Severe air pollution and labor productivity

Teng Li, Haoming Liu, and Alberto Salvo*

Department of Economics, National University of Singapore, Singapore 117568

January 4, 2015

Preliminary–Comments most welcome

Abstract

We examine daily worker-level output over 15 months for a panel of 98 workers at an indoor textile plant in the northern Chinese province of Hebei, close to Beijing. Workers earn piece-rate wages, with no base pay or minimum pay, for homogeneous tasks performed over fixed 8-hour shifts. Over the sample period, ambient fine-particle (PM2.5) mass concentrations measured at an outdoor air monitor located 2 km from the plant ranged between 3 and 773 micrograms per cubic meter ($\mu g/m^3$, 8-hour means), variation that is an order of magnitude larger than what is observed in the rich world today. We document very large reductions in productivity, of the order of 15%, over the first 200 $\mu g/m^3$ rise in PM2.5 concentrations, with the drop leveling off for further increases in fine-particle pollution. We also find reduced product quality as pollution rises. Our empirical model allows for selection into work attendance, though we do not find particle pollution to be a significant determinant of non-attendance. Subsequent research should verify the external validity of our findings.

Keywords: Air pollution, labor productivity, PM2.5, environmental damage

*Teng Li, Haoming Liu, and Alberto Salvo, Department of Economics, National University of Singapore, 1 Arts Link, Singapore 117570. Email: liteng@nus.edu.sg, ecsliuhm@nus.edu.sg, albertosalvo@nus.edu.sg, Phone: +65 6516 4876, Fax: +65 6775 2646

1 Introduction

At 7pm on January 12, 2013, the mass concentration of particulate matter of diameter up to 2.5 microns (PM2.5) in Beijing's air, as measured by the US Embassy, reached 886 micrograms per cubic meter ($\mu g/m^3$). This severe level of fine-particle pollution in ambient air is over 30 times higher than the World Health Organization's guidelines for exposure sustained over a period of 24 hours—a mean 24-hour concentration for PM2.5 of no more than 25 $\mu g/m^3$ is deemed "healthy." The National Ambient Air Quality Standards (NAAQS) set by the United States Environmental Protection Agency establish a PM2.5 "primary standard"—intended to provide public health protection including the health of vulnerable groups—of 12 $\mu g/m^3$ over one year of exposure, and a 24-hour standard of 35 $\mu g/m^3$.[1] China's severe environmental degradation is the by-product of the nation's fast growth in economic activity, and fear of dampening this growth rate is widely regarded to be one key reason behind the Chinese government's revealed hesitation in taking swift action to abate air pollution (Ebenstein et al., 2014). While research on the social benefits of cleaner air has focused on public health, such as longer life expectancy (e.g., Chen et al., 2013), another potentially large source of gains, labor productivity, has received considerably less attention in the policy debate. We conjecture that the relative absence of labor productivity in policy discussions concerning the damage caused by China's air owes at least in part to the paucity of rigorous empirical work on this potentially pervasive manifestation of morbidity.

To the best of our knowledge, only four previous studies have directly addressed the impact of air pollution on labor supply or labor productivity while attemping to overcome some major hurdles that the literature faces. First, people who live in areas with higher ambient air pollution may differ from residents in areas with better air quality, making it difficult to use cross-sectional data to establish a causal relationship between air quality and labor productivity. The second challenge is that because administrative datasets typically do not record worker-level output or earnings on a daily basis, one cannot use daily variation in air quality to examine its effect on worker productivity

[1] See http://www.epa.gov/air/criteria.html.

while flexibly controlling for individual heterogeneity. Nevertheless, the limited existing empirical evidence indeed suggests that air quality has a significant adverse effect on labor productivity and labor supply.[2]

Using a dataset that contains the daily work performance and operating conditions of citrus pickers in southern California in 1973/74, Crocker and Horst (1981) find that in-sample ozone pollution reduces the productivity of the outdoor workers by up to 2%. This finding is confirmed by Graff-Zivin and Neidell (2012) who use payroll data from a farm in California's Central Valley. Graff Zivin and Neidell show that a 10 ppb (parts per billion) increase in average ozone exposure results in a 4% reduction in the productivity of workers picking berries and grapes. Interestingly, in their setting, workers are paid in proportion to their individual output, i.e., on a piece-rate basis, which acts as an incentive for workers to perform. In a recent working paper, Chang et al. (2014) find that outdoor air pollution also in California—but now examining fine-particle (PM2.5) rather than ozone pollution—also impacts the indoor work environment. Chang et al find that a 10 μg/m^3 increase in outdoor PM2.5 concentrations decreases worker productivity by roughly 6%, while having no significant effect on working hours. One challenge Chang et al face is that the indoor work environment they study is naturally ventilated and, as the authors point out, variation in temperature, if not adequately controlled for, may confound variation in fine-particle concentrations, which is their main variable of interest. Finally, Hanna and Oliva (2014) explore the exogenous closure of a large refinery in Mexico City, a policy that induced variation in air quality. Examining aggregate data from an administrative source, and using a difference-in-difference design, Hanna and Oliva find that the reduction in air pollution induced by the policy change led to a 4% increase in weekly working hours for workers living within a 5 km radius of the refinery.

By examining a panel of individual workers at a textile mill located in the province of Hebei, in northern China, our paper adds to this sparse empirical literature. In

[2]Ostro (1983), for example, investigates the relationship between particle pollution (total suspended particles, TSP) and labor supply using cross-sectional survey-based, (though individual-level) data from the 1976 US Health Interview Survey. Workers responded to a survey question asking them how many days in the past two weeks had illness or injury prevented them from working. Ostro reports that "a 10% decrease in ambient levels of TSP is related to a 4.4% decrease in WLD (work loss days)."

particular, we are able to examine the effect of severe air pollution on labor productivity. As alluded to in the opening paragraph, the variation in air pollution observed in China today is much larger than what the extant literature has studied. For instance, daily PM2.5 concentrations range between 2 and 60 μg/m^3 in Chang et al. (2014) whereas it varies between 3 and 963 μg/m^3 in our data. The large variation in particulate pollution that we observe enable us to examine the magnitude and the shape of the relationship between air quality and labor productivity, which can potentially be nonlinear. Knowing whether there is a nonlinear dose-response will allow us to more accurately estimate the economic cost of air pollution to which workers in developing countries today, from China to India to Indonesia, are exposed—levels that are an order of magnitude higher that in developed countries in North America and Europe. We note that even at ambient air levels observed in the United States, PM2.5 is understood to be a major source of health damage (Fann et al., 2012).

We gained access to daily payroll data for over 100 workers at a common workplace during 15 months that span the four seasons of the year. We argue that the labor institution we study coupled with the structure of the data we have access to are ideally suited to the the research question at hand, sharing several features of recent impactful studies, at lower levels of pollution, in particular, Graff-Zivin and Neidell (2012). Like Graff Zivin and Neidell (GZN hereafter), we are able to observe labor supply choices and outcomes for the same worker over time, as ambient air pollution fluctuates day in day out, allowing us to control for worker heterogeneity. Similar to GZN's setting, the workers in our study are paid on a piece-rate basis for their individual output, which rewards them for the level of effort that they individually supply (and has the added feature, as GZN argue, that this plausibly reduces measurement error in this relatively high-frequency output variable).[3] The piece rate does not vary across workers, consistent

[3]To be precise, some of the workers in GZN are paid in proportion to their joint output, which is not the case in our labor market. Further, in our setting there is no minimum wage or base pay, and workers are on long-term contracts. Moreover, in contrast to Chang et al. (2014), the work shift in our setting is of fixed duration, so workers do not choose how many labor hours to supply, conditional on work attendance (which we model). Otherwise, it is conceivable that a worker might choose the number of hours worked in part as a function of the level of airborne contaminants, and this choice might be made jointly with the effort level. We also need not worry that workers might slacken to prolong their normal work hours to earn an overtime rate.

with the homogeneity across tasks and workstations.[4] Unlike GZN's outdoor setting, however, the indoor work environment we study is temperature controlled,[5] which we argue enables us to directly control for an important possible confounder of the effect of pollution on labor productivity. An added benefit of workplace temperature control is that it provides us with an exclusion restriction to identify possible selection into work attendance, and thus control for selection bias, to the extent that outdoor temperature—and weather more generally, including precipitation and wind—shifts the probability of showing up to work, e.g., by changing the value of the outside option such as leisure (Graff-Zivin and Neidell, 2014). Importantly, while temperature controlled, indoor air is not filtered, and we are planning to measure the presumed tight correlation between PM2.5 concentrations in the indoor and outdoor environments. (By "presumed" we mean in light of the significant labor effect we estimate since, like Chang et al. (2014), we observe only outdoor concentrations; this is also consistent with the literature on "indoor-outdoor ratios" in environmental engineering.) Finally, our labor output data consists not only of quantity produced (meters of fabric) but also the quality of production (meters of defective fabric), allowing us to examine the effect of air pollution on the product of labor along this additional margin.

We find that higher mass concentrations of PM2.5 in outdoor air, as recorded by an air monitor situated 2 km from the textile plant, has a significant negative impact on worker productivity. While this sign is consistent with what has been found in the small extant literature, the magnitudes are larger than those, to the best of our knowledge, ever examined: of the order of a 100-meter output reduction out of about 500 meters per worker per shift for a 150 μg/m^3 increase in ambient PM2.5 levels from low levels. In addition, we find that once fine-particle concentrations reach a certain threshold, the adverse impact on worker productivity levels off. These results are new to the literature.

Should the firm invest in an air purifying system?

