0.561	0.219	0.167	0.284
Panel B: Pollutants in concentrations						
$PM_{2.5}$ ($\mu g/m^3$)	0.00057	0.00027	0.12536	-0.00213	-0.18966	-0.01858
	(0.00041)	(0.00018)	(0.10026)	(0.00050)	(0.04308)	(0.00453)
CO (mg/m^3)	-0.00690	-0.00696	0.47441	-0.23014	-19.2388	-2.14862
	(0.03403)	(0.01512)	(8.36261)	(0.05848)	(4.88507)	(0.55221)
O_3 ($\mu g/m^3$)	0.00010	-0.00018	0.13920	-0.00281	-0.21449	-0.02836
	(0.00057)	(0.00026)	(0.13798)	(0.00094)	(0.08089)	(0.00866)
Mean Dep. Var.	0	0.70973	334.502	0	109.661	12.777
SD Dep. Var.	1	0.45389	242.377	1	86.308	9.281
N	138,337	138,337	138,337	94,679	94,679	94,679

Notes: All regressions include worker and date fixed effects. Standard errors that are two-way clustered by worker and date are reported in parentheses. Indexed outcomes are standardized such that the in-sample mean is zero and standard deviation is one. Pollutants in the top panel are measured in standard deviations relative to *zero pollution levels.* All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

More than particulate matters:
Pollution and productivity in Indian call centers

Deepshikha Batheja, Sarojini R. Hirshleifer, Jamie T. Mullins

Supplemental Appendix for Online Publication

SA1 Data

SA1.1 Pollution and weather data

Our pollution data comes from monitors that are maintained by the Central Pollution Control Board of India (CPCB), though our main specifications use measures maintained by the World Air Quality Index (WAQI) Project and accessed from the Air Quality Historical Data Platform. Pollution levels are assigned from the closest monitor to each of the five call-center offices in the productivity data.[28] We rely on the WAQI data since it has been cleaned.[29] We demonstrate that our results are robust to alternative pollutants, including the raw CPCB data, in Section SA2 below.

We use concentration measures for $PM_{2.5}$, CO and O_3 because ambient concentrations are an objective, scientific measure of pollution.[30] In contrast, pollutant-specific AQI measures are scaled based on a subjective assessment of the harm caused by different concentration levels of a given pollutant. Thus, different AQI measures can be constructed by different agencies, and rely on different research. They can have a nonlinear relationship to concentrations given agencies' assumptions and calibration methods. Even within agencies, AQI measures can change as they are updated based on new research. Furthermore, since we are continuing to learn more about the impacts of pollution, it is not possible to have a fully accurate AQI at this point.

Weather data comes from the ERA5 (cloud cover) and ERA5-Land (all other weather variables) data sets, which are maintained by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Hersbach, 2023; Muñoz Sabater, 2019). These are hourly, gridded datasets (0.25°x 0.25°for ERA5 and 0.1°x 0.1°for ERA5-Land), and we assign daily average weather conditions to each call center office based on the closest grid-point to each.

[28] The five pollution monitors from which data are used for this project are: i) IGSC Planetarium Complex, Patna - BSPCB (Bihar), ii) Ashok Nagar, Udaipur - RSPCB (Rajasthan) iii) Deshpande Nagar, Hubli - KSPCB (Karnataka), iv) Bandra, Mumbai - MPCB (Maharashtra) v) Sector 62, Noida - IMD (Uttar Pradesh).

[29] See https://aqicn.org/faq/2015-03-15/air-quality-nowcast-a-beginners-guide/ for details.

[30] We obtain these concentrations by converting from air quality index (AQI) values to concentration measures based on the guidance in U.S. EPA technical documents (EPA, 2018).

SA1.2 Call center data

We worked with two business process outsourcing companies (BPO) companies, called Call-2-Connect and Five Splash. The individual processes included in this study were originally selected for a field experiment on gender peer effects in the workplace (Batheja, 2019). That said, this dataset includes almost three times the number of workers as the experimental dataset. This is because we have pre- and post-experiment data from some processes for several months and there is turnover within processes. In other cases, however, the processes were short term, and thus, we do not have additional data. In addition, the experiment had a staggered launch, and began in one call center firm later than the other. Thus, we do not have data on all cities across the same range of months.[31] Nonetheless, there is substantial evidence that this data is broadly representative as discussed below (see Section SA1.3.)

The experiment randomized the gender composition of teams within processes. Thus, processes were selected into the experiment that met the requirements of the experimental design. In particular, the experiment design required constructing at least three teams (two mixed and one same gender team). So, only large processes with a size of at least 60 workers were selected for the experiment and therefore this analysis.

We included all the data available from all experiment processes from the beginning of the time we worked with these two companies up until March 1, 2020. We never analyzed data collected after this date, since we prefer to avoid analyzing data collected during the pandemic. We expect that this data would have substantial issues as the call centers abruptly shifted to work from home. Furthermore, the relationship between pollution and productivity may change, and overall pollution is expected to be much lower during this period. Our full sample for this study includes 2,777 workers across Hubli (17.72%), Mumbai (10.84%), Noida (20.38%), Patna (16.78%) and Udaipur (34.28%).

SA1.3 Representativeness of the data

Although we do not have productivity data for all dates in all cities (Figure SA2), the available indications suggest that our sample is largely representative of pollution levels in India, and thus

[31] In the first phase, between December 2018 to April 2019, Call-2-Connect employees (from Mumbai, Noida and Patna) were randomly allocated into teams using a matched pair randomization method based on past productivity data. In the second phase, between May 2019 and September 2019, employees from FiveS Digital (from Udaipur and Hubli) were made part of the experiment. The only exception is a process in Patna from Call-2-Connect, which was included in the experiment between May and August in 2019. To conduct the randomization, pre-experiment productivity data was collected and used. Therefore, we include the additional data collected as part of the randomization process in this analysis, along with the data used in the experiment. For some of the processes, we could access data for post-experiment additional months, especially from processes in FiveS Digital.

captures its likely overall impact on the BPO sector in India. As discussed in the main text, the average level of $PM_{2.5}$ across all of the city-worker-day observations in our sample is 66.54 $\mu g/m^3$. This is similar to the yearly average for all cities in India, which had a mean of 61.41 $\mu g/m^3$ in 2018 and 55.57 $\mu g/m^3$ in 2019. The distance between these means is small compared to their distance from the global average, which was 27 $\mu g/m^3$ in 2018 (WHO, 2022).[32] Unfortunately, the WHO data does not include the same global averages for the other pollutants of interest in this study or for additional moments of the data.

Thus, we also compare our own pollution measures for the city-day observations for which we have productivity data to those measures for all day-level observations across the entire period of our study (2018 through February 2020) for the cities in our study. The comparison of these two distributions indicates that the pollution levels captured by the days in our sample are remarkably similar to the full distribution of days in those cities (Figure SA3). Although these distributions do indicate that there is a small handful of very high CO and O_3 that are not captured in our sample, having productivity data for these few outlier days would be unlikely to substantively change our results. Furthermore, we find in this study that $PM_{2.5}$ contributes a relatively modest share of the pollution damages in our sample due to the high concentration of low to moderate $PM_{2.5}$ days. These distributions suggest that, if anything, we may be undercounting such days. Thus, the comparison of these distributions further confirms our results regarding the importance of CO to the share of damages.

