Title
Dynamically Regulating Emissions of Stock Pollutants

Author
Wang, Qian

Publication Date
2023

Peer reviewed|Thesis/dissertation

Dynamically Regulating Emissions of Stock Pollutants

By

QIAN WANG
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Agricultural and Resource Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Kevin Novan, Chair

Erich Muehlegger

Michael Springborn

Jamie Hansen-Lewis

Committee in Charge

2023

Abstract

Air pollution has plagued cities around the world for years. Major metropolitan areas often install air pollution regulations that vary in stringency over time. Regulations on polluters are lax on days when the air quality is good. For days with high ambient air pollution, governments impose stricter measures to avoid exacerbating the already poor air quality. This study focuses specifically on the Heavy Air Pollution Emergency Plan (HAPEP) of Chengdu, China. During the winter months, Chengdu frequently faces periods of sustained, high levels of ambient particulate matter (PM). Importantly, these high levels of ambient PM are not ruled by temporary increases in the flow of PM being emitted each day. Instead, it is driven largely by the weather conditions. For example, during periods of stagnant wind conditions, the stock of PM grows in the air. This accumulation process continues until the weather conditions change – i.e., the wind picks up – and the stock of PM in the air is cleared out. This pollution pattern is representative of the cities which are in a closed basin terrain. The Chengdu HAPEP requires monitoring short-term PM forecasts. If the predicted PM exceeds certain concentration and duration thresholds, an air pollution alert is issued. Subsequently, a set of measures intended to lessen air pollution are imposed, including driving restrictions and production suspensions.

In the first chapter, I develop a dynamic optimization framework taking into account the characteristics of PM pollution by directly modeling the stock nature of pollutants and the cleanup process of pollutants. The model leads to an important conclusion: the optimal amount of pollution emission should always increase over time if there is no pollutant dispersion. Although in reality pollution dissipation is never absolutely zero, this conclusion indicates an incentive to delay the pollution for social welfare improvement. However, the variations in expected pollution dissipation should also be incorporated when deciding the optimal action.

In the second chapter, I empirically estimate the dynamic process of pollution dissipation by identifying how daily weather conditions drive the change in ambient PM levels from one day to the next. I also show that the weather conditions in the near future can be forecasted with high accuracy. By altering the timing of historical interventions according to both existing conditions and expected upcoming weather patterns, a 12.1% more PM pollution reduction and a 25.5% more bronchitis hospital visit reduction can be achieved within a period in 2018.

In the third chapter, I estimate the health effects of PM pollution. I find that pollution exposures up to six days ago are associated with contemporaneous bronchitis hospitalization. I also assess the curvature of the response function of respiratory hospitalization to PM pollution. I fail to find any convincing evidence that supports the presence of nonlinearity. One caveat of this conclusion is that I only consider possible non-linear effects of the contemporaneous pollution concentration but ignore the possible non-linear effects of the lagged pollution concentration, due to the complexity of the specification and the difficulty in identification if both of them are incorporated. Although this simplification might introduce a bias in an unclear way, it is of less concern since prediction instead of specific point estimates is the focus of this analysis.

Acknowledgement

I would like to express my sincere appreciation to Kevin Novan and Erich Muehlegger. You offered me invaluable support and guidance from the formation of research ideas to the completion of this dissertation. I have learned a lot from each of you. I truly appreciate your wisdom, advice, and inspiration. I am also grateful for the brilliant comments and suggestions from Michael Springborn and Jamie Hansen-Lewis. Without your help, I would not have had the opportunity to finish the work presented here.

I had a really wonderful journey with the Department of Agricultural and Resource Economics (ARE) as a graduate student. The program is of the highest quality for both teaching and research. The department spares no efforts to help students develop as an economist and a person. I would like to thank Pierre Mérel, Richard Sexton, Y. Hossein Farzin, James Chalfant, Quirino Paris, Travis Lybbert, Daniel Sumner, Rachael Goodhue, Julian Alston, Dalia Ghanem, and Katrina Jessoe for inspiring me into and guiding me through this field. My appreciation also goes to my peers in the program for creating a friendly and supportive environment. I have also benefited greatly from the courses I took in the Department of Economics, especially with Dave Rapson, Burkhard Schipper, and Andrés Carvajal. During my ARE program, I was hired by the Institute of Transportation Studies (ITS) as a graduate student researcher for many years. I am very grateful for the continued funding and research projects made available to me. This role also provided me with precious opportunities to interact with leaders in business, government, and NGOs. I was very fortunate to have extensive exposure to real-world issues so that I was not restricted by academic thinking. Many thanks to Lewis Fulton and Marshall Miller for their support and supervision of my work at ITS.

Most importantly, I would like to send my deepest gratitude to my family, especially my

parents. You brought me to this beautiful world and taught me how to live a wonderful life. Both of you have been my role models since childhood. I inherited curiosity, optimism, and passion from my mother, and integrity, diligence, and responsibility from my father. Although we were unable to spend much time together for the past ten years, your unconditional and endless love from the other side of the earth gives me tremendous courage to tackle the challenges in life.

Contents

List of Figures

List of Tables

1 Chapter 1: Theoretical Framework

1.1 Introduction

Major metropolitan areas around the world often install air pollution regulations that vary in stringency over time. Regulations on polluters are lax on days when the air quality is good. For days with high ambient air pollution, governments impose stricter measures to avoid exacerbating the already poor air quality. In this chapter, I model the decision of a planner who aims to maximize the net benefit of pollution. I consider the economic outputs as the benefit of pollution, and the health hazard as the damage of pollution. I examine how the stringency of pollution control measures should optimally vary over time. This study focuses specifically on Chengdu, China. During the winter months, Chengdu frequently faces periods of sustained, high levels of ambient particulate matter (PM). Importantly, these high levels of ambient PM are not ruled by temporary increases in the flow of PM being emitted each day. Instead, it is driven largely by the weather conditions. For example, during periods of stagnant wind conditions, the stock of PM grows in the air. This accumulation process continues until the weather conditions change – i.e., the wind picks up – and the stock of PM in the air is cleared out. This pollution pattern is representative of the cities which are in a closed basin terrain. I develop a dynamic optimization framework taking into account the characteristics of PM pollution by directly modeling the stock nature of pollutants and the cleanup process of pollutants. The model leads to an important conclusion: the optimal amount of pollution emission should always increase over time if there is no pollutant dispersion. Although in reality pollution dissipation is never absolutely zero, this conclusion indicates an incentive to delay the pollution for social welfare improvement. However, the variations in expected pollution dissipation should also be incorporated when deciding the optimal action.