(To be completed. If benefit-cost analysis suggests that the firm should filter air as

[4] Again for comparison, GZN observe workers picking different crops, with the piece rate varying across the crops. Since the authors pool observations across crops, they need to standardize output units.

[5] For comparison, as noted above, there is no temperature control in the pear-packing plant that Chang et al. (2014) study.

it penetrates the indoor environment from the outside, as we are aware some schools in large Chinese cities have recently begun doing, this indicates that as research as ours is shared more widely, more central purifying systems will be installed in manufacturing and other buildings across China.)

Back of the envelope estimate of the effect on China's productivity.

(To be completed. We quantify the economic magnitude of the productivity effects we find by applying our estimates to manufacturing workers across different parts of China during different seasons, with varying spatial and temporal levels of particle pollution.)

2 Data and descriptive analysis

Our data on worker output is made available thanks to a special agreement with management at a textile operation in a city in the northern province of Hebei, China.[6] The operation is part (namely, a "department") of a larger industrial complex, obtaining yarn from an upstream operation, processing the input into fabric rolls which are then sold to buyers across China and abroad.[7] The operation we examine runs around the clock, 7 days per week, with the exception of China's main public holidays when the operation shuts down, in particular the weeks of Chinese New Year and National Day. "Line" (production) workers, who are permanent employees, have the right to 3 days of paid leave per month, paid at CNY 30 per leave.[8] The non-attendance we observe in the data (see below) indicates that leave, both planned and unplanned, is fairly balanced within worker over the seven days of the week (Monday through Sunday). Non-attendance is also balanced across workers and over time, such that the operation is always adequately

[6]For reference, the province of Hebei surrounds Beijing. We do not name the city to protect the operation's confidentiality.

[7]The two upstream operations to the "textile" department that we examine are: (i) the "preparatory" department, which receives cotton from the nearby provinces of Shandong and Jiangsu and the northwestern province of Xinjiang, and then cleans and draws the cotton; and (ii) the "yarn" department, whose main task is to spin and pass yarn. The textile department that we examine takes the yarn and weaves threads into fabric. We are unaware of any input (including electricity) shortages during the sample period. Also, in view of its widespread geographic nature, demand for the operation's produce is unlikely to depend on local economic activity.

[8]For a rough conversion into USD, divide CNY by 6. Thus CNY 30 translates into approximately USD 5. In our sample, the median worker pay per shift is CNY 50 (see below). Thus a worker's earnings when working tend to exceed those on planned leave.

staffed, at about 25 workers per shift per day (date-shift).[9]

In total, three similarly sized teams of line workers take turns working during three shifts of fixed, 8-hour duration, starting at: (i) 0 am (to 8 am), (ii) 8 am (to 4 pm), and (iii) 4 pm (to 0 am the next day). The teams rotate among these shifts every four days, such that over every short period in the sample each worker typically works in each of the three shifts. Each worker's workstation consists of 10 machines (or looms, valued at about USD 30,000 each), and the job description is to supervise the looms as they weave threads into fabric. Typically, 1-2 threads will break apart per machine per hour, requiring skill and effort from the worker to reconnect the thread, a task that might take several minutes. Thus, a worker will typically reconnect 10-20 threads across the 10 machines every hour (and this number can rise to 50-60 threads in extreme cases). The work being performed is homogeneous across workers and workstations. The machines are quite new, having been purchased in block when the firm was privatized in the early 2000s.[10] Day in, day out, a worker tends to return to her workstation. While the operation is capital intensive, its productivity depends critically on the quality of the labor. Our understanding is that the quality of labor is a function of the skill (experience) and attentiveness (effort level) of the worker. The appropriate unit of analysis is the individual worker, as there is minimal complementarity across workers.

With worker attention being critical and complementarity across workers being minimal, workers are paid by the amount of good-quality fabric they individually produce in each shift. The sample period is April 1, 2013 to June 30, 2014 and, during this period, the piece rate was increased once, from CNY 0.07 per meter of fabric prior to September 29, 2013 to CNY 0.1 subsequently.[11] This piece rate does not vary across

[9]To keep machines from staying idle, "bench" workers who work across the three manufacturing departments are brought in to substitute for non-attendant textile department workers (or "line" workers). The number of bench (substitute) workers is low, with a median of 1 bench worker across date-shifts. For these occasional workers we also observe their output, but choose to drop them from our baseline sample. In addition, workers from administrative functions have occasionally been brought in to operate the machines, at times of low attendance (we do not observe the output of these workers but understand that their use is minimal).

[10]We do not observe machine breakdowns. The supervisor for the textile department informed us that machines break down rarely. In the event that a machine breaks down, the worker's pay is prorated based on her output using the functioning machines. Every batch of 10 machines (a workstation) undergoes maintenance every fortnight.

[11]This is one reason why we control for quarter fixed effects in the empirical analysis.

workers, consistent with the homogeneity across tasks and workstations. There is no minimum wage, base pay or threshold level above which variable compensation applies (as there is, for example, in GZN). At the end of her shift, a worker marks the point at which her production ended and at which the production of her colleague, working the subsequent shift at the same workstation, begins. The rolls of fabric, once complete, are subsequently inspected for defects and payroll records are updated. While workers are paid based on the amount of defect-free production, they do not receive pecuniary punishment for producing defects. To the best of our knowledge, there are no disputes over what constitutes defect-free fabric, or how individual-level output is recorded. Across workers in our sample, the mean tenure at the firm is 19 years and the mean employment contract duration is 7 years.[12]

The key aspect of the individual level productivity data we have access to is its panel structure. Similar to GZN, we are able to follow the same worker, date-shift by date-shift, which allows us to control for worker heterogeneity. In addition to the quantity of individual output—total and defective, in meters—produced by each worker (ID) over each shift, the payroll records contain the worker's age, gender, education level, years of work experience, and years worked for the plant, which we can use to learn more about worker heterogeneity. As we illustrate below, our source of identification, as in GZN, is the day-to-day covariance between the state of outdoor air pollution and the individual worker's productivity, once we control for seasonality and other potential confounders such as temperature in the workplace.

Similar to GZN, the worker chooses whether to attend work. Besides planned leave, which is predetermined and thus unlikely to depend on day-to-day variation in environmental quality, the worker might not attend work due to herself or a family member falling sick, meteorological and other shocks raising the cost of commuting, or the value of the outside option shifting (e.g., a leisure day spent outdoors). Such unplanned non-attendance may in part depend on the state of air, or be correlated with it (e.g., via

[12]This reflects low hiring and firing during our sample period. In our sample, we observe 20 line workers who were transferred to other departments (or fired) and 28 workers who were transferred from other departments (or hired).

weather shocks), and we subsequently model the worker's selection into work. Our understanding, however, is that while health may drive work attendance, this is less likely to be the case for shifts in the commuting cost and possibly the leisure option, since about 90% of workers live in nearby housing subsidized by the employer. Excluding four or more consecutive days of non-attendance, as such events are likely due to planned leave, the mean number of non-attendance days in our sample is 1-2 per worker per month. We also note that in our setting the worker does not choose her labor supply in terms of hours worked, as this margin is fixed at 8 hours (likely due to the complementary capital whose cost is incurred by the plant owners).

To illustrate the importance of worker heterogeneity, for every worker in our sample we calculate her mean individual output per shift during the entire sample period. The distribution of worker performance is plotted in panel (a) of Figure 1 (the figure plots output including defective fabric). The mode is around 525 meters per worker-shift. The panel indicates that the most productive workers can sustain a production rate that is up to 40% higher than that of the least productive workers. Besides computing the mean productivity within worker over time, we compute the mean productivity per worker for each date-shift. The day-to-day variation in the average productivity of the workforce is plotted in panels (b) to (d) of Figure 1, separately for the 0 am, 8 am and 4 pm shifts. This variation is of key importance to our empirical strategy—our task is to uncover the extent to which this temporal variation in output is driven by variation in ambient PM2.5 concentrations.

Figure 2 provides the same descriptive statistics as Figure 1 but for defective, rather than total, output. Some workers produce as little as 0 meters of defective fabric per shift during the sample period,[13] whereas others produce as much as 7 meters of defects per shift on average. The variation of defective output is similar across shifts, with a mode at about 2 meters per worker, less than 0.5% of the modal output.

Pollutant concentrations, annual and weekly cycles, and meteorology. Hourly mass concentrations for several ambient air pollutants, in $\mu g/m^3$, namely for

[13]For one of the three teams of workers, data on defective output is missing. We are currently investigating why.

PM2.5, PM10, SO_2, NO_2, CO, and O_3, are obtained from the Chinese Ministry of Environmental Protection. The closest official measurement site is located in an outdoor environment, 1.7 km from the plant we analyze. There are three other air monitoring stations in the city, located at distances ranging from 3 to 5 km from the plant. Pollutant concentrations are highly correlated across monitoring sites. Specifically, we compute daily mean concentrations (namely, means over 24 hour readings, 0 am to 11pm) for each monitor-date pair over all dates in our worker sample, and then obtain pairwise correlation coefficients. For PM2.5, cross-monitor correlations range between 0.85 and 0.93.

We hypothesize that fine particle pollution is driven largely by determinants that act at the regional, rather than local, level, such as meteorological and atmospheric conditions (e.g., precipitation, wind, thermal inversions) and economic activity (e.g., fossil fuel combustion that responds to both annual and weekly cycles as well as to idiosyncratic shocks to anthropogenic activity). To audit the Chinese Ministry's PM2.5 concentration series, we thus obtain hourly PM2.5 concentrations measured at the US Embassy in Beijing, several hundred km away from our plant, over the same recent sample period.[14]

To further audit the Chinese Ministry's air quality data over our sample period, we regress daily mean PM2.5 concentrations reported for the nearby site on week-of-year fixed effects (to account for the natural and anthropogenic annual cycle), day-of-week fixed effects (weekly cycle), indicators for China's main public holidays as well as local public holidays, and daily mean meteorological parameters. Meteorological conditions are obtained from China's Department of Meteorology, namely for precipitation, ground temperature, wind speed, and wind direction.[15] Table A.1 in the appendix reports OLS estimates for this regression, first without controlling for meteorology (column I) and then with meteorological controls (column II). Figure A.1 in the appendix plots the values of the dependent variable over time (raw daily mean PM2.5 concentrations) in panel

[14]Again, we do not provide the exact distance given our non-disclosure agreement.