SA2 Robustness Analysis

SA2.1 Additional analysis for main pollutant outcomes

Table SA2 presents results which provide additional context for the main results presented in Table 1. In particular, these robustness checks confirm the independence of the impacts driven by the three pollutants in two ways. In particular, Panel A presents a specification that only includes $PM_{2.5}$ and does not include the other pollutants. This is to provide a comparison with other work in the literature that only examines the impact of $PM_{2.5}$ on productivity. We find that this specification does not change the estimated effect of $PM_{2.5}$ on productivity relative to our main results (reproduced in Panel B for convenient comparison).

In Panel C, we include the pairwise interaction effects of the pollutants in the main results from Table 1. We do not find meaningful evidence of interactions among the three pollutants, as none of

[32] Data for 2020 is not available, and it is a small portion of our sample. It is not surprising that our average is slightly higher than the average for India, since Noida (near Delhi) is 20% of our sample and has higher pollution levels than many Indian cities.

the interactions are significant or have coefficients of meaningful size. Furthermore, adding these interactions does not have a substantive effect on the magnitude of the pollutant coefficients relative to the original estimates. While the coefficients on O_3 do lose some significance, the coefficients on $PM_{2.5}$ and CO maintain their statistical significance (as well as their magnitudes).

Table SA3 reports the main specification including three days of lagged pollutant levels. We do not find any consistent pattern of lagged effects for any of the three pollutants.

SA2.2 Alternative pollution measures

Table SA4 presents three robustness checks on our pollution measures. In Panel A, we present our results using the data downloaded directly from the CPCB. These data are much noisier than our main data, since the WAQI data are cleaned and processed. Thus, the WAQI data presents a more coherent picture of pollutant levels than do the raw data directly from CPCB (CPCB, n.d.). Nevertheless, our main results are comparable, though somewhat attenuated, when estimates are based on pollutant measures obtained directly from CPCB. We also consider specifications with $PM_{2.5}$ data from Berkeley Earth, while keeping our main measures of CO and O_3 from WAQI (Panel B). Notably, these results are remarkably similar to our original estimates.

Finally, we confirm that our results are not being driven by any outliers. The maximum values for each pollutant are substantially higher than the 99^{th} percentile values (Table SA1). Thus, in Panel C, we report the main results from Table 1, but with all three pollutants winsorized at the 99^{th} percentile. We find that the coefficients on the winsorized regressions are remarkably similar to our main specifications, suggesting our results are not primarily driven by outliers. This is not surprising, given that we find effects across the distribution in our non-parametric results.

SA2.3 Disaggregated results

We consider two sets of disaggregated results. Table SA5 presents the results disaggregated by city. This analysis splits our sample five ways and thus statistical power is substantially attenuated. Furthermore, we cannot use date fixed effects in this analysis since we are not able to exploit within date variation for this analysis. Thus, the purpose of this analysis is not to separately estimate effects for each city, since these specifications are not well-suited for such estimations.

Thus, this analysis is simply to confirm that the patterns of impacts are broadly similar across locations even though we expect significant noise and even some sign changes given the limits of identification in this approach. In that sense, this analysis is reassuring in that we do not find that a single city is entirely driving our overall results. In addition, the coefficients are generally the expected sign. Of the sixteen statistically significant coefficients on the intensive margin outcomes, for example, only one is positive, and it is only marginally significant.

We also analyze the results separately for inbound and outbound processes (Table SA6). Of course, given that we are splitting our sample, statistical power is attenuated and not all results are statistically significant. Nonetheless, we confirm with this analysis that there are pollution effects on both process types.

SA2.4 Joint impact of the pollutants

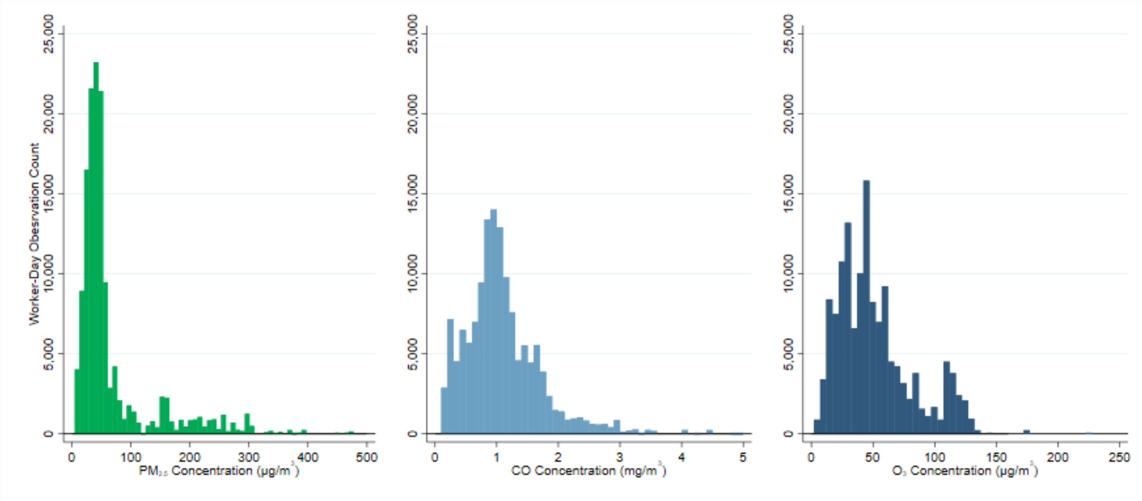
Finally, we consider the total impact of pollution and the contribution of each pollutant to productivity loss (Figure SA8). To calculate the combined impact of the pollutants observed in our data, we add together the binned non-parametric estimates of the impacts of the pollutants for each of the observed combinations of the three pollutants in our data at the city-date level.[33] Since a high pollution day for one pollutant is not necessarily a high pollution for another (see Section 3.2), we report these results framed separately by the distribution of each pollutant.

Examining these three distributions demonstrates the significant damage to productivity induced by the combined effect of these three pollutants. It is not surprising that on days that are in the upper tail of the distribution for a given pollutant that pollutant contributes a relatively large share of the lost productivity.[34] Overall the distributions of productivity loss for the three pollutants are broadly similar. In the bins including the median day for each of the pollutants, the total damages range from -0.32σ to -0.46σ, while for a day in the bin that includes the 80$^{\text{th}}$ percentile day for each of the pollutants, the total damages range from -0.56σ to -0.76σ. These impacts are substantive from a policy perspective.

[33] The impacts of pollutants are estimated in Figure 2. We do not find interaction effects across the pollutants as indicated in Section SA2.