This chapter makes the following two contributions. First, this study adopts a novel approach

to model PM as stock pollutants instead of flow pollutants. Traditionally, economists treat the damage of PM to be instantaneous like other criteria air pollutants such as carbon monoxide and nitrogen dioxide. Once in the air, these pollutants cause a one-time damage and then disappear - no further harms are considered. However, due to the basin climate characteristics of Chengdu, PM can stay in the air for days and cause damages repeatedly. The novel modeling approach developed in this study is able to better capture the dynamics of formation and dissipation of PM, and thus is a more precise representation of how nature works. Therefore, my analysis leads to conclusions that challenge the conventional wisdom and practice of air pollution control. Second, this study contributes to the active research field of evaluations of urban air pollution regulations. Economists have studied such policies around the globe, for example, Mexico City (Davis, 2008; Salas, 2010), Beijing (Viard and Fu, 2015; Zhong et al., 2017; Sun et al., 2014), Delhi (Greenstone et al., 2017), and Stockholm (Simeonova et al., 2019). These studies mostly seek to establish causal relationships between specific pollution control measures and ambient air pollution levels with reduced-form models. However, very little effort has been done to quantitatively examine the design of urban air pollution management schemes. The dearth of such effort can be mainly attributed to the challenge of modeling the pollution characteristics of an area. This study fills in this gap in the literature.

The rest of this chapter proceeds as follows: Section 1.2 provides details on the pollution regulations and pollution patterns in Chengdu, as well as the data used in this chapter. Section 1.3 describes the theoretical model and its implications. Section 1.4 summarizes the findings of this chapter.

1.2 Background and Data

1.2.1 Air Pollution Regulation in Chengdu

During the past forty years, China has experienced major economic growth. The nominal per capita GDP skyrocketed from just above $200 to over $10,000. However, air pollution, a byproduct of economic growth, has become a severe problem in China. National average air pollution levels are six to eight times higher in China compared to the U.S. (Greenstone and Hanna, 2014). In 2015, ambient air pollution accounted for approximately 4.2 million premature deaths globally, almost 40% of which occurred in China (Cohen et al., 2017). The suspended particulate has an economic cost of about $22.4 billion to China in 2005 (Matus et al., 2012). Chengdu, the largest city in Western China with a population of over 20 million and a GDP of more than $300 billion in 2021, has also suffered from excessive air pollution, especially from particulate matter. According to the Chinese air quality classification, Chengdu had only 132 days with good air quality in 2013, accounting for merely one-third of that year. To tackle the problem, the government of Chengdu adopted the Heavy Air Pollution Emergency Plan (HAPEP) in 2012. There have been three major revisions to the plan, in 2014, 2017, and 2020. Effective from November 2017 to March 2020, the 2017 HAPEP (referred to as HAPEP hereafter) is the latest version with an extended period of practice, and is thus the focus of this study. HAPEP targets PM pollution. The local EPA is required to monitor $PM_{2.5}$ forecast into future days. Based on the prediction of concentration and duration of $PM_{2.5}$ pollution, EPA decides whether to issue an emergency alert. It is worth pointing out that the EPA ignores predictions more than two days ahead. For example, if on a Monday the forecasted air quality for Tuesday or Wednesday is bad, then the EPA announces an emergency alert for the corresponding day right away. If, however, on a Monday the forecasted air quality for Tuesday or Wednesday is good but for Thursday is bad, the EPA waits until Tuesday to decide the proper action. There are four levels of emergency alert which are color-coded. The triggering conditions of each level of alert are

listed below.

1. If the predicted $PM_{2.5}$ concentration is higher than 115 micrograms per cubic meter ($\mu g/m^3$) for more than 24 hours, a blue alert is issued.

2. If the predicted $PM_{2.5}$ concentration is higher than 115 $\mu g/m^3$ for more than 48 hours, a yellow alert is issued.

3. If the predicted $PM_{2.5}$ concentration is higher than 115 $\mu g/m^3$ for more than 72 hours, and the predicted $PM_{2.5}$ concentration is higher than 150 $\mu g/m^3$ for more than 24 hours, an orange alert is issued.

4. If the predicted $PM_{2.5}$ concentration is higher than 115 $\mu g/m^3$ for more than 72 hours, and the predicted $PM_{2.5}$ concentration is higher than 250 $\mu g/m^3$ for more than 24 hours; or if the predicted $PM_{2.5}$ concentration is higher than 350 $\mu g/m^3$ for more than 6 hours; or if the current $PM_{2.5}$ concentration is higher than 350 $\mu g/m^3$ for more than 3 hours, a red alert is issued.

Each level of emergency alert is accompanied by a set of recommended and mandatory measures to reduce air pollution. Recommended measures are not enforced. As an example, below is a summary of important mandatory measures of the yellow alert.

- Certain heavy polluting industries are shut down. More firms are subject to mandatory limited and off-peak production.

- Heavy duty vehicles which produce excessive airborne dust (e.g. refuse trucks for construction waste) are banned in the central area.

- Aged vehicles with lower emission standards are banned in the downtown area from 6 am to 10 pm.
- The standard license plate based driving restriction within the downtown area (7:30 am to 8 pm) is extended to 6 am to 10 pm.

Figure 1 illustrates all the alert events that happened under HAPEP, with colors indicating the levels of alerts. No alert was ever issued outside the November to February window. Therefore, I name the period between November of a year and February of the next year a "pollution season". We can also see that alerts are always multi-day events. Yellow is the most common level of alert.

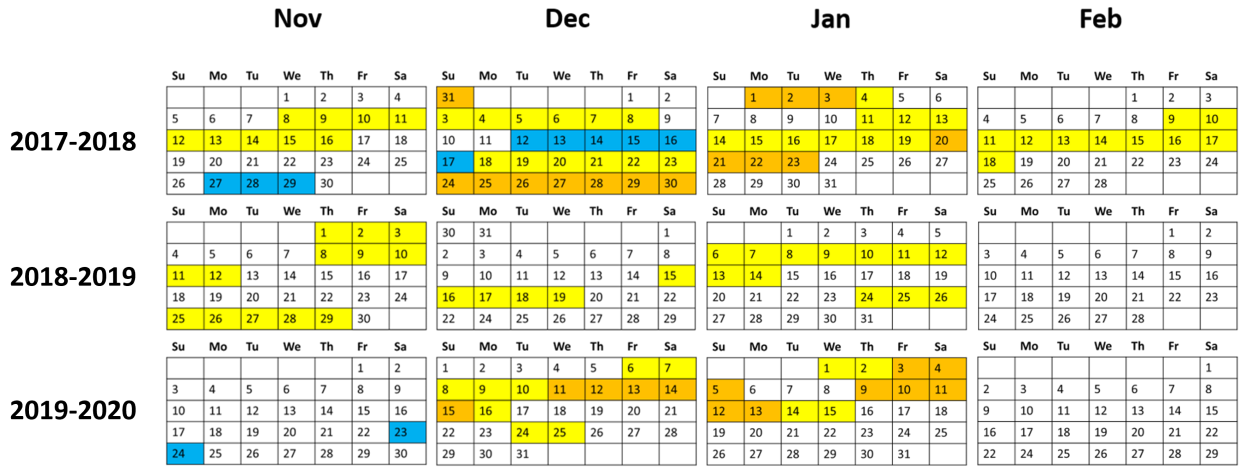

Figure 1: Chengdu HAPEP alert event history from Nov. 2017 to Mar. 2020

1.2.2 Winter Pollution Pattern in Chengdu

It is vital to have a good understanding of the pollution pattern before any modeling effort. The data shown in this chapter reflect the three pollution seasons from November 2017 to

February 2020. Figure 2 shows the daily city-average $PM_{2.5}$ concentration during three pollution seasons.