[15]To control for the effects of wind direction on fine particle concentrations, we generate four dummy variables to indicate the quadrant from which the wind is blowing every hour, when it blows with speed above 0.5 ms^{-1}, and then take means within date.

A, as well as residuals in panels B and C that correspond, respectively, to regressions I and II of Table A.1, without and with meteorological controls. The regression residuals indicate that despite controlling for the annual and weekly cycles, the public holiday calendar, and (even) meteorology, much variation remains. Our empirical strategy to infer the causal effect of particle pollution on worker productivity relies on such residual variation.

To match air quality with the worker output data, we aggregate hourly pollutant concentration readings during our sample period into 8-hour averages with each 8-hour period corresponding to a date-shift. Figure 3 plots the kernel density function of 8-hour PM2.5 concentrations together with some reference levels as per the US Environmental Protection Agency. Less than 8% of the observations fall in the *Healthy* or *Moderate* categories, while more than 8% of them fall in the *Hazardous* category. This wide variation in ambient particle pollution enables us to compare the benefits on worker productivity from abating pollution at the high versus the moderate levels of pollution we observe in our sample. This information can have important policy implications as Chang et al. (2014) indeed show that the adverse impact of air pollution on productivity becomes stronger as the concentration level of PM2.5 increases.[16]

Table 1 reports sample statistics. The average PM2.5 concentration level is 124 μg/m^3, which falls in the *Unhealthy I* category (over a 24-hour exposure) according to standards set by the US Environmental Protection Agency. Less than one-fifth of the workforce is male, consistent with most workers in China's textile industry being female. Years of schooling exceed 9 years, indicating that most workers completed middle school. The average worker has worked at the employer for about 19 years of total work experience, suggesting low turnover.

Figure 5 plots average worker output per shift against the contemporaneous 8-hour mean PM2.5 concentration measured 2 km from the plant. Each observation in the plot is a date-shift pair in the sample, and we average across workers in that date-shift. Panel (a), depicting total output including defects, shows several striking patterns. First,

[16]It should be noted that the PM2.5 concentration level varies between 2 to 60 μg/m^3 in Chang et al. (2014). In contrast, the PM2.5 concentration level exceeds 60 μg/m^3 in about 75% of our sample.

when fine-particle pollution is high, mean worker productivity lies at the lower end of the productivity distribution. Mean output rarely exceeds the sample mean of 509 meters per worker per shift when PM2.5 concentrations rise above 250 μg/m^3. In particular, mean output per worker per shift exceeds the sample mean only on four times out of the 100 date-shifts with PM2.5 levels above 250 μg/m^3. Second, even among date-shifts with PM2.5 concentrations below 250 μg/m^3, the figure clearly indicates a negative relationship between labor productivity and PM2.5 concentrations. Panel (b) shows that the mean production of defective fabric rises with the air pollution. The mean quantity of defects lies below the sample mean of 3.8 meters only 12 times out of the 100 date-shifts with PM2.5 value above 250 μg/m^3. For date-shifts with PM2.5 levels higher than 500 μ g/m^3, the quantity of defects never falls below 5 meters per worker per shift, and for date-shifts with PM2.5 exceeding 600 μg/m^3, defects never fall below 10 meters.

3 Conceptual framework and empirical model

To fix ideas, consider a worker who works individually over a fixed shift of eight hours. During this shift, the worker is exposed to ambient air pollution Ω^I (the superscript denotes the *indoor* work microenvironment), chooses effort level e, and produces gross output quantity $\tilde{q}$, a fraction $0 \leq \xi \leq 1$ of which comes out free of defects. The worker incurs an effort cost $c(e, \Omega^I)$, which is increasing at an increasing rate in the first argument, the level of effort, and shifts with the level of indoor pollution. We specify net, or defect-free, output quantity $q = \tilde{q}\xi = q(e)$ as an increasing and concave function of the effort level. The worker is paid a piece rate p per unit of defect-free output, and does not receive pecuniary punishment for producing defects. There is no daily minimum wage (as there is in the setting studied by GZN). The piece rate is invariant to air quality and does not vary across workstations, with workers performing the same parallel tasks on machines of similar vintage and quality.[17] Conditional on coming to work, the worker

[17]Where relevant, we adopt the notation in GZN). Notice that we could have followed that paper and specified ambient pollution as impacting output, rather than the cost of effort. Modeling the probability of job retention as increasing in the level of output would be both trivial (as in GZN) and unnecessary,

solves:

$$\arg\max_{e} pq(e) - c(e, \Omega^I)$$

and the first-order condition is simply:

$$p\frac{\partial q(e)}{\partial e} = \frac{\partial c(e, \Omega^I)}{\partial e}$$

We can further model the worker's choice of attending work, with the reservation utility ϕ being a possible function of *outdoor* air pollution Ω^O. For example, poor air quality may impact the health of the worker or of a family member who demands home care from the worker (e.g., Diette et al., 2000), or, alternatively, good air quality raises the value of outdoor leisure relative to work. The plant's indoor air is not purified or filtered, thus to the extent that fine particles in outdoor air penetrate into the indoor microenvironment, Ω^I and Ω^O will be highly correlated.[18] In addition to Ω^O, we posit that meteorological conditions Λ, such as temperature or precipitation, may drive selection into work attendance, since meteorology may directly impact own or family health, the relative value of leisure, or the cost of commuting to work. However, meteorology is less likely to directly affect worker productivity, given the sheltered work microenvironment that is temperature controlled both in winter and in summer—to the best of our knowledge, precipitation and extreme temperatures could damage the yarn, fabric and/or electricity operated machines.[19] This assumption provides an exclusion restric-

since employment contracts in our setting are *not* short-lived (in contrast to GZN's setting). Notice that in GZN, ξ is implicitly equal to 1, with one observable output, $q = \tilde{q}$. We could go further and, more richly, allow the worker to choose different types of effort, each differentially impacting gross quantity $\tilde{q}$ and production quality ξ.

[18]Like Chang et al (2014), the outdoor/indoor air pollutant we primarily have in mind is PM2.5. Whereas Chang et al assume that such fine particles in outdoor air penetrate indoors, we plan to conduct field measurements to ensure that this indeed is the case for the particular microenvironment we study. We emphasize that Chang et al's assumption (as well as ours thus far) finds support in the environmental sciences and engineering literature, e.g., Morawska et al. (2001), Cyrys et al. (2004), Gupta and Cheong (2007). For example, Cyrys et al (2004) find, for a given microenvironment they study, that with "closed windows, the I/O (indoor-to-outdoor) ratios for PM2.5 are ... 0.63 ... (and) that more than 75% of the daily indoor variation could be explained by the daily outdoor variation for those pollutants." It is also conceivable that the worker arrives to work already feeling unwell due to her exposure to pollution in the preceding hours, while commuting or at home. As we show, PM2.5 concentrations tend to be persistent over adjacent hours, but not necessarily across days (particularly under meteorological shocks such as strong wind or precipitation).

[19]Further, in principle one can use meteorology to instrument for pollution in the productivity equation,

tion to identify an overall selection-plus-productivity effect of pollution. Further, the exclusion of meteorology from the productivity equation, thanks to indoor temperature control, in principle allows one to better control for an otherwise potentially important confounder—ambient temperature—in any evaluation of the labor supply effects of pollution (e.g., Crocker and Horst 1981). The worker's problem is then:

$$\max \left\{ \phi(\Omega^O, \Lambda), \max_e pq(e) - c(e, \Omega^I) \right\} \tag{1}$$

with the worker choosing between an unplanned absence or non-attendance (i.e., excluding predetermined leave days, such as consecutive days of vacation) or attending work as planned, and, conditional on attending work, optimizing over the effort level. We note that one feature of this labor market, that the worker does not choose how many hours to work, conditional on showing up to work, implies that we need not model this additional margin of labor supply.