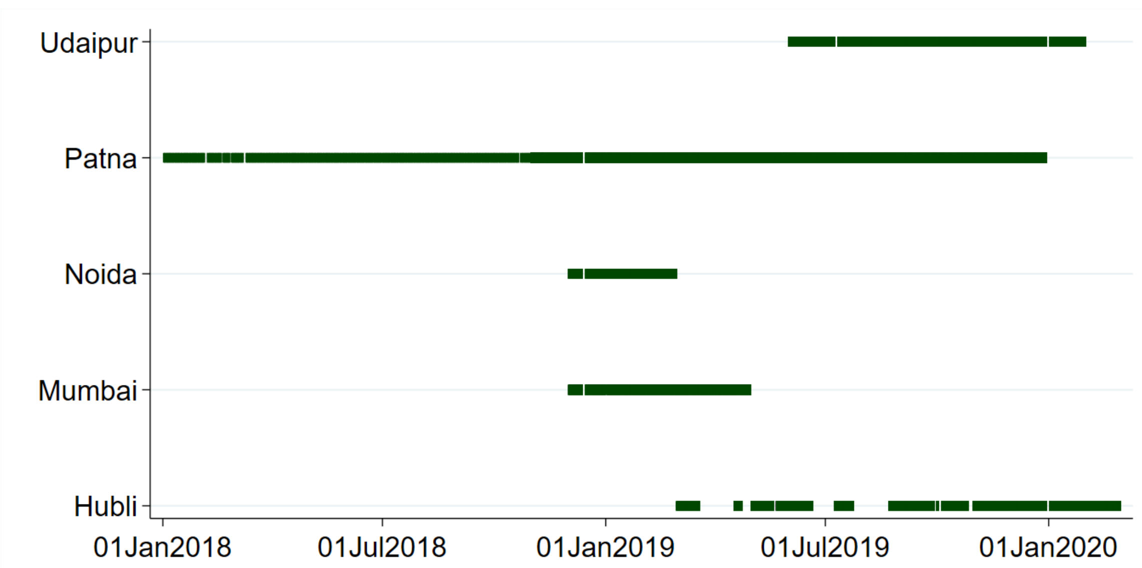
[34] Because our specification omits the lowest bin for each pollutant, there is no estimated damage from each pollutant at the low end of its owndistribution.

Figure SA1: Pollutant histograms



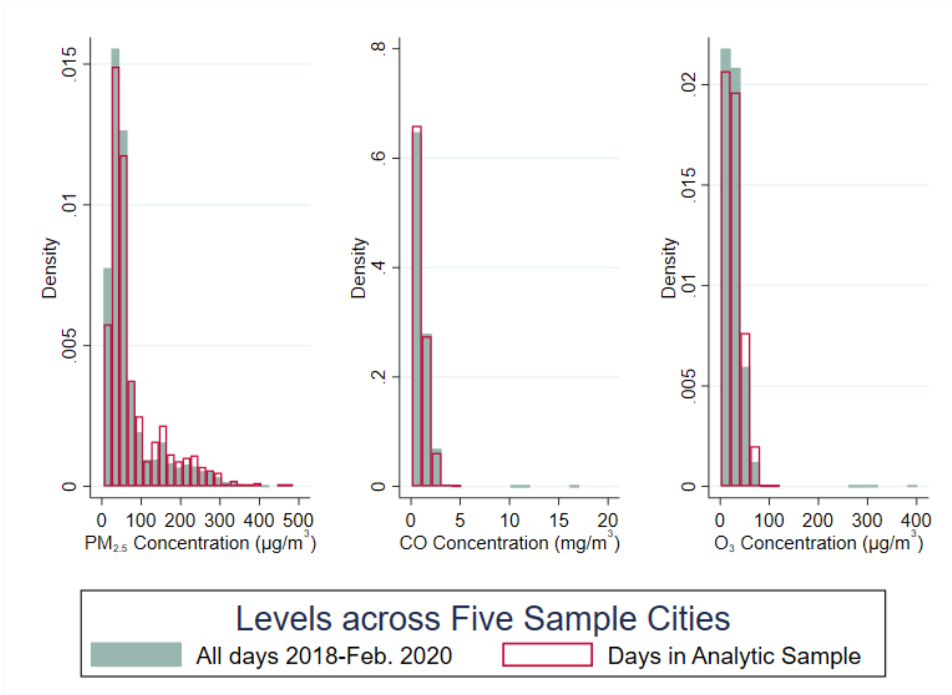
Notes: Histograms represent counts of the number of worker-days that are observed in the full at a given pollution level in the data. Histogram bin widths are 8, 100, and 5 $\mu g/m^3$ for PM$_{2.5}$, CO, and O_3 graphs, respectively.

Figure SA2: Sample dates by city

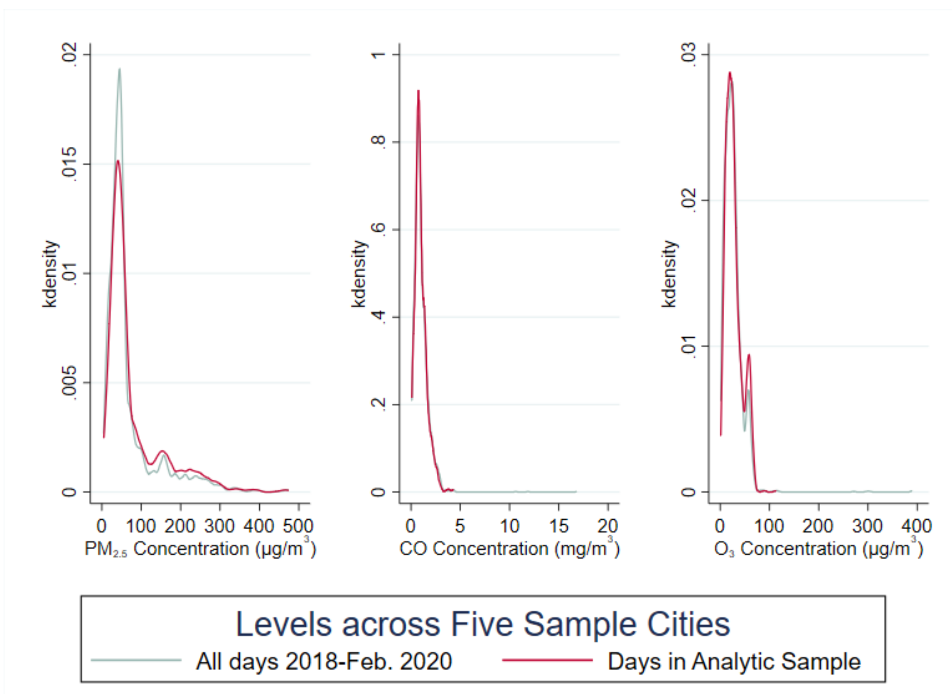


Notes: Bars indicate dates of observations used in the main sample. Short gaps are due to missing pollutant values.