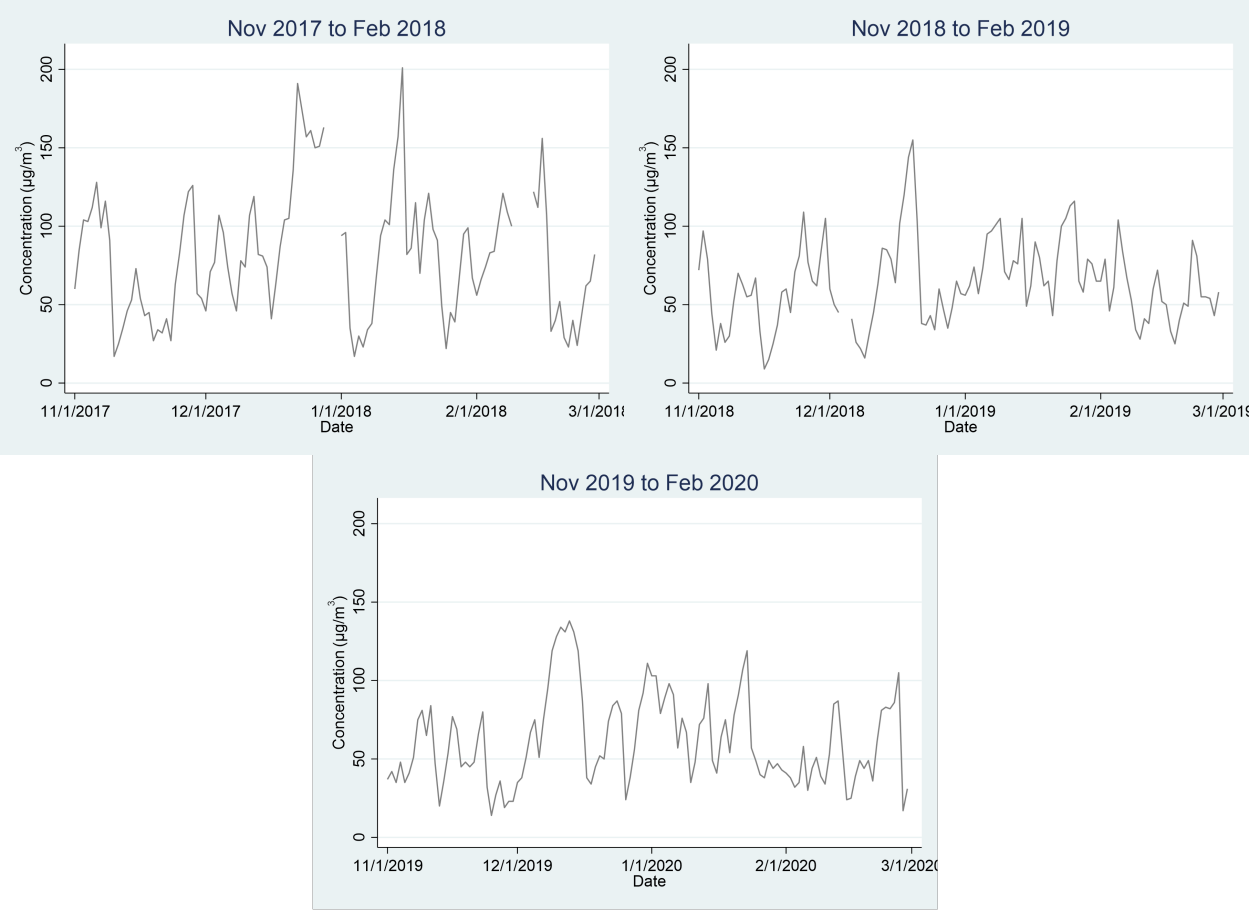

Figure 2: Daily city-average $PM_{2.5}$ concentration ($\mu g/m^3$) during three pollution seasons

Two characteristics of winter particulate matter pollution in Chengdu are noticed. First, the level of $PM_{2.5}$ concentration is rather high. The average $PM_{2.5}$ concentration during three pollution seasons is 68.6 $\mu g/m^3$, corresponding to a US air quality index (AQI) of 157 which falls into the "unhealthy" category. The worst $PM_{2.5}$ pollution during three pollution seasons occurred on January 15, 2018, with a concentration of astonishing 201 $\mu g/m^3$. This day had a US AQI reading of 251, accompanied by a "very unhealthy" air quality category. Second, the $PM_{2.5}$ concentration is extremely variable: it can easily double or drop by half in a few days. For example, the concentration skyrocketed from 38 $\mu g/m^3$ (January 8, 2018)

to 94 $\mu g/m^3$ (January 10, 2018) in two days, while it was able to plummet from 155 $\mu g/m^3$ (December 20, 2018) to 38 $\mu g/m^3$ (December 22, 2018) in two days. It is also noted that the level of $PM_{2.5}$ pollution in the first pollution season is higher than in the other two. Figure 3 shows the daily station-average $PM_{2.5}$ concentrations ($\mu g/m^3$) at individual monitoring stations during the first pollution season.

In Figure 3, Panel A and B correspond to the daily station-average $PM_{2.5}$ concentrations at monitoring stations in downtown and suburban area[1], respectively. The symbols on two sides indicate the relative location of the monitoring stations in the city area[2]. The red and blue vertical lines indicate the beginnings and endings of alert events, respectively. Each time series represents the daily station-average $PM_{2.5}$ concentration at one monitoring station. It is very obvious that no matter where the monitoring stations locate, they have a very similar level and trend of $PM_{2.5}$ concentration[3]. This pattern continues into the next two pollution seasons, of which the figures can be found in the appendix (Figure 5 and 6 in Section 1.5.3).

The observed pollution characteristics of Chengdu are well explained by physical science (Liao et al., 2017; Li et al., 2017; Lv et al., 2015). Chengdu is surrounded by Longmen-Qionglai Mountains[4] and Longquan Mountain on the west and east margin, respectively. This gives Chengdu a complex and closed basin terrain that generally makes pollutants difficult to diffuse. The basin climate renders wind and precipitation scarce in winter, slowing down the cleanup of pollutants. Humidity, which facilitates the formation of certain pol-

[1]The boundary is defined as the 4th Ring Road.

[2]The square box represents the city boundary. The dashed and solid circle stand for the 3rd and 4th Ring Road, respectively. The blue round area represents the downtown area. The red dot shows approximately where a monitoring station locates.

[3]As a further confirmation of this observation, Table 2 in the appendix (Section 1.5.2) shows summary statistics of daily station-average $PM_{2.5}$ concentration at different monitoring stations during the three pollution seasons, while Table 3 in the appendix (Section 1.5.2) demonstrates pairwise correlation coefficients of daily station-average $PM_{2.5}$ concentration at different monitoring stations during the three pollution seasons.

[4]They are part of the eastern edge of the Tibet Plateau.