3.1 Empirical model

The empirical counterpart of a worker's problem can be written as

$$\ln q_{ijt} = \alpha_0 + \gamma \Omega^O_{jt} + X'_{ijt}\alpha_1 + \alpha_t + \alpha_j + \eta_i + \epsilon_{ijt}, \tag{2}$$

$$d^*_{ijt} = \beta_0 + \delta \Omega^O_{jt} + \tilde{X}'_{ijt}\beta_1 + \beta_t + \beta_j + \mu_i + \zeta_{ijt}; \quad d_{ijt} = 1[d^*_{ijt} > 0] \tag{3}$$

where 1[.] is an indicator function, which equals one if its argument is true, and zero otherwise. In terms of subscripts, q_{ijt} is worker i's defect-free output quantity ($q = \tilde{q}\xi$) observed during the 8-hour shift j on date t, which is only observable if $d_{ijt} = 1$. In the output equation, Ω^O_{jt} is the mean PM2.5 concentration (in μg/m$^3 \times 100$) contemporaneously recorded outdoors during the 8-hour work shift, assumed (for now) to correlate tightly with unobserved indoor PM2.5 concentrations, $\Omega^I_{jt} = \iota\Omega^O_{jt} + \omega_{jt}$, where ι captures the "indoor-outdoor ratio" and ω is classical measurement error. Other controls

to the extent that meteorology shifts fine particle pollution but does not affect productivity directly.

include: (i) a vector of individual characteristics, X_{ijt}, to capture observed heterogeneity, such as gender, age (or a quadratic in age), schooling, whether the worker suffers from asthma, and how long the worker has been employed at the firm; (ii) year-month and day-of-week fixed effects, α_t, to capture seasonality over the year and within the week; (iii) shift fixed effects, α_j; and (iv) worker fixed effects, η_i, to account for unobserved worker productivity. The random error term is denoted ϵ_{ijt}. Clearly, parameters α_1 cannot be identified if we control for worker fixed effects in the regression, but we can subsequently regress fitted worker fixed effects on observed individual characteristics to obtain mean effects. Other variations are possible such as combining controls (iii) and (iv) into worker-shift fixed effects, η_{ij}; replacing year-month fixed effects in α_t (15 months) by year-week fixed effects (65 weeks in the sample, which is currently being expanded as we collect more data); or even adding a quadratic trend to α_t. Another possibility is that the impact of fine particle pollution on labor productivity might be nonlinear. To address this possibility, we can include either a third-order polynomial function or a spline function of PM2.5 concentrations in the regression.

The selection equation contains all the control variables of the output equation (2) in addition to other variables that affect d^*_{ijt} but not q_{ijt}, such as meteorology. Presumably, these variables could affect a worker's reservation utility. However, since the work environment is temperature controlled, these meteorological variables should not affect workers' productivity. μ_i is an unobserved time invariant worker specific effect that affects a worker's attendance decision, and ζ_{ijt} is a random error term.

In a model without the worker fixed effects η_i, if the random error terms in the selection equation and the output equation are not correlated, we can estimate γ using pooled OLS. As in Dustmann and Rochina-Barrachina (2007), a sufficient condition to obtaining consistent estimate of γ using pooled OLS is:

$$E(\eta_i + \epsilon_{ijt} | \Omega^O_{jt}, X_{ijt}, \alpha_t, \alpha_j, d_{ijt} = 1) = 0, \forall t. \tag{4}$$

If workers with higher η_i are more or less likely to take leave when pollution is high, the OLS estimate for γ will be biased. For instance, PM2.5 might have a larger impact on

the health of a particular group of workers, such as workers who suffer from asthma, who are then more likely to call in sick during polluted days. If these workers are less (more) productive than other workers, then the estimate for γ will underestimate (overestimate) the true average effect.

We can then estimate a model with worker fixed effects to remove the potential bias caused by the selection on η_i. A sufficient condition for obtaining a consistent estimate of γ using a worker fixed-effect model is

$$E(\epsilon_{ijt} - \epsilon_{ijs} | \Omega^O_{jt}, X_{ijt}, \alpha_t, \Omega^O_{js}, X_{ijs}, \alpha_s, d_{ijt} = d_{ijs} = 1) = 0, \quad s \neq t, \tag{5}$$

where s and t are time periods. The above condition will be violated if workers who experience a positive productivity shock, i.e., higher ϵ_{ijt}, are more or less likely to take leave when pollution is high, the pooled OLS estimate will be biased as well. Wooldridge (1995) recommends deriving an expression for (4), and adding it as an additional regressor to output equation (2).

4 The impact of PM2.5 on non-attendance

We focus our analysis on the decision of starting a non-attendance spell. The reason for focusing on the first day of a non-attendance spell is that factors that trigger a worker not to work might differ from factors that keep her from coming back to work. The average duration of these spells is 1.7 days. Out of a total of 909 person non-attendance spells in our sample, 585 lasted for only one day, 179 for two days and 73 for more than three days. We exclude non-attendance spells that are longer than 3 days from our analysis, as factors that lead a worker to absent from work for several days might differ from these that keep her from working for only one or two days. Because a worker's decision on whether to come to work or not may more likely depend on the pollution level she is experiencing, we analyze the impact of the average concentration level in the past 24 hours prior to the start of the shift rather than the contemporaneous PM2.5

concentration.[20]

Table 2 reports the estimation results. In column (1), we only include the cubic function of the past 24-hour average PM2.5 concentration level in the regression. None of the coefficients are statistically significant even at at the 10% level. Since the PM2.5 concentration level generally follows a seasonal pattern and is affected by meteorological variables, the insignificance might be the result of failing to control for these variables. In column (2), we control for whether the daily temperature falls between 16^0C and 28^0C, whether the daily precipitation is higher than 5mm, workers' personal characteristics, year-month dummies, and day-of-week dummies. Adding these additional controls has little impact on our estimate although the coefficient on the linear term of the past 24-hour average PM2.5 concentration level becomes marginally significant (at the 10% level). Nevertheless, coefficients on several of variables are statistically significant. For instance, workers are more likely not to work when the daily temperature is in the range of 16 ^{0}C to 28^0C. Since outdoor activities are more attractive when the temperature is mild, the higher probability of not working on these days suggests that worker's non-attendance decision indeed depends on the option value of not working. Following a similar reasoning, when the daily precipitation is above 5mm, workers are also more likely to work. This observation is also partially driven by the fact that most workers either live in dorms or residential compound built by the factory, heavy rainfall does not have a large impact on their commuting cost. In column (3), we control for worker fixed effects, which raises the R^2 from 0.005 in column (2) from 0.045 in column (3). However, none of the coefficients is affected by the inclusion of worker fixed effects. This is likely due to the fact that the time invariant worker specific factors are unlikely to be correlated with either air pollution or meteorological variables, such as temperature or rainfall. In column (4), we control for the contemporaneous concentration level of SO_2 and O_3. Again, adding these controls makes little impact on the estimates, although the coefficient on the linear term of the past 24-hour average PM2.5 concentration level becomes statistically insignificant even at the 10% level. To check the overall impact of

[20]We have also tried to include the contemporaneous PM2.5 concentration in the regression. The coefficient on the contemporaneous PM2.5 value is small and never statistically significant.

the PM2.5 concentration on non-attendance, we have also tested the joint significance of the linear, quadratic and cubic terms for the past 24-hour average PM2.5 concentration level. Except for the regression where we only control for variables related to PM2.5, the joint impact of these three terms are only statistically significant at the 10% level, suggesting that the impact of air pollution on attendance is likely to be weak.

To further check whether the impact of PM2.5 concentration level on non-attendance varies across shifts or workers, we run separate regressions for each shift, and for *more-productive* and *less-productive* workers. *More-productive* workers are those whose average per shift productivity is above the median and *less-productive* workers are those whose average productivity is below the median.[21] The results show that the impact is strongest for the 0 am shift (from 0 am and 8 am), and weakest for the 8 am shift (from 8 am to 4 pm). Even for the 0 am shift, an $10\mu g/m^3$ increase in PM2.5 concentration only reduces the probability of non-attendance by 0.1 percentage points when evaluated at the median of $95\mu g/m^3$. We do not find any significant difference between *more-productive* and *less-productive* workers although the latter has a higher probability of not working. The similarity between the results for *more-* and *less-* productive workers suggests that mean of unobserved worker quality at any given shift is unlikely to be correlated with the PM2.5 concentration level, which suggests that the estimated impact of air pollution on labor productivity is unlikely to be sensitive to whether we control for the self-selection term in equation 2.

5 The impact of PM2.5 labor productivity

Table 4 reports the estimated impact of air quality on labor productivity. In column (1), we run a pooled OLS regression with the cubic function of the contemporaneous PM2.5 concentration level as the only control. The estimates suggest that a 10 μg/m^3 reduction in 8-hours average of PM2.5 concentration from the median level of 95 μg/m^3 will increase the number of fabric produced per worker per shift by about 3.5 meters (about 0.7%). As we argued in section 3, the estimates could be biased if more productive

[21] Because of the small sample size for non-local workers, we cannot run a separate regression for them.

workers are more or less likely to absent from work when the pollution level is high. In column (2), we control for worker characteristics, the shift, year-month fixed effects, and day-of-week fixed effects. Adding these controls only slightly changes the coefficients on PM2.5. When evaluated at the median pollution level, our estimates suggest that a 10 μg/m^3 reduction in the PM2.5 concentration will increase the number of fabric produced per worker per shift by about 3.9 meters (about 0.7%). Controlling for worker fixed-effect instead of worker characteristics make little difference. Adding the concentrations of SO_2 and O_3 almost has no effect on our estimates.

As shown in Figure 5, variations in the PM2.5 concentration level does not only affect the amount of output, but also the amount of defective output. Table 6 reports the impact of PM2.5 on the net output, defined as the total output minus the amount of defective output. Given the mean of defective output per worker per shift is only 3.8 meters or 0.7% of the mean gross output, our estimates are not sensitive to whether we the gross output or net output as the dependent variable. When we only control for the cubic function of the PM2.5 concentration level, the estimates suggests that a 10 μg/m^3 increase in PM2.5 concentration will reduce the net output per worker per shift by 3.9 meters when evaluated at the median, which increases to 4.2 meters after we control for worker fixed effects, shift, year-month fixed effects, day-of-week fixed effects, and co-pollutants. The sightly stronger effect compared with what have reported in Table 4 suggests that the negative effect of pollution on output quality strengthens its impact on product quantity. In our case, although air pollution has a significant negative effect on both the product quantity and quality, the main channel is via its effect on quantity.