Figure SA3: Pollutant distributions for analytic and full samples



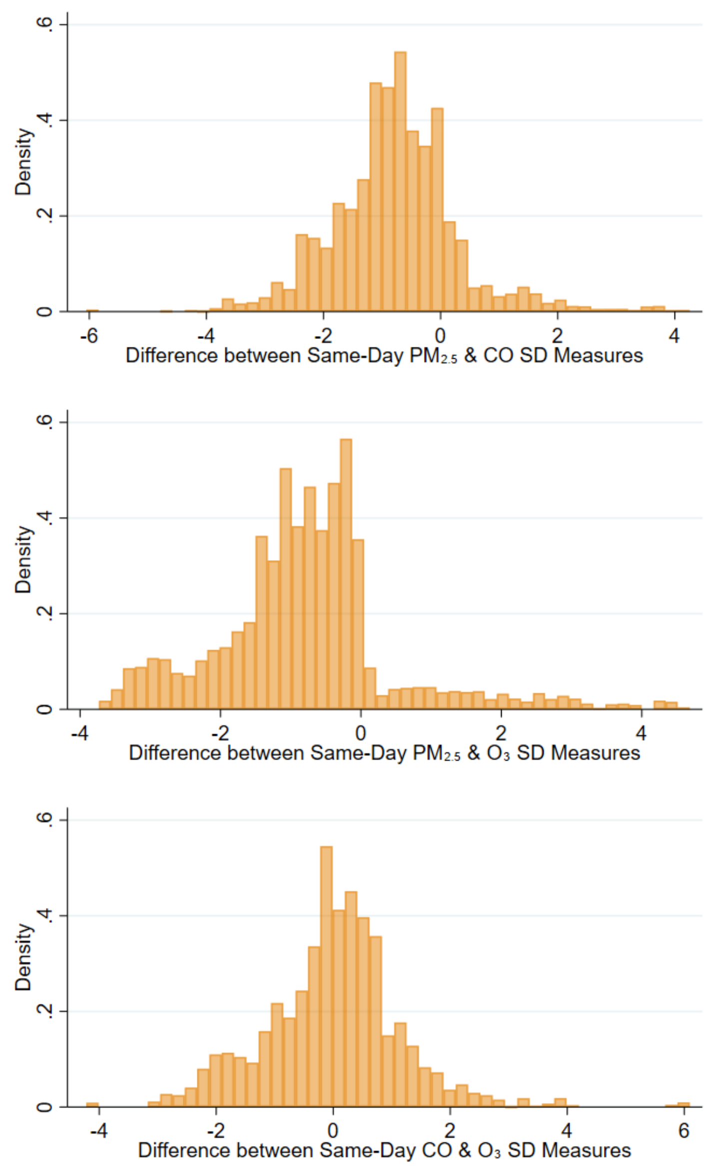
(a) Histograms



(b) K-densities

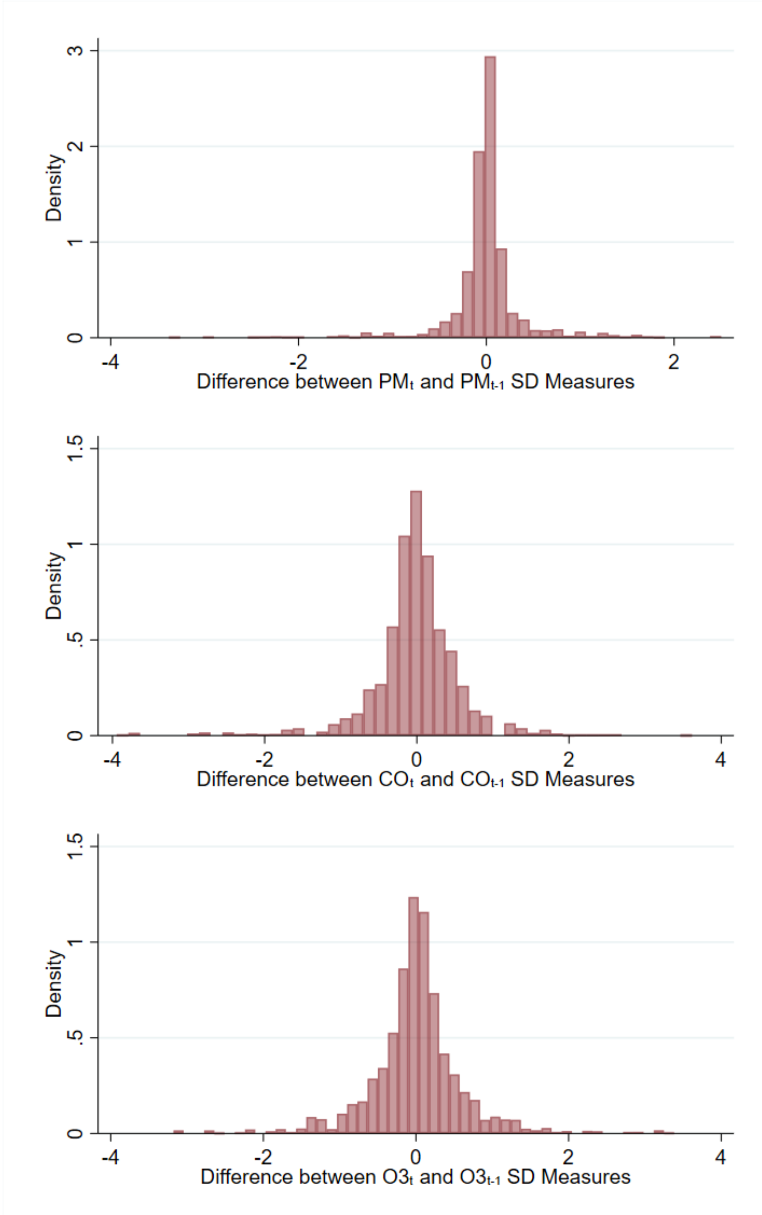
Notes: Depicts densities of pollution levels for city-days in our analytic sample and days in the five cities we study over the full period of our sample from 2018 to February 2020. Panel A depicts histograms of bin widths are 20 $\mu g/m^3$, 1 mg/m^3, and 20 $\mu g/m^3$ for $PM_{2.5}$, CO, and O_3 graphs, respectively. Panel B represents k-densities.

Figure SA4: Differences in same-day pollutant levels



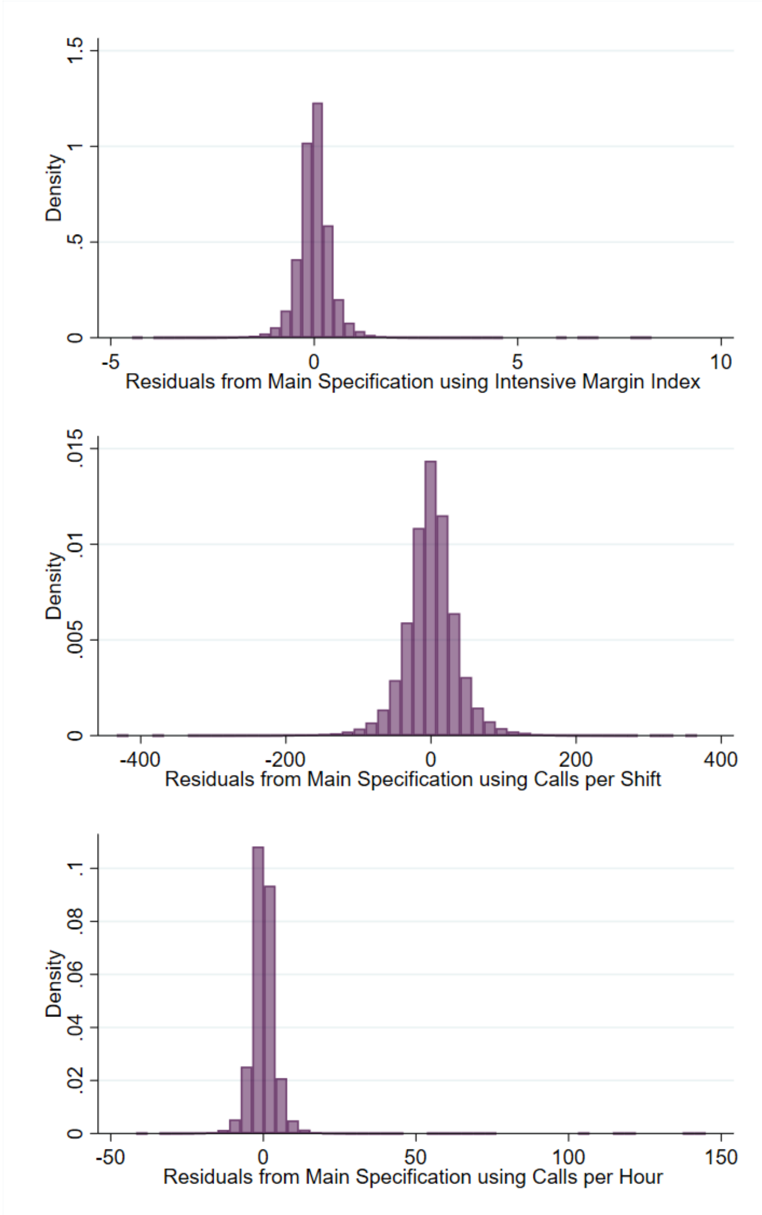
Notes: Each graph represents the differences between the measures of the two indicated pollutants experienced by a worker on a date in the full sample. Pollutants are measured in standard deviations and winsorized at the 99th percentile before differencing.