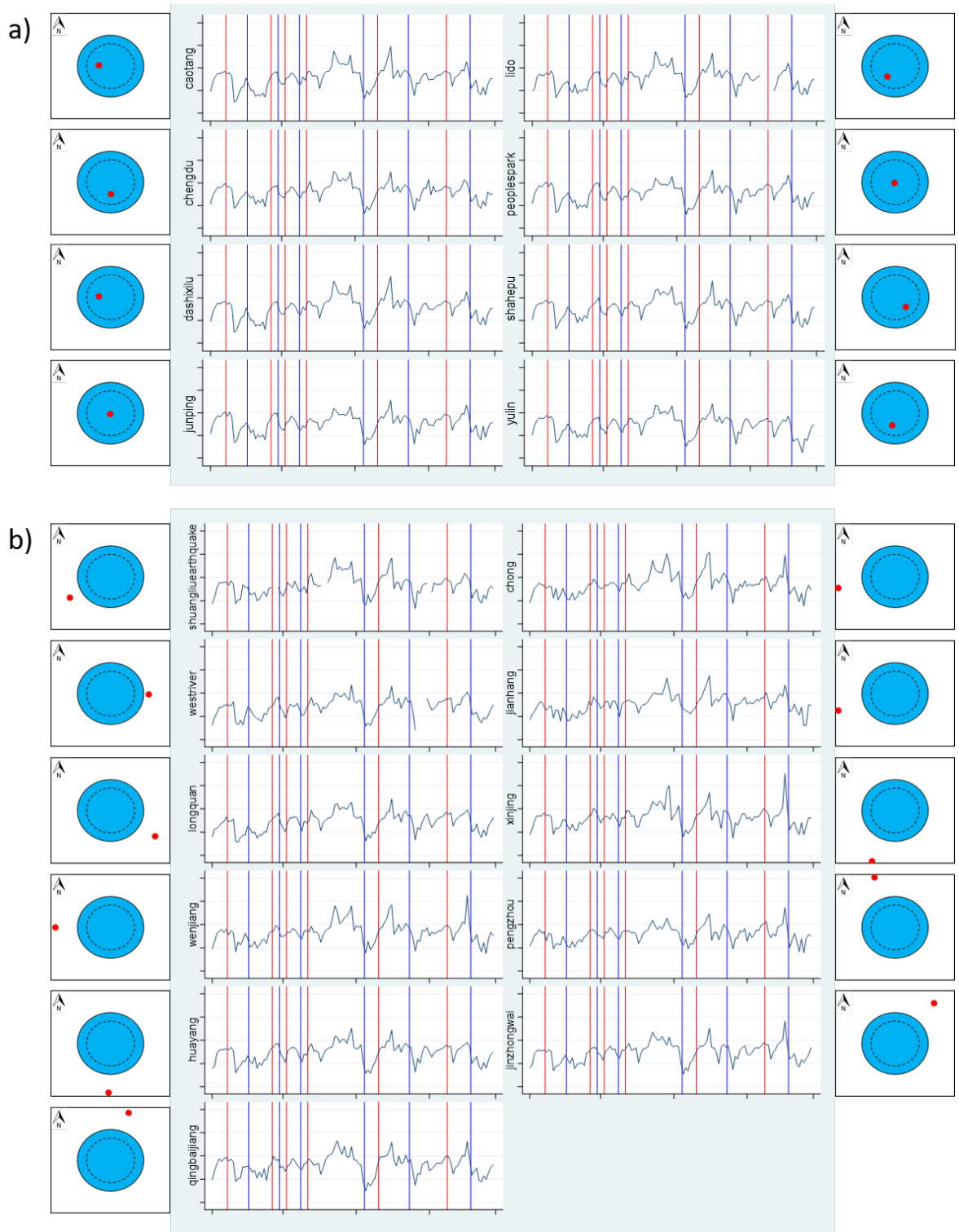

Figure 3: Daily station-average $PM_{2.5}$ concentrations ($\mu g/m^3$) at individual monitoring stations during the first pollution season

lutants, is relatively high in winter. Additionally, temperature inversion, an atmospheric phenomenon that inhibits the vertical movement of air mass and thus the dispersion of pollutants, is more frequent in winter. These factors combined with the large population and high vehicle ownership are the main contributors to the high winter PM pollution level in Chengdu. Also due to these topographic and climatic conditions, Liao et al. (2017) claim that PMs tend to accumulate without significant wind or precipitation. This explains the upsurges of daily city-average $PM_{2.5}$ concentration observed in Figure 2. However, a stretch of heavy pollution is ended when strong wind or precipitation happens. The pollutants are removed drastically from the city, resulting in the plunges of $PM_{2.5}$ concentration observed in Figure 2. By analyzing the $PM_{2.5}$ pollution pathways and sources, Lv et al. (2015) argue that $PM_{2.5}$ concentration in Chengdu is more sensitive to wind speed rather than wind direction. This conclusion is consistent with the geographic and climatic characteristics of Chengdu, and the similar level and trend of $PM_{2.5}$ concentration at individual monitoring stations located in different directions relative to the city center observed in Figure 3. To further examine the importance of wind speed, I look into the correlation between $PM_{2.5}$ concentration and wind speed (and its lags). Table 1 demonstrates the pairwise correlation coefficients between daily city-average $PM_{2.5}$ concentration and average wind speeds.

	$PM_{2.5}$	Wind (contemporaneous)	Wind (1-day Lag)	Wind (2-day Lag)	Wind (3-day Lag)
$PM_{2.5}$	1				
Wind (contemporaneous)	-0.29	1			
Wind (1-day Lag)	-0.43	0.42	1		
Wind (2-day Lag)	-0.31	0.08	0.42	1	
Wind (3-day Lag)	-0.22	-0.03	0.08	0.42	1

Table 1: Pairwise correlation coefficients between daily city-average $PM_{2.5}$ concentration and average wind speeds

As shown in Table 1, $PM_{2.5}$ concentration is highly correlated with wind speed. Specifically, the contemporaneous $PM_{2.5}$ concentration has the highest correlation with 1-day lagged wind speed. This is reasonable since the 1-day lagged wind speed has the most influence on the contemporaneous $PM_{2.5}$ concentration among the wind speed of different days. In

Figure 4, I overlap the 1-day lagged wind speed onto the time series shown in Figure 2. The grey and orange time series represents the daily city-average $PM_{2.5}$ concentration and 1-day lagged average wind speed, respectively. The patterns of $PM_{2.5}$ concentration and 1-day lagged wind speed align well with each other, which is consistent with the aforementioned conclusion from physical science that wind speed is the main driver of variation in $PM_{2.5}$ concentration. Precipitation, however, is very rare during these three pollution seasons - the average daily precipitation is only 0.01 inches. December 28, 2018 and November 24, 2019 have the most precipitation at 0.22 inches[5]. This level of rainfall is unlikely to have a meaningful impact on $PM_{2.5}$ concentration.

