The relationship between pollution level and worker productivity*

Jiayi Du, Kaiyuan Yan

Abstract

In recent years, With the rapid development of China's economy, there is also a big environmental pollution problem, especially with the explosive growth of cars on the road, air pollution has become a very intractable problem. We conduct study on the effect of air pollution on the worker productivity in the service sector by focusing on two call centers in China. We measure the worker productivity by precise daily output of each worker and finally find out that higher levels of air pollution decrease worker productivity. These results have certain reference significance for other work in other industries throughout the developing and developed world. Besides, it can also arouse people's attention to protecting the environment.

Contents

*Code and data are available at: https://github.com/dujiayi1/The-relationship-between-pollution-level-and-worker-productivity and the DOI is https://doi.org/10.48152/ssrp-kzfg-5z15

1 Introduction

1.1 Introduction to the effect on worker productivity

Increasing evidence indicates that pollution can reduce the productivity of workers in physically demanding occupations. However, nowadays most of the modern economic output is from the performance of the service and knowledge sectors, which results in the modest importance of the physically demanding occupations. Thus, it's quite challenging to judge whether pollution has an impact on productivity in higher skilled, cognitively demanding professions. In this report, we examine the effect of pollution on workers whose primary tasks require cognitive rather than physical performance.

In particular, we analyze the effect of pollution on call center workers at a firm with offices in Shanghai and Nantong, China.

Since the the workers in our sample perform "desk jobs" that require minimal physical effort, the way we assess the extent to which pollution affects productivity also differs. Besides, the working environment of these workers is indoors and the indoor climate is basically constant, so that it suggests that the effect is caused by particulate matter (PM) pollution largely. Exposure to PM pollution can potentially impair productivity through changes in cardiovascular and lung functioning (Seaton et al. 1995); irritation of the ear, nose, throat, and lungs (Pope 2000); as well as through direct impacts on cognitive performance (Ebenstein, Lavy, and Roth 2014).

What's more, China has gone through a dramatic growth in economy as well as dramatic increase in pollution during the past few decades. Under different kinds of indicators, the quality of China's environment is almost the worst all over the world. And the impact on the worker productivity may imply that the environmental problems caused by the disorderly expansion of China's industries in the past decades may affect the established economic development goals and the economic growth rate.

In this report, in the section of Data, we choose the dataset from a company. The company we study is the Ctrip, China's largest travel agency. Several aspects of the company's operations ensure the causal effect of pollution on the marginal product of labor. First, the workers have little discretion over their labor supply, which means that we can estimate the effect on the productivity without much bias. Second, daily change in pollution in Shanghai and Nantong is likely unrelated to the firm's output. The company's customers come from all over the country, so the demand for service on any given day is highly unlikely to correlate with the pollution levels on any given day. Finally, a pilot program carried by Ctrip reminds us of another potential variable: traffic. Traffic may have huge influence on the environment, create emotional stress and make employees late for work, which may result in decrease in work productivity.

In the section of result, our analysis of the dataset points out a statistically significant, negative impact of pollution on the productivity of workers at the firm. Moreover, productivity declines are largely linearly increasing in pollution levels.

1.2 Background on Pollution

Like most of the countries, China also releases API every day, which is a composite measure of pollution that ranks air quality based on its associated health risks as a means to facilitate comprehensibility by the public. The higher the API , the worst air quality is. Generally, when the value is less than 100, it is acceptable. When the value is between 100 and 150, the air is unhealthy for sensitive groups. And impacts on the general population begin to emerge at levels greater than 150.

While API is calculated by determining the maximum value for a nonlinear transformation of three separate pollutants, particulate matter is almost always the driver of API in our sample. More specifically, for the subset of data for which the driving pollutant is reported, particulate matter is the determining pollutant 99.2 percent of the time, and 100 percent of the time for days with an API greater than 100.

The API in China converts concentrations of three criteria air pollutants into a single index. The pollutant that has the highest index which is called the primary pollutant, determines the API on a given day.

In our dataset, for API levels of greater than 50, particulate matter smaller than 10 $\mu g/m^3$ in diameter (PM_{10}) is the primary pollutant over 99 percent of the time.

While the more dramatic health effects due to pollution exposure may lead to changes in labor supply, milder impairment of respiratory, cardiovascular, and cognitive function may reduce productivity on the intensive margin. Focus, concentration, and critical thinking are all essential components of office-based job performance and depend heavily on a well-functioning brain and CNS. The goal of our analysis is to estimate the effect of pollution on the marginal product of labor, independent of any possible effects on the extensive margin of labor supply. Given the ubiquity of office work and the value it adds to global economic output, the welfare implications of any link between pollution and productivity in this setting are potentially enormous.