Figure SA5: Differences across contemporaneous and one-day lag pollution levels



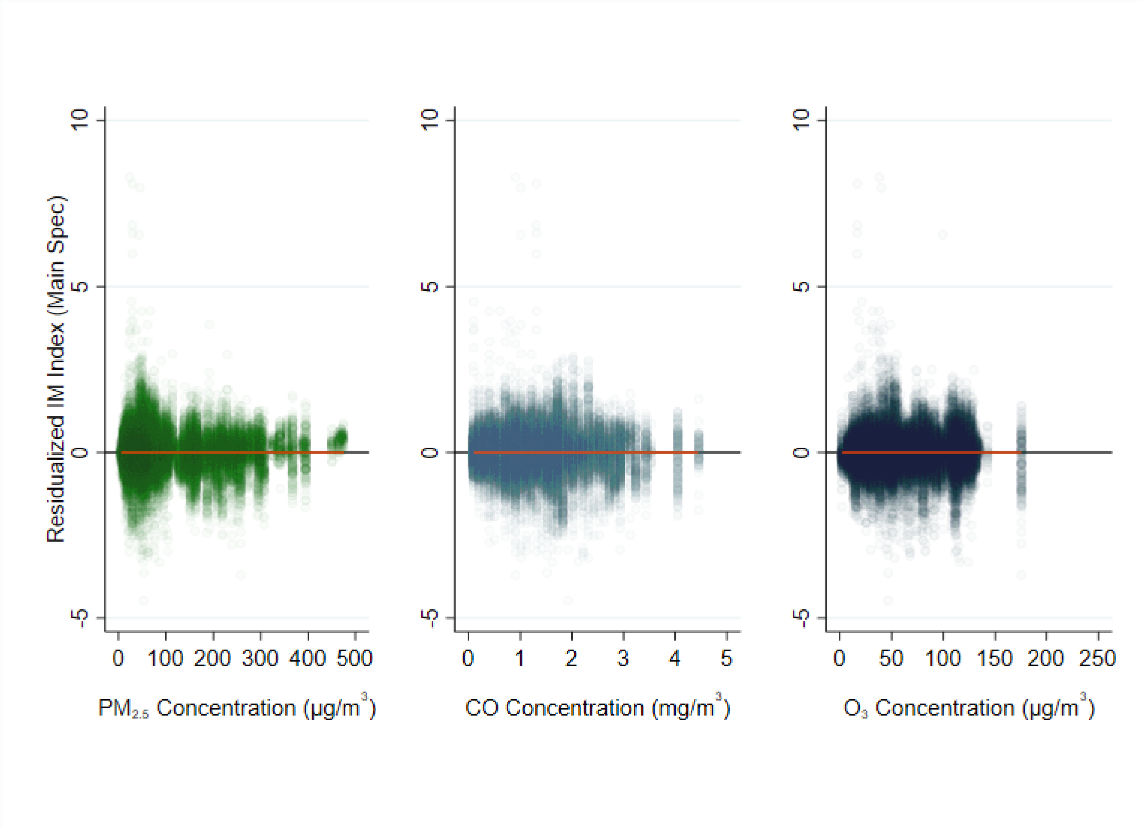
Notes: Each graph represents the differences between the measures of the contemporaneous and one-day-lag levels of the indicated pollutant experienced by a worker on a day in our extensive margin sample. Pollutants are measured in standard deviations and winsorized at the 99th percentile before differencing.

Figure SA6: Residuals from the main specification by intensive margin outcome



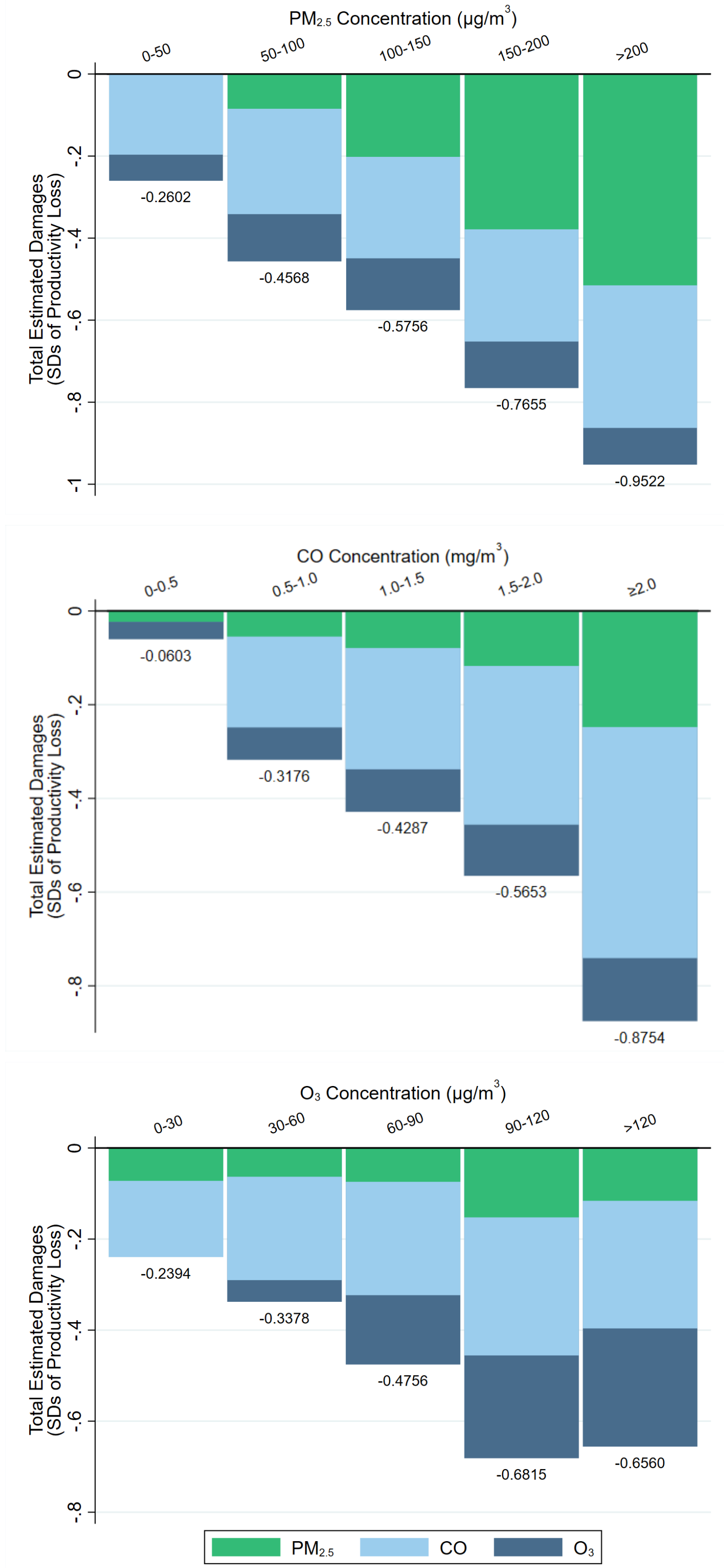
Notes: Each graph represents the distribution of the residuals that result from estimates in Columns 4-6 of Panel B of Table 1 for the indicated outcome variable.