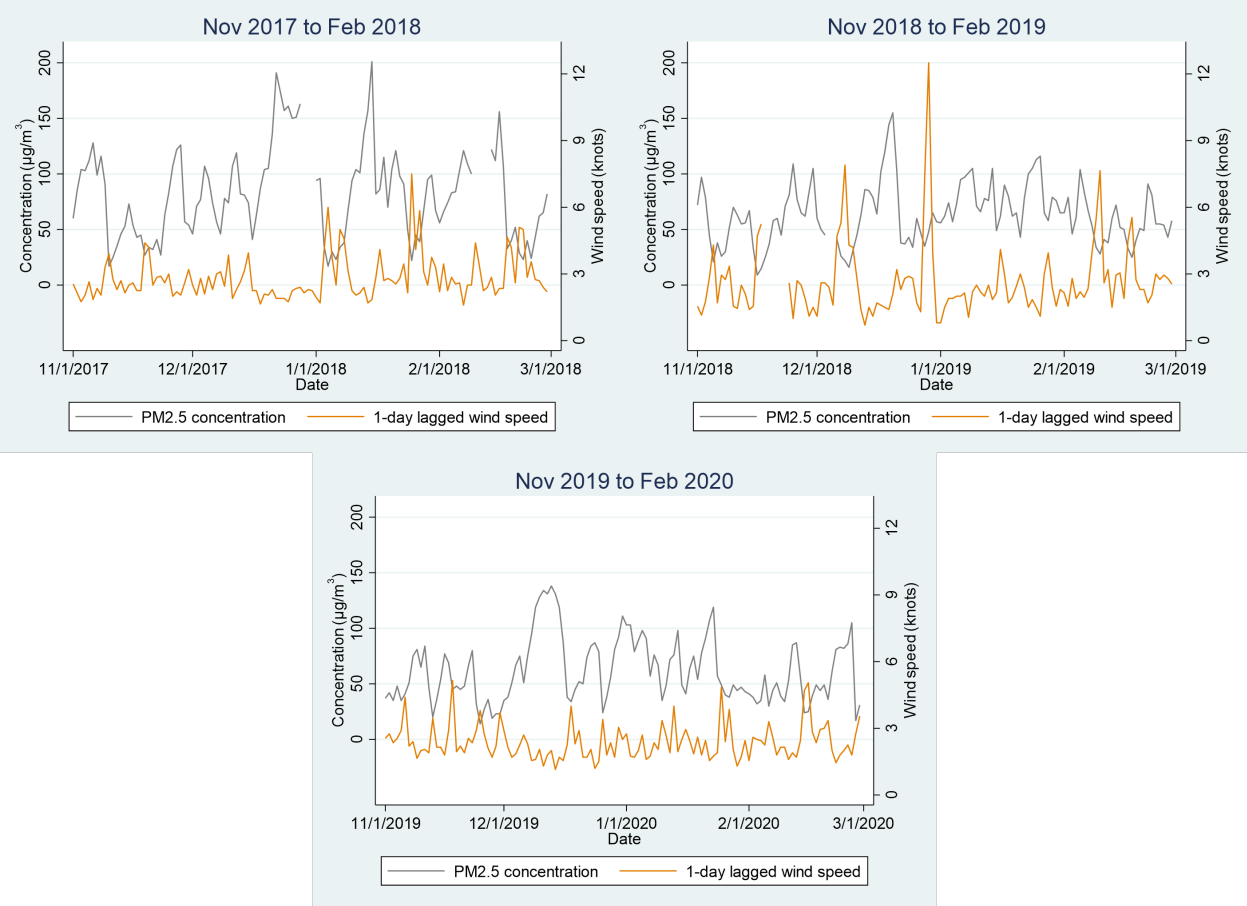

Figure 4: Daily city-average $PM_{2.5}$ concentration and 1-day lagged average wind speed during three pollution seasons

[5]Light rain is conventionally defined as precipitation that is less than 2.4 inch/day

To sum up, the physical science and evidence from data suggest the following PM pollution pattern: without significant wind, the PMs accumulate and mix well in the whole city area. An episode of heavy pollution is interrupted or ended by a "removal process" caused by wind. Contrary to the conventional wisdom that the damage of PMs is instantaneous, winter PM pollution in Chengdu should be treated as a temporary stock. PMs emitted earlier stay in the atmosphere longer, thus causing more damage compared to the PMs emitted closer to a "removal process". A focus on welfare then dictates that the optimal amount of emission should account for the time-varying marginal cost of pollution emissions. Moreover, since the PMs form a temporary pollution stock over days, the optimal action should incorporate its consequences in the future. Therefore, the static optimization framework will be insufficient in this context. Modeling the pollution control problem in Chengdu calls for a dynamic framework.

1.2.3 Data

The time span of the data in this chapter is from November 2017 to February 2020, pollution seasons only. Historical daily city-average air quality data were retrieved from *aqistudy.cn* which include AQI and concentrations of $PM_{2.5}$. Historical daily station-average $PM_{2.5}$ concentrations at individual monitoring stations were downloaded from *aqicn.org*. Official documentation of historical HAPEP emergency alert events indicating the starting and ending time of each event was acquired from Chengdu EPA.

1.3 Theoretical Framework

In this section, I develop a deterministic theoretical framework to model the decision of the pollution control planner, who aims to maximize the net benefit of pollution.

1.3.1 Basic Assumptions

Below are the basic assumptions of the modeling approach.

1. Pollution causes damages and brings benefits (e.g. economic outputs).
2. The planner aims to maximize the net benefit of pollution.
3. The time step is a day and the time horizon is infinity.
4. The planner can perfectly predict future meteorological/weather conditions.

Assumption 4 is not very unrealistic. In Chapter 2, I show empirically that the near future weather conditions in Chengdu can be predicted with high accuracy. The rest of the assumptions are self-explanatory. In this setup, the planner decides the optimal amount of pollution emission for the future days according to predicted weather conditions.

1.3.2 Functions

I define a social benefit function of pollution $B(c_t)$, where c_t is the amount of pollution emitted on day t. I assume $B(c_t)$ is non-negative, and strictly increasing in c_t. I also define a social damage function of pollution $D(c_t + s_t)$, where s_t denotes the pollution stock at the beginning of day t. I assume $D(c_t + s_t)$ is non-negative, and strictly increasing in $(c_t + s_t)$. The benefit and damage of pollution are assumed to be additively separable across time.

1.3.3 Optimization Problem

Below is the optimization problem of the planner.

$$\underset{c_t}{\text{Max}} \sum_{t=0}^{\infty} \delta^t [B(c_t) - D(c_t + s_t)] \qquad \text{s.t.} \quad s_{t+1} = (1 - r_t)(c_t + s_t)$$

$$c_t \geq 0$$

$$s_t \geq 0$$

$$\delta,\ r_t \in [0, 1]$$

r_t denotes the fraction of pollution that dissipates on day t (referred to as "dissipation factor" hereafter), which is exogenously determined by meteorological conditions. δ denotes the discount factor. For simplicity of derivation, I define $f(c_t, s_t) = B(c_t) - D(c_t + s_t)$ and $g(c_t, s_t) = s_{t+1} = (1 - r_t)(c_t + s_t)$. Now we can write the sequence problem as below.

$$\underset{c_t}{\text{Max}} \sum_{t=0}^{\infty} \delta^t f(c_t, s_t) \qquad \text{s.t.} \quad s_{t+1} = g(c_t, s_t)$$

$$c_t \geq 0$$

$$s_t \geq 0$$

$$\delta,\ r_t \in [0, 1]$$

Then I define a value function V at each t.