Since the coefficient on the level of PM2.5 is always negative and on its squared term is always positive across all the specifications, our results suggests that the marginal effect first decreases as the PM2.5 level increases, which is inconsistent with Chang et al. (2014) who find that the marginal effect increases with PM2.5. To check whether the decreasing marginal effect is mostly driven by the function form restriction, we replace the cubic function with a spline function with the kink points set at the first, second and third quartile of the PM2.5 distribution. The results are reported in column (1)

of Table 7. The results show that a 10 μg/m^3 rise in PM2.5 concentration reduces the net output by 4.5 meters if the level is below 60 μg/m^3, the magnitude of the negative impact declines slightly to 3.2 meters if the PM2.5 is in the range of 60-95 μg/m^3, but then increases to 6.6 meters once the PM2.5 concentration level is in the range of 95-145 μg/m^3. Once the concentration level exceeds 145 μg/m^3, the impact declines again to 0.63 meters for every $10mug/m^3$ increase in the PM2.5 concentration level. These results suggest that the impact of air pollution on labor productivity is highly non-linear. As a result, we cannot simply extrapolate the findings from countries with relatively low pollution levels to highly polluted countries.

To check whether the impact of PM2.5 concentration on labor productivity differs across shifts, we run separate regressions for the three shifts and report the results in columns (2) to (4) of Table 7. The impact is nonlinear for all the shifts, but the relationship differs across shifts. The marginal effect peaks when PM2.5 concentration is in the range of 95-145 μg/m^3 for the 0 am and 8 am shifts, but at the 0-60 μg/m^3 range for the 4 pm shift. The peak values varies from 4.3 meters for every 10 μg/m^3 change in the PM2.5 concentration level for the 0 am shift to 7.7 meters for the 8 am shift. Nevertheless, once the PM2.5 concentration exceeds 145 μg/m^3, the marginal impact of air pollution varies little across shifts, from 0.6 meters for the 0 am shift to 0.7 meters for the 8 am shift. We also run separate regressions for *more* and *less productive* workers and report the results in columns (5) and (6). The results suggest that air pollution has a larger impact on *more productive* workers.

6 Conclusions

In this paper, we use the payroll information from a Chinese textile mill to analyze the impact of air pollution on labor productivity. Because our sample workers worked in an air conditioned factory, variations in other weather related variables, such as temperature or wind, have little impact on our estimation results. Our results show air pollution, measured by PM2.5 concentration level, affects non-attendance, labor productivity, and output quality. The relationship is far from linear. An increase in PM2.5 actually

reduces non-attendance when the pollution is at the moderate range. This is likely due to the fact that non-working activities become less attractive as the air pollution worsens. Nevertheless, once the pollution pass a certain threshold, a further increase in PM2.5 will indeed increase worker's probability of non-attendance. The impact does not sensitive to whether we control for worker's characteristics, suggesting that workers who absent from work in polluted days are comparable to these who absent in not so polluted days. Similar to Graff-Zivin and Neidell (2012) and Chang et al. (2014), we also find that air pollution has a significant negative impact on labor productivity. Because the variation in air pollution in our sample is much larger than these two previous studies, we can examine the impact of air pollution on labor productivity at a very high concentration level with relatively high precision. Our results show that the marginal impact is smaller when the PM2.5 is in the range of 60-95$\mu g/m^3$ than in the range of 0-65$\mu g/m^3$, but increases once the PM2.5 concentration falls into the 90-145$\mu g/m^3$ range, and decreases again once when the concentration level is above 145$\mu g/m^3$. Nevertheless, while these results show the nonlinear relationship between labour productivity and air pollution, the turning points are sensitive to the model specifications.

Another novel contribution of our study to the literature is that we can examine the impact of air pollution on both the gross output and quality adjusted output. Our results show that air pollution has a stronger negative impact on quality adjusted output than on the gross output. This suggests that the negative impact of air pollution on labor productivity goes beyond it impact on quantity. Because our sample workers worked in a air conditioned factory, our findings could potential be generalized to firms where the output of each worker is not easily quantifiable, such as office worker, or the quantity of output is mostly determined by the technology, such as assemble line workers.

References

Chang, T., Graff-Zivin, J. S., Gross, T., and Neidell, M. J. (2014). Particulate pollution and the productivity of pear packers. Working Paper 19944, National Bureau of Economic Research.

Table 1: Descriptive Statistics

Variables	(1) Mean	(2) S.D.	(3) Minimum	(4) Maximum	(5) N
Total output (meters per worker per shift)	509.3	96.14	128.3	1,137	27,065
Defective output (meters per worker per shift)	3.823	7.872	0	82	20,649
Proportion of absence	0.0604	0.238	0	1	28,805
First day of non-attendance	0.0325	0.177	0	1	27,980
Non-attendance duration (days)	1.736	1.564	1	16	909
Days since last absence	20.12	19.49	1	112	27,065
PM2.5 (8-hr ave. $100\mu g/m^3$)	1.242	0.986	0.030	7.731	1,177
Ozone (8-hr ave. $\mu g/m^3$)	0.567	0.522	0.010	4.279	1,155
PM10 (8-hr ave. $100\mu g/m^3$)	2.097	1.256	0.060	9.425	1,139
SO2 (8-hr ave. $\mu g/m^3$)	0.537	0.485	0.016	3.090	1,168
CO (8-hr ave. $\mu g/m^3$)	0.022	0.017	0.002	0.141	1,154
NO2 (8-hr ave. $\mu g/m^3$)	0.690	0.417	0.020	4.354	1,152
0 am shift	0.333	0.472	0	1	1,177
8 pm shift	0.335	0.472	0	1	1,177
Male worker	0.184	0.389	0	1	98
Age of worker	40.79	9.309	17	58	98
Years of employment at firm	18.52	12.07	1	35	98
Length of contract (years)	7.071	3.033	1	10	98
Worker's years of schooling	9.735	1.615	6	12	98
Local worker (Hebei)	0.806	0.397	0	1	98
Ave. daily temperature (outdoor,^{0}C)	18.44	10.13	-5.250	35.25	403
Daily precipitation (mm)	0.941	4.642	0	69.60	403
Ave. daily wind speed (outdoor, mph)	6.373	2.355	2	13.25	403
Ave. daily humidity, (outdoor)	0.454	0.166	0.106	0.894	403
Temperature between 16 and 28^{0}C	0.114	0.318	0	1	403
Daily precipitation > 5mm	0.0447	0.207	0	1	403

Table 2: The impact of PM2.5 on non-attendance

	(1)	(2)	(3)	(4)
Past 24-hr ave. PM2.5 $100\mu g/m^3$	-0.011	-0.016*	-0.016*	-0.015
	(0.009)	(0.009)	(0.009)	(0.009)
Lagged PM2.5 squared	0.005	0.006	0.006	0.005
	(0.004)	(0.004)	(0.004)	(0.004)
Lagged PM2.5 cubic	-0	-0.001	-0.001	-0
	(0)	(0)	(0)	(0)
Temperature between 16 and 28^0C		0.016**	0.016**	0.014*
		(0.008)	(0.008)	(0.008)
Precipitation > 5mm		-0.014**	-0.014**	-0.016**
		(0.007)	(0.007)	(0.007)
Ave. daily wind speed (outdoor,mph)		0	0	0
		(0.001)	(0.001)	(0.001)
0 am shift		0.002	0.002	-0.001
		(0.003)	(0.003)	(0.003)
8 am shift		-0	0	-0.001
		(0.003)	(0.003)	(0.003)
Worker's age		0*		
		(0)		
Male		-0.008***		
		(0.002)		
Years of exp. in the factory		0		
		(0)		
Length of contract (years)		-0.001**		
		(0.001)		
Worker's years of schooling		-0.001*		
		(0.001)		
Local worker		-0.015***		
		(0.003)		
SO2 (8-hr ave. $\mu g/m^3$)				-0
				(0.003)
Ozone (8-hr ave. $\mu g/m^3$)				-0.006
				(0.004)
Constant	0.036***	0.058***	0.044***	0.046***
	(0.005)	(0.013)	(0.010)	(0.010)
Observations	27,143	27,143	27,143	26,665
R-squared	0	0.005	0.041	0.040
Month and year effect	No	Yes	Yes	Yes
Day-of-week effect	No	Yes	Yes	Yes
Worker fixed-effects	No	No	Yes	Yes
Mean value of dependent variable	0.031	0.031	0.031	0.031

Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%.

Table 3: The impact of PM2.5 on non-attendance, by group

	Winter	Summer	0 am Shift	8 am Shift	4 pm Shift	More Productive	Less Productive
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Past 24-hr ave. PM2.5 $100\mu g/m^3$	-0.021*	0.053	-0.036**	-0.020	0.009	-0.017	-0.015
	(0.012)	(0.057)	(0.016)	(0.013)	(0.015)	(0.011)	(0.013)
Lagged PM2.5 squared	0.008*	-0.054	0.015**	0.007	-0.005	0.005	0.008
	(0.005)	(0.052)	(0.007)	(0.005)	(0.007)	(0.005)	(0.006)
Lagged PM2.5 cubic	-0.001	0.014	-0.001*	-0.001	0.001	-0	-0.001
	(0)	(0.014)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Temperature between 16 and 28^0C	0.002	0.016**	0.012	0.015*	0.020*	0.015**	0.017
	(0.006)	(0.008)	(0.011)	(0.008)	(0.012)	(0.008)	(0.011)
Precipitation > 5mm	-0.029***	-0.015*	-0.013	-0.016	-0.015*	-0.009	-0.019**
	(0.007)	(0.008)	(0.010)	(0.010)	(0.009)	(0.008)	(0.009)
Ave. daily wind speed (outdoor,mph)	0.002*	-0.001	-0	0	0	0.001	0
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
00-07 hr shift	0.001	0.002				0	0.003
	(0.004)	(0.003)				(0.003)	(0.004)
08-15 hr shift	0.002	-0.001				0.001	-0.001
	(0.004)	(0.004)				(0.003)	(0.005)
Constant	0.039***	0.029*	0.059***	0.039***	0.036**	0.040***	0.048***
	(0.012)	(0.015)	(0.015)	(0.014)	(0.016)	(0.012)	(0.013)
Observations	11,053	16,090	9,025	9,019	9,099	13,690	13,453
R-squared	0.067	0.038	0.052	0.053	0.046	0.032	0.051
Month and year effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%. *More productive* workers refer to those whose average productivity is above the median and *less productive* workers refer to those whose average productivity is below the median.