Figure SA7: Residuals graphed against pollutant predictors



Notes: Each panel represents the distribution of the residuals that result from estimates in Columns 4-6 of Panel B of Table 1 for the indicated outcome variable. Each point is plotted with 2% opacity. Orange lines denote the best linear fit of the data in each panel.

Figure SA8: Combined productivity losses from $PM_{2.5}$, CO, and O_3



Notes: Productivity losses are estimated using the intensive margin productivity index and the binned specification from Figure 2. Estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover, as well as worker and date fixed effects. The total productivity losses are estimated as the average of the dot product between the vector of pollutant coefficient estimates and the corresponding vector of variable values for each worker-day observed in our data.

Table SA1: Summary statistics

	Mean	SD	Min	p10	Median	p90	p99	Max
Pollutants								
$PM_{2.5}$ ($\mu g/m^3$)	66.54	68.28	5.04	21.61	43.22	163.62	307.01	474.66
CO (mg/m^3)	1.05	0.62	0.10	0.30	0.91	1.72	3.04	4.96
O_3 ($\mu g/m^3$)	51.57	30.77	2.12	19.05	44.45	107.8	127.6	222.65
Weather controls								
Avg. Daily Temp. ($^\circ C$)	24	5.62	11.85	15	25.09	30.23	35.66	38.98
Total Daily Precip. (m)	0.03	0.08	0	0	0	0.11	0.33	0.95
Dew Point ($^\circ C$)	16.82	6.1	0.52	8.2	17.1	24.05	26.86	28.14
Cloud Cover (%)	0.42	0.36	0	0	0.32	0.98	1	1
City-Days: 1,224		**Workers: 2,777**		**Worker-Days: 138,337**				

Notes: Extensive Margin Sample includes non-work days as long as they are both preceded and followed by work days and are not part of a stretch of non-work days that is 6 days or longer. Only observations for which local measures of temperature, $PM_{2.5}$, CO, and O_3 are available are included in the analysis.

Table SA2: Individual and interacted pollutant measures

	Extensive margin			Intensive margin		
	(1)	(2)	(3)	(4)	(5)	(6)
	EM index	At work	Net login time	IM index	Calls per shift	Calls per hour
Panel A: Model with only $PM_{2.5}$						
$PM_{2.5}$ ($\mu g/m3$)	0.00055	0.00031	0.10810	-0.00238	-0.20436	-0.01913
	(0.00039)	(0.00015)	(0.08395)	(0.00049)	(0.03617)	(0.00389)
N	138,337	138,337	138,337	94,679	94,679	94,679
Panel B: Main Estimates Repeated for Ease of Comparison						
$PM_{2.5}$ ($\mu g/m3$)	0.00057	0.00027	0.12536	-0.00213	-0.18966	-0.01858
	(0.00041)	(0.00018)	(0.10026)	(0.00050)	(0.04308)	(0.00453)
CO (mg/m3)	-0.00690	-0.00696	0.47441	-0.23014	-19.2389	-2.14862
	(0.03403)	(0.01512)	(8.36261)	(0.05848)	(4.88507)	(0.55221)
O_3 ($\mu g/m3$)	0.00010	-0.00018	0.13920	-0.00281	-0.21449	-0.02836
	(0.00057)	(0.00026)	(0.13798)	(0.00094)	(0.08089)	(0.00866)
N	138,337	138,337	138,337	94,679	94,679	94,679
Panel C: Pollutants interacted						
$PM_{2.5}$ ($\mu g/m3$)	0.00117	0.00054	0.25758	-0.00298	-0.26380	-0.02630
	(0.00076)	(0.00035)	(0.17799)	(0.00093)	(0.08144)	(0.00839)
CO (mg/m3)	-0.04836	-0.01508	-14.69258	-0.27782	-24.4150	-2.46563
	(0.07466)	(0.03310)	(18.30665)	(0.11104)	(9.34961)	(1.04075)
O_3 ($\mu g/m3$)	0.00065	0.00015	0.22603	-0.00318	-0.23279	-0.03327
	(0.00124)	(0.00060)	(0.28216)	(0.00201)	(0.16751)	(0.01901)
$PM_{2.5}$ ($\mu g/m3$)$\times$ CO (mg/m3)	0.00009	0.00001	0.03641	0.00037	0.03858	0.00258
	(0.00031)	(0.00014)	(0.07619)	(0.00041)	(0.03497)	(0.00384)
$PM_{2.5}$ ($\mu g/m3$)$\times$ O_3 ($\mu g/m3$)	-0.00001	-0.00000	-0.00293	0.00000	0.00020	0.00005
	(0.00001)	(0.00000)	(0.00162)	(0.00001)	(0.00091)	(0.00010)
CO (mg/m3) $\times$ O_3 ($\mu g/m3$)	0.00045	0.00010	0.15568	-0.00005	-0.00311	-0.00059
	(0.00072)	(0.00033)	(0.17165)	(0.00127)	(0.10645)	(0.01191)
N	138,337	138,337	138,337	94,679	94,679	94,679

Notes: All regressions include worker and date fixed effects. Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and the standard deviation is one. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