$$V(s_t) = \underset{c_t}{\text{Max}} \sum_{t=0}^{\infty} \delta^t f(c_t, s_t) \qquad \text{s.t.} \quad s_{t+1} = g(c_t, s_t)$$

According to Bellman's principle of optimality, the solutions to the above sequence problem correspond to the solutions to the following functional equation.

$$V(s_t) = \underset{c_t}{\text{Max}} \left[f(c_t, s_t) + \delta \cdot V(g(c_t, s_t)) \right] \tag{1}$$

1.3.4 Solving the Model

The first order condition of (1) is

$$\frac{\partial f(c_t, s_t)}{\partial c_t} + \delta \cdot \frac{\partial V(g(c_t, s_t))}{\partial s_{t+1}} \cdot \frac{\partial g(c_t, s_t)}{\partial c_t} = 0 \tag{2}$$

Rearrange Equation (2), we have

$$\frac{\partial V(g(c_t, s_t))}{\partial s_{t+1}} = -\frac{\frac{\partial f(c_t, s_t)}{\partial c_t}}{\delta \cdot \frac{\partial g(c_t, s_t)}{\partial c_t}} \tag{3}$$

By envelope theorem, we have

$$\frac{\partial V(s_t)}{\partial s_t} = \frac{\partial f(c_t, s_t)}{\partial s_t} + \delta \cdot \frac{\partial V(g(c_t, s_t))}{\partial s_{t+1}} \cdot \frac{\partial g(c_t, s_t)}{\partial s_t} \tag{4}$$

Plug Equation (3) into Equation (4).

$$\frac{\partial V(s_t)}{\partial s_t} = \frac{\partial f(c_t, s_t)}{\partial s_t} - \frac{\partial f(c_t, s_t)}{\partial c_t} \cdot \frac{\frac{\partial g(c_t, s_t)}{\partial s_t}}{\frac{\partial g(c_t, s_t)}{\partial c_t}} \tag{5}$$

Combine Equation (2) and Equation (5), we have

$$\frac{\partial f(c_t, s_t)}{\partial c_t} + \delta \cdot [\frac{\partial f(c_{t+1}, s_{t+1})}{\partial s_{t+1}} - \frac{\partial f(c_{t+1}, s_{t+1})}{\partial c_{t+1}} \cdot \frac{\frac{\partial g(c_{t+1}, s_{t+1})}{\partial s_{t+1}}}{\frac{\partial g(c_{t+1}, s_{t+1})}{\partial c_{t+1}}}] \cdot \frac{\partial g(c_t, s_t)}{\partial c_t} = 0 \tag{6}$$

Now we substitute in $f(c_t, s_t) = B(c_t) - D(c_t + s_t)$ and $g(c_t, s_t) = (1 - r_t)(c_t + s_t)$.

$$B'(c_t) - \frac{\partial D(c_t + s_t)}{\partial c_t} + \delta \cdot [-\frac{\partial D(c_{t+1} + s_{t+1})}{\partial s_{t+1}} - (B'(c_{t+1}) - \frac{\partial D(c_{t+1} + s_{t+1})}{\partial c_{t+1}})] \cdot (1 - r_t) = 0 \tag{7}$$

Rearrange Equation (7), we have

$$\delta(1 - r_t)B'(c_{t+1}) - B'(c_t) = \delta(1 - r_t)(\frac{\partial D(c_{t+1} + s_{t+1})}{\partial c_{t+1}} - \frac{\partial D(c_{t+1} + s_{t+1})}{\partial s_{t+1}}) - \frac{\partial D(c_t + s_t)}{\partial c_t} \tag{8}$$

I assume $\delta = 1$ hereafter, which means there is no discounting for the future. Note that this is most likely to be a valid assumption in the context of HAPEP as a pollution season only lasts four months and there is no particular reason why the planner would put a higher value on the present than the near future. Then Equation (8) becomes

$$(1-r_t)B'(c_{t+1}) - B'(c_t) = (1-r_t)(\frac{\partial D(c_{t+1}+s_{t+1})}{\partial c_{t+1}} - \frac{\partial D(c_{t+1}+s_{t+1})}{\partial s_{t+1}}) - \frac{\partial D(c_t+s_t)}{\partial c_t} \quad (9)$$

It is impossible to determine if c_{t+1} or c_t is larger according to Equation (9) without further information. However, we can evaluate several extreme cases.

Case 1: $\mathbf{r_t = 1}$ $\forall$t

According to Equation (9), we have

$$B'(c_t) = \frac{\partial D(c_t+s_t)}{\partial c_t}$$

In this case, I assume full dissipation for each period, thus breaking the stock nature of the pollutant. The dynamic optimization problem boils down to a classic static optimization problem where the planner chooses the optimal emission quantity which equates the marginal benefit and marginal damage of pollution in each period. Since there is no carry-over of pollution stock, the optimal quantity of emission in each period is the same.

Case 2: $\mathbf{r_t = 0}$ $\forall t$

According to Equation (9), we have

$$B'(c_{t+1}) - B'(c_t) = (\frac{\partial D(c_{t+1}+s_{t+1})}{\partial c_{t+1}} - \frac{\partial D(c_{t+1}+s_{t+1})}{\partial s_{t+1}}) - \frac{\partial D(c_t+s_t)}{\partial c_t}$$

Since $D(c_t+s_t)$ is strictly increasing in c_t, we know $-\frac{\partial D(c_t+s_t)}{\partial c_t} < 0$. And because $\frac{\partial D(c_{t+1}+s_{t+1})}{\partial c_{t+1}} = \frac{\partial D(c_{t+1}+s_{t+1})}{\partial s_{t+1}}$, we have

$$B'(c_{t+1}) < B'(c_t)$$

Due to the strict concavity of the benefit function, it must be that $c_{t+1} > c_t \ \forall t$. In this case, I assume no dissipation, which gives rise to an increasing pollution stock. The stock nature leads to a very important conclusion: without pollutant dispersion, the optimal amount of pollution emission should always increase over time.

In reality, the dissipation factor is never zero. Even without any wind, pollution dispersion happens to some extent. Therefore, the scenario that the dissipation factor is always zero in Case 2 is not realistic. However, this case provides an important intuition that pollution emitted later creates less damage due to its shorter presence in the air. This connects to the actual setting of this study because days with very high dissipation factors exist during and after the pollution season when almost all pollutants in the air are removed. In expectation of these days, there is an incentive to delay pollution closer to the removal process in order to improve social welfare, although it is unclear how this incentive exactly compares to the incentive driven by expected increases and decreases in dissipation factors.

1.4 Conclusions

Air pollution has plagued major cities around the world for years. In this chapter, I develop a dynamic programming framework to model the decision of a planner who aims to maximize the net benefit of air pollution in Chengdu, China. This framework takes into account the characteristics of PM pollution by directly modeling the stock nature of pollutants and the cleanup process of pollutants. The model provides important policy implications. If there is no pollutant dispersion, the optimal amount of pollution emission should always increase over time. Although in reality pollution dissipation is never absolutely zero, this conclusion

indicates an incentive to delay the pollution for social welfare improvement. However, the variations in expected pollution dissipation should also be incorporated when deciding the optimal action. When pollutants completely dissipate in each period, i.e. there is no carry-over of pollution stock, the dynamic framework is equivalent to the classic static optimization model, matching the conventional wisdom of pollution control.