Table 4: The impact of PM2.5 on output

	(1)	(2)	(3)	(4)
PM2.5 (8-hr ave. $100\mu g/m^3$)	-52.871***	-61.430***	-61.194***	-61.467***
	(6.834)	(6.082)	(5.846)	(5.923)
PM2.5 squared	9.993***	12.820***	12.734***	12.861***
	(2.370)	(2.028)	(2.016)	(2.021)
PM2.5 cubic	-0.573***	-0.800***	-0.802***	-0.816***
	(0.220)	(0.192)	(0.200)	(0.201)
0 am shift		7.923***	7.555***	7.166***
		(2.758)	(2.096)	(2.245)
8 am shift		5.126**	4.866***	4.488**
		(2.166)	(1.768)	(1.815)
First day of the shift		0.746	0.606	0.009
		(2.155)	(2.151)	(2.102)
First day of work after non-attendance		-7.190	-5.927	-5.756
		(4.542)	(4.243)	(4.297)
Last day of work before non-attendance		11.163**	13.801***	14.592***
		(4.818)	(4.289)	(4.285)
First day of work after holiday		5.018	4.353	4.963
		(5.295)	(4.991)	(4.973)
Last day of work before holiday		-5.386	-7.152	-7.255
		(5.190)	(5.798)	(5.838)
Worker's age		0.286***		
		(0.099)		
Male		0.396		
		(1.157)		
Years of exp. in the factory		-0.233***		
		(0.081)		
Length of contract (years)		1.718***		
		(0.238)		
Worker's years of schooling		-1.694***		
		(0.343)		
Local worker		9.399***		
		(1.509)		
SO2 (8-hr ave. $\mu g/m^3$)				0.418
				(2.451)
Ozone (8-hr ave. $\mu g/m^3$)				0.211
				(2.541)
Constant	554.553***	556.099***	566.295***	566.416***
	(4.714)	(8.112)	(6.392)	(6.685)
Observations	27,065	27,065	27,065	26,587
R-squared	0.046	0.099	0.275	0.274
Month and year effect	No	Yes	Yes	Yes
Day-of-week effect	No	Yes	Yes	Yes
Worker fixed-effects	No	No	Yes	Yes
Mean value of dependent variable	509.3	509.3	509.3	509.3

Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%.

Table 5: The impact of PM2.5 on output, by season and shift

	Winter	Summer	0 am Shift	8 am Shift	4 pm Shift	More Productive	Less Productive
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PM2.5 (8-hr ave. $100\mu g/m^3$)	-74.784***	-9.421	-63.027***	-64.801***	-61.310***	-75.935***	-44.395***
	(11.469)	(12.832)	(9.554)	(7.421)	(8.138)	(6.837)	(6.523)
PM2.5 squared	16.732***	-28.803***	14.975***	12.368***	13.707***	16.997***	7.536***
	(3.674)	(10.324)	(3.439)	(2.488)	(3.195)	(2.396)	(2.260)
PM2.5 cubic	-1.132***	9.104***	-1.125***	-0.702***	-0.952***	-1.170***	-0.325
	(0.342)	(2.277)	(0.338)	(0.226)	(0.358)	(0.234)	(0.222)
0 am shift	7.893**	7.332***				6.275**	9.218***
	(3.556)	(2.519)				(2.463)	(2.576)
8 am shift	5.956**	3.780*				2.545	7.608***
	(2.732)	(2.237)				(2.516)	(2.222)
First day of the shift	-0.183	0.874	4.685	1.590	-4.118	-1.824	3.120
	(3.573)	(2.467)	(3.135)	(3.168)	(3.076)	(2.599)	(2.380)
First day of work after non-attendance	-5.190	-6.603	-8.215	-10.985	0.920	-9.227*	-3.210
	(6.244)	(5.401)	(5.752)	(7.639)	(6.306)	(4.900)	(5.611)
Last day of work before non-attendance	20.163***	8.171	15.436**	10.821*	16.650**	3.151	23.311***
	(5.469)	(6.151)	(6.132)	(6.264)	(6.773)	(4.922)	(5.745)
First day of work after holiday	5.740	5.413	12.031	2.399	2.695	2.074	6.763
	(8.171)	(5.218)	(9.596)	(8.336)	(8.036)	(4.969)	(5.953)
Last day of work before holiday	-13.792***	-12.657**	-12.302*	-10.056	4.865	-5.152	-9.219
	(4.221)	(6.410)	(7.188)	(8.827)	(8.949)	(5.884)	(7.404)
Constant	578.707***	551.497***	571.253***	582.649***	560.174***	616.837***	513.955***
	(11.213)	(5.932)	(10.593)	(8.345)	(8.333)	(7.342)	(7.098)
Observations	11,072	15,993	8,964	9,091	9,010	13,671	13,394
R-squared	0.283	0.284	0.280	0.280	0.299	0.192	0.174
Month and year effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean value of dependent variable	503.1	513.6	509.6	509.0	509.4	542.1	475.8

Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%.

Table 6: The impact of PM2.5 on net output

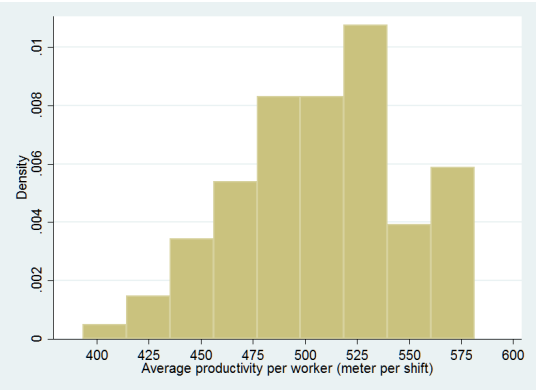
	(1)	(2)	(3)	(4)
PM2.5 (8-hr ave. $100\mu g/m^3$)	-57.437***	-65.328***	-64.916***	-65.645***
	(7.003)	(5.869)	(5.539)	(5.631)
PM2.5 squared	10.710***	13.387***	13.296***	13.464***
	(2.497)	(2.053)	(1.984)	(1.989)
PM2.5 cubic	-0.605**	-0.824***	-0.835***	-0.847***
	(0.235)	(0.200)	(0.202)	(0.203)
0 am shift		2.942	2.454	1.942
		(2.753)	(2.206)	(2.434)
8 am shift		5.679**	5.142**	4.756**
		(2.573)	(2.020)	(2.088)
First day of the shift		0.802	0.566	0.284
		(2.421)	(2.420)	(2.439)
First day of work after non-attendance		-12.253**	-11.589**	-11.672**
		(5.331)	(4.866)	(4.937)
Last day of work before non-attendance		8.664	10.726*	11.672**
		(6.328)	(5.726)	(5.754)
First day of work after holiday		10.304	9.983	10.224*
		(6.257)	(6.109)	(6.191)
Last day of work before holiday		-3.094	-4.662	-4.712
		(6.902)	(7.286)	(7.320)
Worker's age		0.290***		
		(0.109)		
Male		6.894***		
		(1.332)		
Years of exp. in the factory		-0.728***		
		(0.091)		
Length of contract (years)		3.016***		
		(0.280)		
Worker's years of schooling		-0.852**		
		(0.377)		
Local worker		13.731***		
		(1.759)		
SO2 (8-hr ave. $\mu g/m^3$)				0.234
				(2.918)
Ozone (8-hr ave. $\mu g/m^3$)				-0.177
				(2.835)
Constant	549.756***	537.755***	562.826***	563.890***
	(4.768)	(8.148)	(6.283)	(6.598)
Observations	20,649	20,649	20,649	20,305
R-squared	0.058	0.124	0.309	0.310
Month and year effect	No	Yes	Yes	Yes
Day-of-week effect	No	Yes	Yes	Yes
Worker fixed-effects	No	No	Yes	Yes
Mean value of dependent variable	504.9	504.9	504.9	504.8

Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%.