2 Data

I used R (R Core Team 2019), the tidyverse package (Wickham et al. 2019), the ggplot2 package (Wickham 2016),the haven package (Wickham 2021) , and the readxl package(Wickham 2019) to analyze the data and make plots and tables.

2.1 The Source of Data

Our data about worker productivity come from Ctrip, largest travel agency in China, which is available at https://www.openicpsr.org/openicpsr/project/113706/version/V1/view?pat h=/openicpsr/113706/fcr:versions/V1/data&type=folder. Ctrip's main business is arranging trips for its customers. Its main profit comes from commissions from hotels, airlines, and tour operators. In contrast to US and European-based agencies that operate in markets with deep internet penetration, Ctrip conducted much of its business on the telephone during our study period. The firm was listed on the NASDAQ in 2003 and its market capitalization was more than $17 billion by now.

2.2 Data Overview

Our analysis focuses on Ctrip's call center workers in Shanghai and Nantong, who book travel for clients throughout China. The offices are mainly located in large climate controlled buildings, and are filled with cubicles and modern telecommunications and computing hardware. Equipment and staffing practices are identical at both offices, and both sets of workers follow the same procedural guidelines. A central server automatically routes customer calls and assignments to workers logged into the system, based on a computerized call-queuing system. We focus on three aspects that show workers' productivity, respectively are: number of phone calls handled per shift, number of minutes spent on the phone, and number of minutes logged in to the call center's computer system (i.e., the number of minutes within the workday that workers are available to handle calls).

The data span from January 1, 2010 to December 9, 2012.

In addition, during the study period, Ctrip carried out a controlled experiment to analyze whether working from home affected worker performance. About 250 employees from the Shanghai office were asked to work from home. The experiment spanned from December 6, 2010 until August 14, 2011, during which time the workers who worked from home were provided with the necessary equipment to allow them to perform duties of their usual work.

More details on call center operations and the work-from-home experiment can be found in Bloom et al. (2015).

As to daily pollution data, they are obtained from the China National Environmental Monitoring Center (CNEMC). These data provide a measure of pollution that is based on the average API score across all monitors within a city. During the study period there were ten operating national pollution monitors in Shanghai and five in Nantong.

2.3 Data Visualization

As is shown in Figure 1, the pollution in both Shanghai and Nantong are continuously distributed and the API values are between 30 and 70 in most of the time.

```
## Rows: 1196 Columns: 3

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): cityname, date
## dbl (1): API
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.

## Warning: Removed 2 rows containing missing values (geom_bar).
```