Table SA3: Main estimates with lagged pollution

	Extensive margin			Intensive margin		
	(1)	(2)	(3)	(4)	(5)	(6)
	EM index	At work	Net login time	IM index	Calls per shift	Calls per hour
$PM_{2.5}$ ($\mu g/m3$)						
Day -3	0.00081	0.00035	0.19468	0.00065	0.05845	0.00569
	(0.00055)	(0.00026)	(0.12805)	(0.00048)	(0.04026)	(0.00450)
Day -2	-0.00081	-0.00032	-0.21423	-0.00075	-0.06589	-0.00671
	(0.00055)	(0.00024)	(0.13683)	(0.00084)	(0.07052)	(0.00799)
Day -1	-0.00031	-0.00015	-0.06391	0.00010	0.00299	0.00152
	(0.00056)	(0.00025)	(0.14143)	(0.00095)	(0.08250)	(0.00862)
Day 0	0.00060	0.00024	0.15604	-0.00172	-0.15018	-0.01532
	(0.00058)	(0.00026)	(0.14679)	(0.00071)	(0.06182)	(0.00649)
CO (mg/m3)						
Day -3	0.03914	0.00713	14.59895	-0.15303	-12.27600	-1.48426
	(0.04603)	(0.02096)	(11.44398)	(0.06829)	(5.78086)	(0.64080)
Day -2	-0.01136	-0.00315	-3.65905	-0.10881	-8.92320	-1.03435
	(0.04954)	(0.02394)	(11.35266)	(0.09155)	(7.73923)	(0.85292)
Day -1	-0.02814	-0.00656	-9.73490	-0.09474	-7.99692	-0.87624
	(0.06535)	(0.03041)	(15.45230)	(0.09195)	(8.05455)	(0.83643)
Day 0	0.01720	0.00513	5.35152	-0.12015	-10.38409	-1.08512
	(0.04400)	(0.01948)	(11.02443)	(0.05761)	(4.74511)	(0.55715)
O_3 ($\mu g/m3$)						
Day -3	0.00054	0.00025	0.11939	-0.00080	-0.06896	-0.00727
	(0.00078)	(0.00036)	(0.18678)	(0.00105)	(0.08780)	(0.01000)
Day -2	0.00059	0.00011	0.22172	0.00013	0.03268	-0.00114
	(0.00098)	(0.00047)	(0.22238)	(0.00125)	(0.10535)	(0.01169)
Day -1	-0.00043	-0.00013	-0.13420	-0.00164	-0.15001	-0.01391
	(0.00090)	(0.00043)	(0.20446)	(0.00136)	(0.11571)	(0.01265)
Day 0	-0.00028	-0.00031	0.03037	-0.00125	-0.08155	-0.01411
	(0.00076)	(0.00035)	(0.17854)	(0.00112)	(0.09501)	(0.01046)
N	122,693	122,693	122,693	83,197	83,197	83,197

Notes: All regressions include worker and date fixed effects. Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and the standard deviation is one. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover for days -3, -2, -1 and 0. Fewer observations are available for these analyses due to missing pollution data in lagged days.

Table SA4: Main estimates using alternative pollutant measures

	Extensive margin			**Intensive margin**		
	(1)	(2)	(3)	(4)	(5)	(6)
	EM index	At work	Net login time	IM index	Calls per shift	Calls per hour
Panel A: Pollutant measures from CPCB						
$PM_{2.5}$ ($\mu g/m3$)	0.00033	0.00015	0.06272	-0.00149	-0.15026	-0.01469
	(0.00041)	(0.00016)	(0.09551)	(0.00057)	(0.04454)	(0.00495)
CO (mg/m3)	0.00328	-0.00655	0.28448	-0.20117	-12.5115	-1.43000
	(0.03132)	(0.01241)	(6.99523)	(0.05715)	(3.83317)	(0.42254)
O_3 ($\mu g/m3$)	-0.00013	-0.00015	0.22775	-0.00205	-0.15966	-0.02553
	(0.00090)	(0.00038)	(0.19861)	(0.00101)	(0.07241)	(0.00820)
N	138,226	165,882	165,882	94,584	113,551	113,551
Panel B: $PM_{2.5}$ Measure from Berkeley Earth						
$PM_{2.5}$ ($\mu g/m3$)	0.00150	0.00053	0.42572	-0.00336	-0.28286	-0.03120
	(0.00063)	(0.00029)	(0.14812)	(0.00104)	(0.09095)	(0.00942)
CO (mg/m3)	-0.01228	-0.00777	-1.62786	-0.23030	-19.3513	-2.13941
	(0.03498)	(0.01559)	(8.56059)	(0.06634)	(5.53888)	(0.62556)
O_3 ($\mu g/m3$)	0.00012	-0.00015	0.14039	-0.00272	-0.20568	-0.02780
	(0.00058)	(0.00026)	(0.14493)	(0.00101)	(0.08677)	(0.00923)
N	123,586	123,586	123,586	84,532	84,532	84,532
Panel C: Pollutant measures winsorized at 99^{th}-Percentile						
$PM_{2.5}$ ($\mu g/m3$)	0.00043	0.00020	0.09489	-0.00255	-0.22608	-0.02242
	(0.00041)	(0.00018)	(0.10020)	(0.00053)	(0.04560)	(0.00480)
CO (mg/m3)	0.00425	-0.00241	3.28536	-0.22910	-19.2160	-2.13192
	(0.03345)	(0.01500)	(8.24806)	(0.06212)	(5.20599)	(0.58433)
O_3 ($\mu g/m3$)	0.00006	-0.00020	0.13398	-0.00280	-0.21328	-0.02837
	(0.00058)	(0.00027)	(0.14165)	(0.00091)	(0.07739)	(0.00839)
N	138,337	138,337	138,337	94,679	94,679	94,679

Notes: All regressions include worker and date fixed effects. Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and standard deviation is one for the main sample, as such only observations included in the main sample have EM and IM Index values. This results in different observation counts for the indexed values when the considered pollutants expand the sample. Model in Panel A includes only the concentration of $PM_{2.5}$ (and controls), excluding any measures of CO or O_3. Panel B is based on pollutant measures directly from the website of the Central Pollution Control Board of India. Panel C replaces the main measure of $PM_{2.5}$ with a gridded measure produced by Berkeley Earth (CO and O_3 measures are those from the main analysis. Panel D repeats the main, concentration-based analysis with each pollutant measure winsorized at the 99^{th} percentile. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