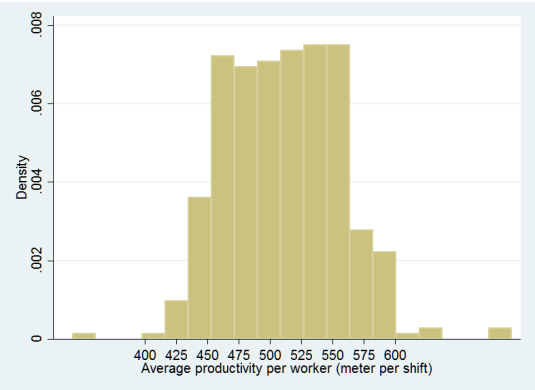
Table 7: The impact of PM2.5 on net output

	Entire Sample (1)	4 pm Shift (2)	0 am Shift (3)	8 am Shift (4)	More Productive (5)	Less Productive (6)
PM2.5 0-0.60	-45.001***	-48.860***	-49.318*	-21.829	-43.002***	-43.763***
	(11.984)	(16.467)	(26.889)	(23.327)	(15.442)	(14.795)
PM2.5 0.60-0.95	-31.601***	-45.704**	-0.759	-57.812***	-48.975***	-15.468
	(11.273)	(18.645)	(21.448)	(20.198)	(12.498)	(13.807)
PM2.5 0.95-1.45	-65.699***	-42.528***	-67.052***	-76.784***	-75.002***	-57.248***
	(8.031)	(14.143)	(12.206)	(11.532)	(8.619)	(9.703)
PM2.5 1.45-	-6.311***	-5.803*	-5.598	-7.181***	-6.869***	-6.010**
	(2.105)	(3.111)	(3.476)	(2.377)	(2.032)	(2.645)
0 am shift	2.399				-2.647	7.091**
	(2.201)				(2.544)	(2.815)
8 am shift	4.667**				0.595	8.749***
	(2.026)				(2.867)	(2.454)
First day of the shift	0.441	-4.734	3.480	3.339	-1.782	2.576
	(2.423)	(3.791)	(3.244)	(3.738)	(2.573)	(2.834)
First day of work after non-attendance	-11.518**	-2.723	-14.691**	-19.678**	-14.266**	-9.188
	(4.869)	(7.710)	(6.790)	(8.438)	(6.283)	(6.468)
Last day of work before non-attendance	10.546*	13.951	14.989*	3.621	0.711	18.572***
	(5.698)	(8.700)	(7.627)	(8.057)	(6.863)	(7.010)
First day of work after holiday	10.388*	24.421***	11.662	7.279	11.691**	9.685
	(5.880)	(7.789)	(8.942)	(7.176)	(5.431)	(7.075)
Last day of work before holiday	-6.111	13.923	-20.262***	-4.002	-8.606	-3.534
	(6.542)	(8.859)	(6.730)	(11.074)	(7.135)	(8.161)
Constant	553.255***	548.171***	549.956***	558.656***	606.486***	503.353***
	(7.397)	(11.128)	(13.298)	(12.425)	(9.358)	(9.185)
Observations	20,649	6,862	6,829	6,958	9,878	10,771
R-squared	0.308	0.343	0.306	0.318	0.220	0.195
Month and year effect	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week effect	Yes	Yes	Yes	Yes	Yes	Yes
Worker fixed-effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean value of dependent variable	501.0	502.1	498.9	504.8		

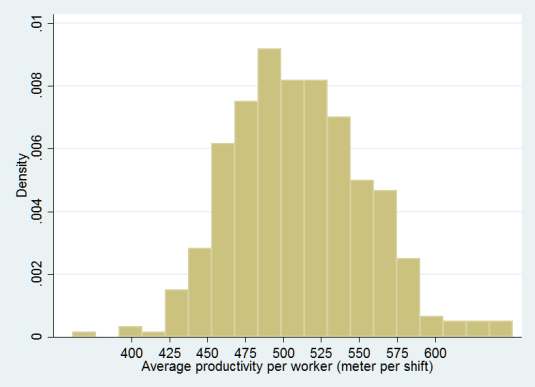
Note: Numbers in parentheses are standard errors, clustered by date. *** significant at 1% ** significant at 5% * significant at 10%.



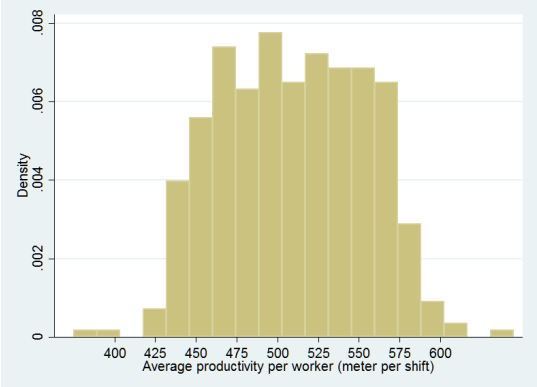
(a) Mean within-worker output per shift (over all days)



(b) Mean within-day output per worker per shift, 0 am shift

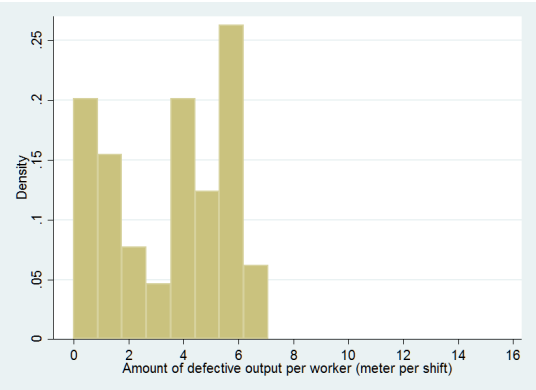


(c) Mean within-day output per worker per shift, 8 am shift

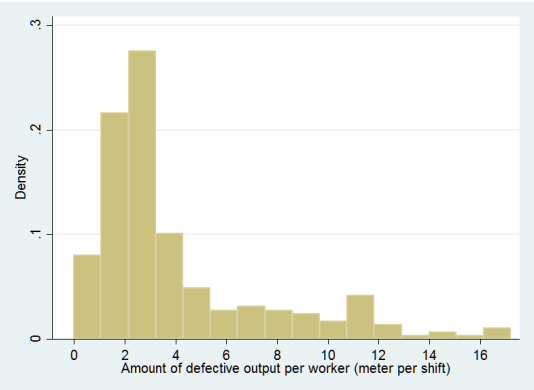


(d) Mean within-day output per worker per shift, 4 pm shift

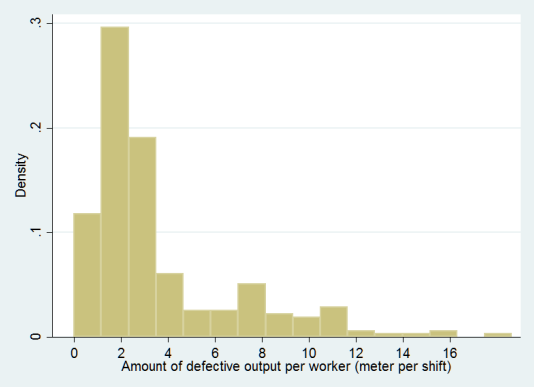
Figure 1: Distribution of mean output across workers and days



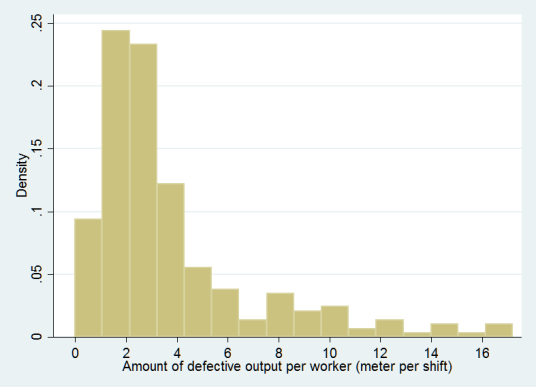
(a) Mean within-worker defective output per shift (over all days)



(b) Mean within-day defective output per worker per shift, 0 am shift



(c) Mean within-day defective output per worker per shift, 8 am shift



(d) Mean within-day defective output per worker per shift, 4 pm shift

Figure 2: Distribution of mean defective output across workers and days

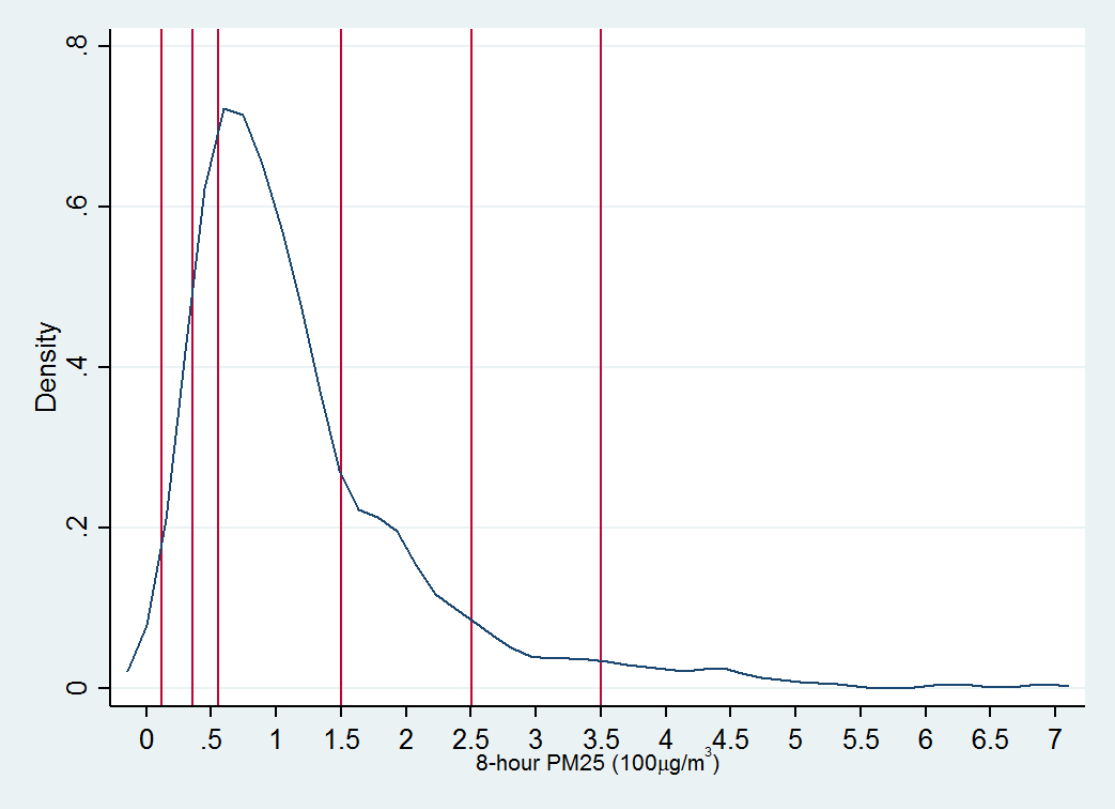
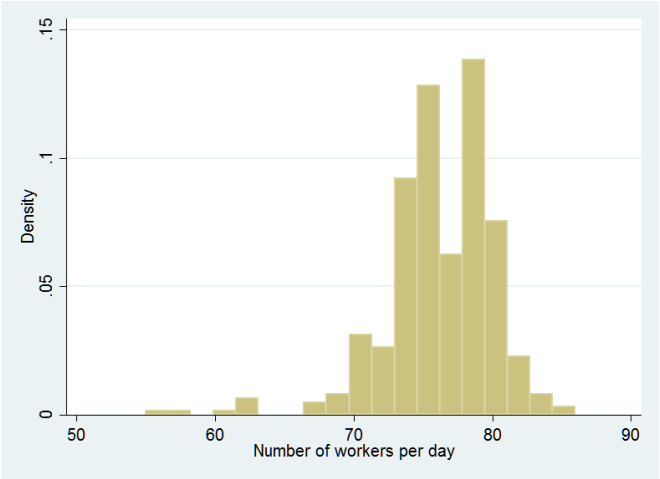
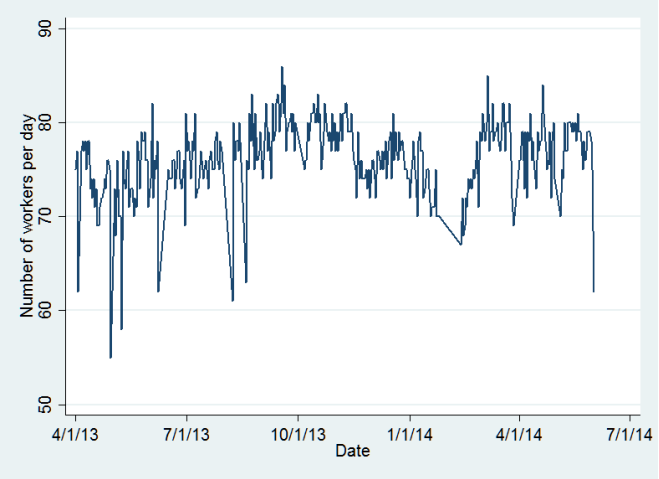