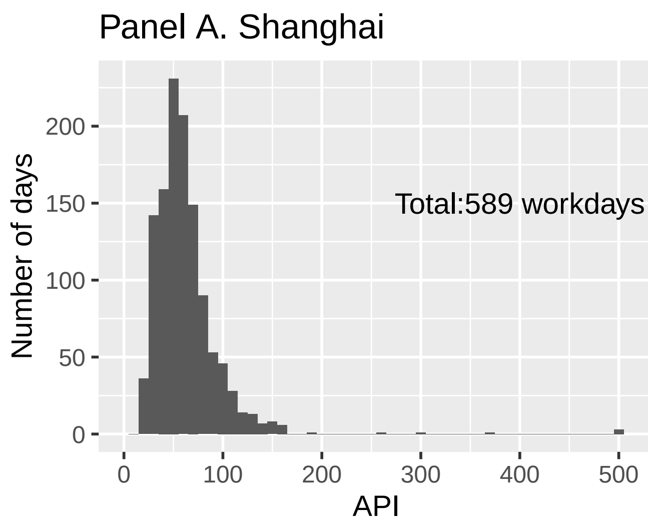


Figure 1: Histograms of Pollution

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```

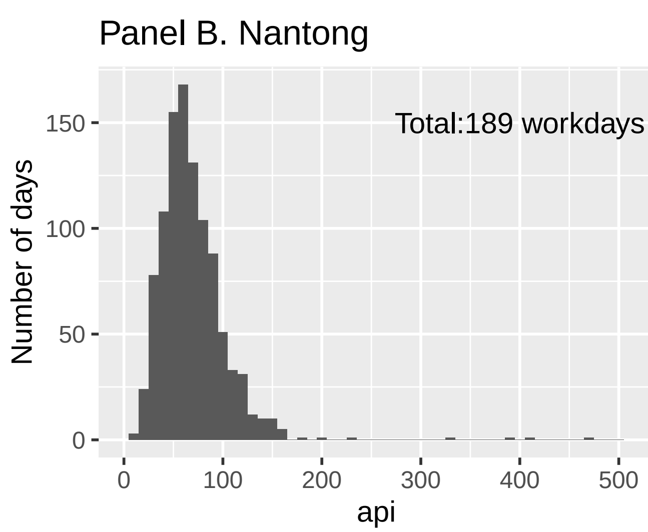


Figure 2: Histograms of Pollution

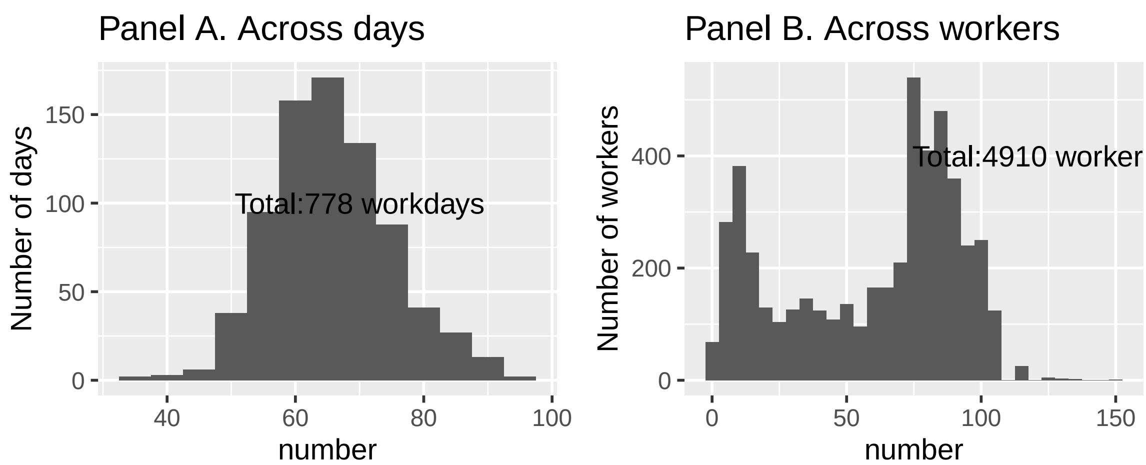
Table 1 presents simple descriptive statistics for the data. The average API is 66, while the average value in Shanghai is 65 and it's 69 in Nantong. The dataset includes the observations of 393 workers from Shanghai and 4499 workers from Nantong. The Shanghai sample is the subset of workers who took part in the work-from-home experiment from January 1, 2010 to August 14, 2011, while the Nantong sample includes all worker observations from September 1, 2011 to December 9, 2012.

Table 1: Sample Statistics

Variable	Overall	Shanghai	Nantong
Panel A. Environmental data			
Air pollution index (API)	65.9	62.8	69.6
(36.1)	(36.1)	(35.8)	
Temperature	16.3	17.4	15.1
(8.5)	(7.7)	(9.1)	
API by bin:			
API 0–50	36.8	37	36.3
—	(8.3)	(8)	(8.8)
API 50–100	68.6	67.2	70
—	(13.6)	(13.2)	(14)
API 100-150	118	117.3	118.6
—	(13.1)	(13)	(13.3)
API 150–200	159.8	160.2	159.3
—	(9.9)	(11.2)	(8.8)
API 200+	404.3	404.3	—
—	(110.8)	(110.8)	—
Panel B. Worker-Day level data			
Number of worker-days	359,013	112,678	246,335
Calls	70.9	64.1	74.0
—	(45.1)	(47.0)	(43.9)
Logged-in minutes	413.3	441.5	400.4
—	(167.6)	(164.5)	(167.4)
Minutes on phone	222.5	182.1	241.0
—	(144.8)	(142.2)	(142.2)
Minutes per call	3.0	2.4	3.2
—	(1.2)	(1.1)	(1.2)

Notes: The temperature is in degree centigrade. The numbers in parentheses are standard deviations.

Figure 2 displays the distribution of worker productivity. Panel A shows the distribution of calls across day, which appears to be normally distributed. Our analysis of the data set found that worker made 66 phone calls a day on average. Panel B describes the average productivity across workers. There are two peaks in the figure.



3 Model

3.1 Original model

According to the factors that affect the worker productivity, we first consider the linear regression model:

$$y = \beta * API + \epsilon$$

where y is the worker productivity, which we are interested in. And β is the effect of API, ϵ is the error term.

3.2 Final model

Our goal is to estimate the effect of pollution on the worker productivity. And then we consider the following hybrid production function:

$$log(y_{ijt}) = \beta * API_{jt} + X_{jt}^{'}\gamma + \delta_j + \alpha_i + \epsilon_{ijt}$$

Among them, y_{ijt} is the daily worker productivity, which is also the response variable. It's measured by either total number of calls per shift or the number of minutes logged in to the call center's computer system. API is a daily average measure of the air pollution index in city j, and β represents the effect of API on our two measures of productivity. The vector X_t includes temperature. The parameter δ includes day-of-week and year-month indicator variables to account for trends within the week and over time. The variable α_i is a worker-specific fixed effect for those specifications that rely only on within-worker variation in pollution exposure to identify impacts on productivity. In addition,ϵ is the error term.