Table SA5: Main estimates by city

	Extensive margin			Intensive margin		
	(1) EM index	(2) At work	(3) Net login time	(4) IM index	(5) Calls per shift	(6) Calls per hour
Panel A: Hubli						
$PM_{2.5}$ ($\mu g/m3$)	0.00092	0.00055	0.13519	-0.00127	-0.08614	-0.01410
	(0.00103)	(0.00050)	(0.22558)	(0.00065)	(0.05279)	(0.00641)
CO (mg/m3)	-0.06998	-0.02669	-18.65677	-0.03162	-2.92583	-0.26490
	(0.12222)	(0.05774)	(26.90874)	(0.09311)	(7.35484)	(0.94009)
O_3 ($\mu g/m3$)	-0.00469	-0.00199	-1.14498	-0.00031	-0.02938	-0.00259
	(0.00210)	(0.00095)	(0.48866)	(0.00115)	(0.08453)	(0.01222)
N	19,928	19,928	19,928	14,535	14,535	14,535
Panel B: Mumbai						
$PM_{2.5}$ ($\mu g/m3$)	-0.00422	-0.00203	-0.90028	0.00058	0.08615	0.00135
	(0.00279)	(0.00127)	(0.64191)	(0.00465)	(0.39409)	(0.04323)
CO (mg/m3)	0.14633	0.07083	30.99771	-0.38538	-33.23706	-3.48809
	(0.07977)	(0.03620)	(18.74925)	(0.19742)	(16.99952)	(1.80437)
O_3 ($\mu g/m3$)	-0.00060	-0.00045	-0.04277	-0.00759	-0.61915	-0.07253
	(0.00129)	(0.00059)	(0.30504)	(0.00242)	(0.20833)	(0.02218)
N	15,765	15,765	15,765	11,526	11,526	11,526
Panel C: Noida						
$PM_{2.5}$ ($\mu g/m3$)	0.00016	0.00011	0.01709	-0.00168	-0.16284	-0.01330
	(0.00053)	(0.00025)	(0.12079)	(0.00051)	(0.05348)	(0.00371)
CO (mg/m3)	-0.03560	-0.01849	-6.86752	-0.01333	0.51777	-0.29988
	(0.06494)	(0.02929)	(14.96298)	(0.06219)	(6.12969)	(0.49517)
O_3 ($\mu g/m3$)	0.00177	0.00070	0.45827	0.00125	0.11700	0.01031
	(0.00109)	(0.00049)	(0.25006)	(0.00105)	(0.09606)	(0.00907)
N	16,195	16,195	16,195	9,053	9,053	9,053
Panel D: Patna						
$PM_{2.5}$ ($\mu g/m3$)	-0.00020	-0.00004	-0.07234	-0.00062	-0.05715	-0.00524
	(0.00044)	(0.00020)	(0.10832)	(0.00030)	(0.02538)	(0.00276)
CO (mg/m3)	-0.01170	-0.00992	-0.20547	-0.02067	-1.84930	-0.17993
	(0.06057)	(0.02698)	(14.39673)	(0.03682)	(3.14382)	(0.34110)
O_3 ($\mu g/m3$)	0.00081	0.00018	0.28560	0.00033	0.05061	0.00059
	(0.00145)	(0.00062)	(0.36213)	(0.00079)	(0.06762)	(0.00744)
N	27,323	27,323	27,323	21,016	21,016	21,016
Panel E: Udaipur						
$PM_{2.5}$ ($\mu g/m3$)	-0.00250	-0.00134	-0.45900	-0.00010	-0.00066	-0.00171
	(0.00133)	(0.00065)	(0.28752)	(0.00034)	(0.02962)	(0.00325)
CO (mg/m3)	0.10175	0.05746	17.16814	-0.04122	-3.20456	-0.41069
	(0.06105)	(0.02856)	(14.89170)	(0.02147)	(1.57372)	(0.24719)
O_3 ($\mu g/m3$)	0.00110	0.00035	0.32899	0.00051	0.06610	0.00219
	(0.00184)	(0.00091)	(0.39131)	(0.00043)	(0.03432)	(0.00461)
N	59,126	59,126	59,126	38,549	38,549	38,549

Notes: All regressions include worker fixed effects. Date fixed effects are omitted because no variation in pollution (or weather conditions) exists within a city on a given day. Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and the standard deviation is one. Each panel reports the results of the main estimation limited to the sub-sample of workers in the indicated city. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

Table SA6: Main estimates by process type

	Extensive margin			Intensive margin		
	(1)	(2)	(3)	(4)	(5)	(6)
	EM index	At work	Net login time	IM index	Calls per shift	Calls per hour
Panel A: Outbound						
$PM_{2.5}$ ($\mu g/m3$)	-0.00006	0.00000	-0.02869	-0.00107	-0.10781	-0.00793
	(0.00064)	(0.00030)	(0.14572)	(0.00060)	(0.05831)	(0.00494)
CO (mg/m3)	-0.01529	-0.00897	-2.39842	-0.12096	-8.99041	-1.24978
	(0.05492)	(0.02485)	(12.62314)	(0.06885)	(6.22932)	(0.60207)
O_3 ($\mu g/m3$)	0.00074	0.00022	0.23001	-0.00219	-0.17173	-0.02173
	(0.00080)	(0.00036)	(0.18832)	(0.00120)	(0.10227)	(0.01106)
N	35,513	35,513	35,513	23,071	23,071	23,071
Panel B: Inbound						
$PM_{2.5}$ ($\mu g/m3$)	-0.00014	-0.00003	-0.04780	-0.00044	-0.04072	-0.00376
	(0.00036)	(0.00017)	(0.08503)	(0.00023)	(0.01962)	(0.00223)
CO (mg/m3)	0.01345	0.00675	2.72075	-0.04123	-3.65129	-0.36299
	(0.04484)	(0.02078)	(10.40064)	(0.02503)	(2.07800)	(0.24206)
O_3 ($\mu g/m3$)	0.00021	-0.00022	0.21568	0.00018	0.04758	-0.00190
	(0.00092)	(0.00041)	(0.22230)	(0.00043)	(0.03623)	(0.00417)
N	102,824	102,824	102,824	71,608	71,608	71,608

Notes: All regressions include worker and date fixed effects. Process type is determined by whether the calls handled were initiated by the work ("Outbound") or received by the worker ("Inbound"). Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and standard deviation is one. Each panel reports the results of the main estimation limited to the sub-sample of workers in the indicated city. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

Table SA7: Estimates for Figure 1

	(1) EM index	(2) IM index
$PM_{2.5}$ 1-2 SDs	-0.021	-0.043
	(0.055)	(0.074)
$PM_{2.5}$ 2-3 SDs	0.036	-0.382
	(0.060)	(0.083)
$PM_{2.5}$ 3-4 SDs	0.034	-0.516
	(0.068)	(0.128)
$PM_{2.5}$ >4 SDs	0.098	-0.492
	(0.112)	(0.139)
CO 1-2 SDs	0.124	-0.167
	(0.050)	(0.056)
CO 2-3 SDs	0.042	-0.239
	(0.059)	(0.082)
CO 3-4 SDs	0.055	-0.274
	(0.067)	(0.115)
CO >4 SDs	0.168	-0.707
	(0.097)	(0.165)
O_3 1-2 SDs	-0.043	-0.062
	(0.036)	(0.044)
O_3 2-3 SDs	0.013	-0.186
	(0.047)	(0.072)
O_3 3-4 SDs	-0.006	-0.275
	(0.049)	(0.080)
O_3 >4 SDs	-0.110	-0.357
	(0.108)	(0.124)
N	138,337	94,679

Notes: Reports the regressions that are used to create Figure 1. All regressions include worker and date fixed effects. Standard errors clustered by worker and date are reported in parentheses. Index outcomes are standardized such that the in-sample mean is zero and the standard deviation is one. Pollutants are measured in standard deviations relative to *zero pollution levels.* All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.

Table SA8: Estimates for Figure 2

	(1) IM index
$PM_{2.5}$ 50-100 $\mu g/m3$	-0.091
	(0.059)
$PM_{2.5}$ 100-150 $\mu g/m3$	-0.216
	(0.086)
$PM_{2.5}$ 150-200 $\mu g/m3$	-0.405
	(0.100)
$PM_{2.5}$ >200 $\mu g/m3$	-0.551
	(0.136)
CO 0.5-1.0 mg/m3	-0.207
	(0.080)
CO 1.0-1.5 mg/m3	-0.277
	(0.095)
CO 1.5-2.0 mg/m3	-0.363
	(0.115)
CO >2.0 mg/m3	-0.528
	(0.134)
O_3 30-60 $\mu g/m3$	-0.051
	(0.046)
O_3 60-90 $\mu g/m3$	-0.163
	(0.073)
O_3 90-120 $\mu g/m3$	-0.242
	(0.078)
O_3 >120 $\mu g/m3$	-0.278
	(0.117)
N	94,679

Notes: Reports the results used to construct Figure 2. All regressions include worker and date fixed effects. Standard errors clustered by worker and date are reported in parentheses. The outcome is standardized such that the in-sample mean is zero and the standard deviation is one. All estimates include quadratic controls for temperature, precipitation, humidity, and cloud cover.