This theoretical model represents an ideal scenario where the planner can update her knowledge about future weather conditions on a daily basis and the planning horizon is infinite. In Section 1.5.1, I present an alternative modeling approach where the planner is required to make weekly planning in advance to ensure the practicability of the regulations. With similar assumptions, the alternative modeling approach reach the same conclusions.

1.5 Appendix

1.5.1 Alternative Modeling Approach

Basic Assumptions

Below are the basic assumptions of the modeling approach.

1. Pollution causes damages and brings benefits (e.g. economic outputs).
2. The planner aims to maximize the net benefit of pollution over a planning horizon.
3. The time step is a day and the time horizon is a week.
4. The planner can perfectly predict meteorological/weather conditions within a planning horizon.
5. The planner takes into account the damage of scrap pollution that remains at the terminal time.

Assumption 3 ensures the practicality of the optimal policy because the planning horizons and periods align well with human activity cycles (i.e. weeks and days). Assumption 5 arises from the fact that the planner takes into account the damage of residual pollution that will realize beyond the current planning horizon. The rest of the assumptions are self-explanatory. In this setup, the planner decides the optimal amount of pollution emission of each day in the upcoming week according to predicted weather conditions.

Functions

I define a social benefit function of pollution $B(c_t)$, where c_t is the amount of pollution emitted on day t. I assume $B(c_t)$ is non-negative, twice continuously differentiable, strictly increasing in c_t, and strictly concave in c_t. The concavity comes from the diminishing marginal return of pollution. I also define a social damage function of pollution $D(c_t + s_t)$, where s_t denotes the pollution stock at the beginning of day t. I assume $D(c_t + s_t)$ is non-negative, twice continuously differentiable, strictly increasing in (c_t+s_t), and strictly convex in (c_t+s_t). The convexity assumes the marginal damage of pollution is increasing in pollution level. The benefit and damage of pollution are assumed to be additively separable across time. I also define a scrap damage function of pollution $F(s_T, T)$. I assume $F(s_T, T)$ is non-positive, twice continuously differentiable, strictly decreasing in s_T, and strictly concave in s_T.

Optimization Problem

Below is the optimization problem of the planner.

$$\underset{c_t}{\text{Max}} \quad \sum_{t=0}^{T-1} B(c_t) - D(c_t + s_t) + F(s_T, T) \qquad \text{s.t.} \quad s_{t+1} = (1 - r_t)(c_t + s_t)$$

$$s_0 \quad \text{given}$$

$$s_T \quad \text{free}$$

$$T \quad fixed$$

$$c_t \geq 0$$

$$s_t \geq 0$$

$$r_t \in [0, 1]$$

T represents the terminal time of a planning horizon. r_t denotes the fraction of pollution that dissipates on day t (referred to as "dissipation factor" hereafter), which is exogenously determined by meteorological conditions. The planner aims to maximize the net benefit of pollution subject to state transition and a number of constraints: The initial pollution stock is known to the planner; There is no constraint on the terminal pollution stock; T is fixed because a planning horizon is a week; The rest of constraints are self-explanatory. Note that discounting is trivial and thus ignored in this framework because the planning horizon is only a week.

Necessary Conditions

The Hamiltonian of the optimization problem is:

$$H(c_t, s_t, \pi_t) = B(c_t) - D(c_t + s_t) + \pi_t[c_t - r_t(c_t + s_t)]$$

π_t is the co-state variable of day t. The optimal solution of this dynamic model must satisfy

the following conditions.

$$\frac{\partial H}{\partial c_t} = B'(c_t) - \frac{\partial D(c_t + s_t)}{\partial c_t} + \pi_t(1 - r_t) = 0 \tag{10}$$

$$-\frac{\partial H}{\partial s_t} = \frac{\partial D(c_t, s_t)}{\partial s_t} + \pi_t r_t = \pi_t - \pi_{t-1} \tag{11}$$

$$\frac{\partial H}{\partial \pi_t} = c_t - r_t(c_t + s_t) = s_{t+1} - s_t \tag{12}$$

$$\pi_T = \frac{\partial F(s_T, T)}{\partial s_T} \qquad \text{(transversality condition)} \tag{13}$$

We can check the second-order condition. According to my assumptions of benefit and damage functions, we have $\frac{\partial^2 H}{\partial c_t^2} = B''(c_t) - \frac{\partial^2 D(c_t, s_t)}{\partial c_t^2} < 0$. Therefore, we do have a maximum. The transversality condition comes from the fact that s_T is free and T is fixed, and the planner considers the scrap damage of pollution.

Path of Co-state Variable π_t

With the above conditions, we can trace out the path of π_t. Rearrange Equation (11):

$$\pi_{t-1} = \pi_t(1 - r_t) - \frac{\partial D(c_t + s_t)}{\partial s_t} \tag{14}$$

With Equation (13) and (14), we have

$$\pi_T = \frac{\partial F(s_T, T)}{\partial s_T}$$

$$\pi_{T-1} = (1 - r_T)\frac{\partial F(s_T, T)}{\partial s_T} - \frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T}$$

$$\pi_{T-2} = (1 - r_T)(1 - r_{T-1})\frac{\partial F(s_T, T)}{\partial s_T} - (1 - r_{T-1})\frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T} - \frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T-1}$$

$$\pi_{T-3} = (1 - r_T)(1 - r_{T-1})(1 - r_{T-2})\frac{\partial F(s_T, T)}{\partial s_T} - (1 - r_{T-1})(1 - r_{T-2})\frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T} - (1 - r_{T-2})\frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T-1} - \frac{\partial D(c_t + s_t)}{\partial s_t}|_{t=T-2}$$

$$\pi_{T-4} = ...$$

Therefore, we have:

$$\pi_t = \frac{\partial F(s_T, T)}{\partial s_T} \cdot \prod_{u=t+1}^{T}(1 - r_u) - \sum_{u=t+1}^{T}[\frac{\partial D(c_u + s_u)}{\partial s_u} \cdot \prod_{v=t+1}^{u-1}(1 - r_v)] \tag{15}$$

u and v are alternative time subscripts. According to my assumptions, $\frac{\partial F(s_T,T)}{\partial s_T} < 0$, $1 - r_t > 0$, $\frac{\partial D(c_t+s_t)}{\partial s_t} > 0$. Therefore, $\pi_t < 0$. It means that during a planning horizon the "shadow price" of pollution stays negative, or damage of marginal pollution stays positive. Equation (15) indicates that the current damage of marginal pollution comes from two sources: "depreciated" scrap damage of pollution (the first term), and the "depreciated" marginal damage of pollution from future periods (the second term). "Depreciation" refers to the dissipation of pollution in this context. We can also write Equation (15) in its difference form:

$$\pi_t - \pi_{t-1} = r_t \cdot \{\frac{\partial F(s_T, T)}{\partial s_T} \cdot \prod_{u=t+1}^{T}(1 - r_u) - \sum_{u=t+1}^{T}[\frac{\partial D(c_u + s_u)}{\partial s_u} \cdot \prod_{v=t+1}^{u-1}(1 - r_v)]\} + \frac{\partial D(c_t + s_t)}{\partial s_t} \tag{16}$$

It is impossible to determine if π_t or π_{t-1} is larger according to Equation (16) without more information. Intuitively, as time moves forward in a planning horizon, the scrap damage of pollution is subject to less "depreciation", thus contributing more to the co-state variable. However, there is also less "depreciated" marginal damage from the future periods, thus contributing less to the co-state variable. Therefore, we are not able to sign the left side of Equation (16) without knowing which effect dominates.