There are several challenges to uncovering a causal effect of pollution on worker productivity. First, individuals can sort into locations based on the amount of pollution in that area, leading to nonrandom assignment of pollution. Moreover, within a location, an individual's labor supply might change in response to pollution.We explicitly test this claim by examining whether API relates to the probability of working and the number of hours worked (using the regression equation above) and find no effect of API levels on either measure of labor supply.

4 Result

Table 2 below shows the result of the model that estimates the effect of pollution on worker productivity. From the table we can conclude that a 10-unit increase in API decreases the number of calls per shift by 0.35 percent. It's quite significant that the worker productivity decreases at the 1 percent level.

Table 2: Effect of Pollution on Call Center Productivity

Dependent Variable	Number of calls
API/10	-0.0035
—	[0.0013]
Temperature	0.0008
—	0.0030
Temperature squared/1,000	-0.0029
—	[0.0230]
R^2	0.136
Observations	359,013

Table 3 presents estimate of the same regression, but considering minutes logged in per shift as the dependent variable. And the result is similar to the effect on the number of phone calls. We find that minutes logged in per shift decreases as the pollution increases. Table 3 implies that a 10-unit increase in API decreases the minutes logged in per shift by 0.25 percent.

Table 3: Effect of Pollution on Call Center Productivity

Dependent Variable	Minutes logged in
API/10	-0.0025
—	[0.0010]
Temperature	0.0025
—	0.0020
Temperature squared/1,000	-0.0124
—	[0.0150]
R^2	0.087
Observations	359,013

Notes: The logarithm of the indicated dependent variable is the outcome of interest. Standard errors in brackets are clustered on date and worker. The sample consists of worker-day observations in both the Shanghai and Nantong offices. All regressions include year-monthspecific fixed effects and a control for the day of the week.

5 Discussion

5.1 Interpretation of Results & Key Findings

In this paper, we discuss the relationship between the pollution level and the worker productivity at a large call center in China. Our analysis also suggests that these productivity losses are largely linearly increasing in pollution levels. From the results above, a 10-unit increase in API decreases the number of calls per shift by 0.35 percent and the minutes logged in per shift by 0.25 percent.

Compared with physical work, the work at the call center is in-door so that many people don't pay much attention to this problem. It is common knowledge that air pollution will affect workers engaged in physical labor because air pollution will affect their physical state to a certain extent and even cause serious diseases imperceptibly. However, in this paper, we find out that air pollution also has a statistically significant effect on the worker productivity at the call center.

While call center work is mental rather than physical work, it is important to emphasize that it remains a semiskilled occupation. If our measured productivity impacts are the result of diminished cognitive function, the negative impacts of pollution may well be larger for high-skilled occupations that form the backbone of the service and information economy. The development of suitable measures of productivity in those occupations and assessing its relationship with environmental quality represents a fruitful area for future research.

5.2 Limitations

Since the dataset only includes the sample of workers at the call center at Ctrip, the final results are not necessarily universal. In addition, since the daily work of call center workers is relatively simple, the productivity of workers can be approximated by the number and duration of calls, but in other departments or even other industries, it may not be possible to measure productivity through a fixed method.

Besides air pollution, there are many other factors that can affect a person's work efficiency every day, such as mood and physical state, etc., but we cannot reflect these factors through quantitative data, so there will be certain errors in the results of the model.

5.3 Future Directions

Since we found that the air pollution can also affect indoor do not need the productivity of manual workers, so we need to push this to more professional more industries to determine the effects of air pollution on productivity, and thus the government can implement different policy for different industries according to the different quantitative indicators on the effect.

In addition, the research on this aspect can be carried out to more countries, because China is not the only developing country with the problem of environmental pollution, and many rapidly developing countries will experience the problem of environmental pollution.

The main purpose of this aspect of the problem is to cause everyone to pay attention to environmental protection, they are not only harmful to people's physical and mental health, but also have a great impact on everyone's daily life and work, so we must take action to protect the environment. It's also beneficial for our own.

6 References

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying. 2015. "Does Working from Home Work? Evidence from a Chinese Experiment." Quarterly Journal of Economics 130 (1): 165–218. https://nbloom.people.stanford.edu/sites/g/files/sbiybj4746/f/wfh.pdf

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. DOI: 10.1257/app.20150213

Hadley Wickham, Evan Miller. 2021. https://CRAN.R-project.org/package=haven

Hadley Wickham, Jennifer Bryan. 2019. https://CRAN.R-project.org/package=readxl

Hadley Wickham . 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https: //ggplot2.tidyverse.org.

R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Seaton A, MacNee W, Donaldson K, Godden D. Particulate air pollution and acute health effects. Lancet. 1995 Jan 21;345(8943):176-8. doi: 10.1016/s0140-6736(95)90173-6. PMID: 7741860.