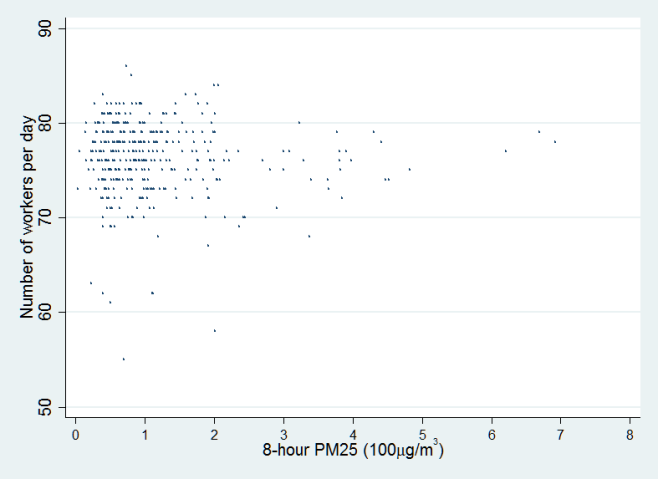
Figure 3: Density function of 8-hour PM2.5 concentration



(a) Variation in the number of workers per day

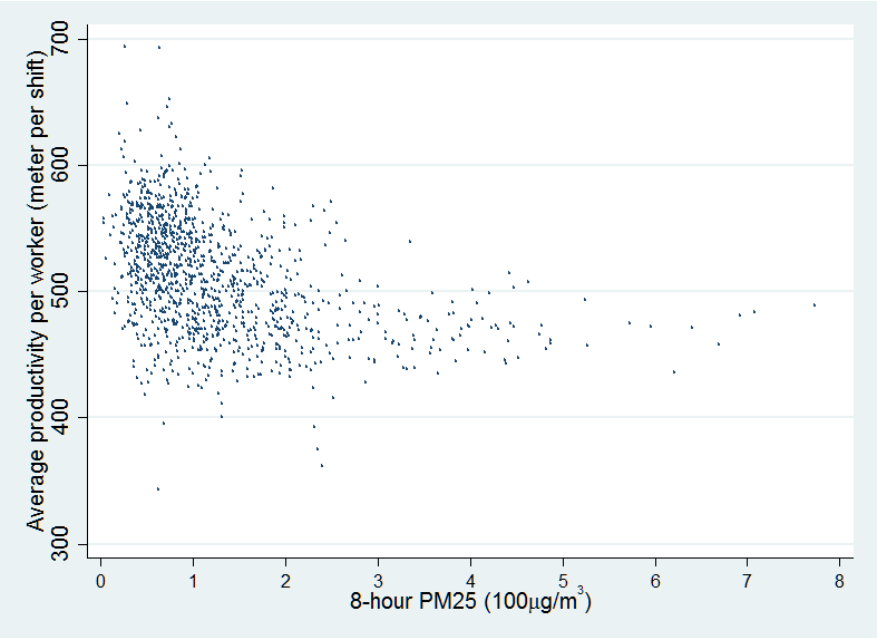


(b) Number of workers, by date

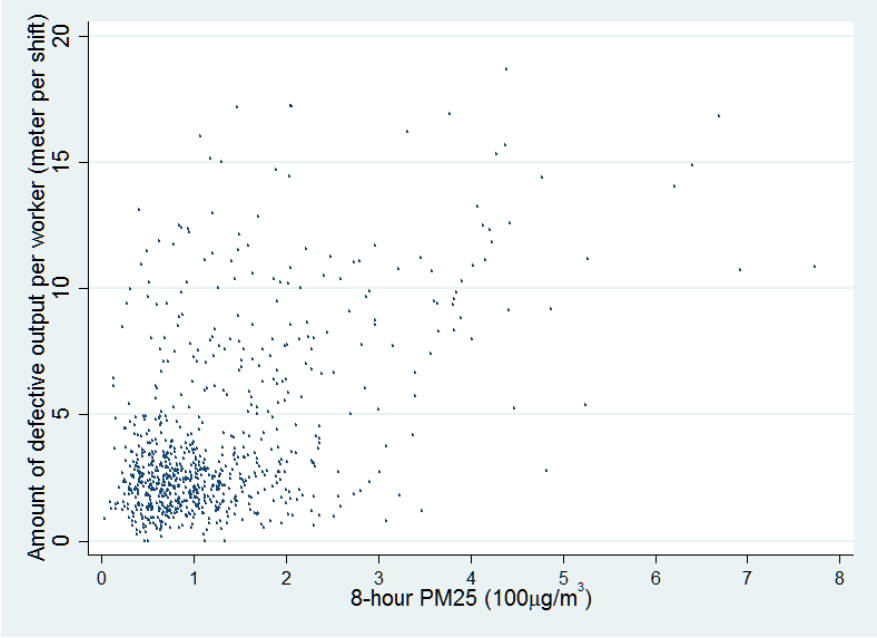


(c) Number of workers, by PM2.5

Figure 4: Number of workers per day



(a) average gross output per shift



(b) Average net output per shift

Figure 5: PM2.5 concentration ($100\mu g/m^3$), output and defective output

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from Chinas Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941.

Crocker, T. D. and Horst, R. L., J. (1981). Hours of work, labor productivity, and environmental conditions: A case study. *The Review of Economics and Statistics*, 63(3):pp. 361–368.

Cyrys, J., Pitz, M., Bischof, W., Wichmann, H.-E., and Heinrich, J. (2004). Relationship between indoor and outdoor levels of fine particle mass, particle number concentrations and black smoke under different ventilation conditions. *Journal of Exposure Science and Environmental Epidemiology*, 14(4):275–283.

Diette, G. B., Markson, L., Skinner, E. A., Nguyen, T. T., Algatt-Bergstrom, P., and Wu, A. W. (2000). Nocturnal asthma in children affects school attendance, school performance, and parents' work attendance. *Archives of Pediatrics & Adolescent Medicine*, 154(9):923–928.

Dustmann, C. and Rochina-Barrachina, M. E. (2007). Selection correction in panel data models: An application to the estimation of females' wage equations. *The Econometrics Journal*, 10(2):263–293.

Ebenstein, A., Fan, M., Greenstone, M., He, G., and Zhou, M. (2014). Growth, pollution, and life expectancy: China from 1991-2012.

Fann, N., Lamson, A. D., Anenberg, S. C., Wesson, K., Risley, D., and Hubbell, B. J. (2012). Estimating the national public health burden associated with exposure to ambient PM2.5 and ozone. *Risk Analysis*, 32(1):81–95.

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

Graff-Zivin, J. and Neidell, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, 321:1–26.

Gupta, A. and Cheong, K. (2007). Physical characterization of particulate matter and ambient meteorological parameters at different indoor–outdoor locations in Singapore. *Building and Environment*, 42(1):237–245.

Hanna, R. and Oliva, P. (2014). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico city. *Journal of public economics*, (forthcoming).

Morawska, L., He, C., Hitchins, J., Gilbert, D., and Parappukkaran, S. (2001). The relationship between indoor and outdoor airborne particles in the residential environment. *Atmospheric Environment*, 35(20):3463–3473.

Ostro, B. D. (1983). The effects of air pollution on work loss and morbidity. *Journal of Environmental Economics and Management*, 10:371–382.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, 68(1):115–132.