Path of Pollution Emission $\mathbf{c_t}$

The optimal path of pollution emission is of the top interest. Rearrange Equation (10) and (14), we have

$$B'(c_t) - \frac{\partial D(c_t + s_t)}{\partial c_t} = -\pi_t(1 - r_t) \tag{17}$$

$$-\pi_t(1 - r_t) = -\pi_{t-1} - \frac{\partial D(c_t + s_t)}{\partial s_t} \tag{18}$$

Combine Equation (17) and (18), we have

$$B'(c_t) = -\pi_{t-1} \tag{19}$$

The implication of Equation (19) is clear: the optimal path of pollution requires that in each period the planner selects the amount of emission (c_t) such that the marginal benefit of pollution equals the damage of marginal pollution in the previous period, which, according to Equation (15), consists of the "depreciated" marginal scrap damage, the "depreciated" marginal damage of pollution from future periods, and the marginal damage of pollution from the current period. Next, I investigate how the optimal c_t change over time. Let $n \in \{1, 2, ..., T-1\}$. We can write Equation (19) in its difference form:

$$B'(c_{n+1}) - B'(c_n) = \pi_{n-1} - \pi_n \tag{20}$$

Plug Equation (16) into Equation (20), we have

$$B'(c_{n+1}) - B'(c_n) = r_n \cdot \{ \sum_{u=n+1}^{T} [\frac{\partial D(c_u + s_u)}{\partial s_u} \cdot \prod_{v=n+1}^{u-1} (1 - r_v)] - \frac{\partial F(s_T, T)}{\partial s_T} \cdot \prod_{u=n+1}^{T} (1 - r_u) \} - \frac{\partial D(c_n + s_n)}{\partial s_n} \tag{21}$$

It is impossible to determine if c_{n+1} or c_n is larger according to Equation (21) without more information. However, we can evaluate several extreme cases.

Case 1: $r_n = 1 \ \forall n$

According to Equation (10), we have

$$B'(c_n) = \frac{\partial D(c_n + s_n)}{\partial c_n}$$

Moreover, according to the state transition equation, we know $s_n = 0 \ \forall n$. Now Equation (21) becomes:

$$B'(c_{n+1}) - B'(c_n) = \frac{\partial D(c_{n+1} + s_{n+1})}{\partial s_{n+1}}|_{s_{n+1}=0} - \frac{\partial D(c_n + s_n)}{\partial s_n}|_{s_n=0} = 0$$

So $B'(c_{n+1}) = B'(c_n)$ must always hold. Since $B''(\cdot) < 0$, thus $c_{n+1} = c_n \ \forall n$.

In this case, I assume full dissipation for each period, thus breaking the stock nature of the pollutant. The dynamic optimization problem boils down to a classic static optimization problem where the planner chooses the optimal emission quantity which equates the marginal benefit and marginal damage of pollution in each period. Since there is no carry-over of pollution stock, the optimal quantity of emission in each period is the same.

Case 2: $r_n = 0$

According to Equation (21), we have

$$B'(c_{n+1}) - B'(c_n) = -\frac{\partial D(c_n + s_n)}{\partial s_n}$$

Since the damage function is strictly increasing in its arguments, $\frac{\partial D(c_n+s_n)}{\partial s_n} > 0$. So we have $B'(c_{n+1}) - B'(c_n) < 0$, and thus $B'(c_{n+1}) < B'(c_n)$. Due to the strict concavity of the benefit function, it must be $c_{n+1} > c_n$.

I assume no dissipation for a specific period in this case. The result indicates that if there is no pollution dispersion in the previous period, the optimal amount of emission of the current period should be higher than the previous period.

Case 3: $r_n = 0 \ \forall n$

We can generalize the previous case by assuming $r_n = 0 \ \forall n$. Then $c_{n+1} > c_n$ must always hold.

In this case, I assume no dissipation, which gives rise to an increasing pollution stock. The stock nature leads to a very important conclusion: without pollutant dispersion, the optimal amount of pollution emission should always increase over time.

Concluding Remarks

In this section, I develop an optimal control framework to model the decision of the pollution control planner. Compared to the dynamic programming framework in Section 1.3, this approach places greater emphasis on practicality - the planning horizons align well with human activity cycles (i.e. weeks). Implementation of the suggested policy is relatively straightforward: the planner makes decisions on c_t for the upcoming week according to

the predicted r_t reported by EPA. Depending on the calculated c_t, corresponding day-by-day measures (or no measure) are announced to the public before the beginning of a new week. The two different approaches reach the same conclusions. Although it is noted that the optimal control approach requires stronger assumptions of the curvature of the functions.

1.5.2 Tables

For a clear presentation, acronyms instead of full names of monitoring stations are displayed in the tables of this section.

Station	Obs	Mean	Std. Dev.	Min	Max
cd	358	148.9	33.9	61	250
ct	291	143.1	40.3	30	296
cy	210	138.9	32.7	57	212
cz	357	147.6	44.0	40	308
ds	358	141.6	39.8	29	296
hb	118	128.9	34.5	42	195
hy	357	138.9	35.0	31	252
jb	237	131.5	33.8	47	205
jh	268	132.9	40.7	44	275
jp	358	145.6	35.8	42	256
jy	118	126.0	33.8	59	194
jz	357	131.5	39.0	21	283
ld	352	142.2	38.1	42	272
lj	118	135.5	35.8	52	219
lq	358	133.3	36.0	33	243
pp	291	146.7	35.8	42	256
pz	357	141.9	34.7	56	274
qb	357	145.0	39.8	44	265
sl	263	150.1	40.4	57	286
sp	358	141.6	37.9	38	275
sx	169	130.4	32.5	39	202
us	221	144.0	32.3	61	298
wj	356	150.2	41.3	47	326
wr	265	130.4	36.7	33	236
xj	268	140.9	44.4	47	352
xm	150	142.0	31.7	66	216
yl	298	138.1	37.6	23	240
yn	148	139.4	35.1	53	226

Table 2: Summary statistics of daily station-average $PM_{2.5}$ concentrations at different monitoring stations during the three pollution seasons

	cd	ct	cy	cz	ds	hb	hy	jb	jh	jp	jy	jz	ld	lj	lq	pp	pz	qb	sl	sp	sx	us	wj	wr	xj	xm	yl	yn
cd	1																											
ct	.94	1																										
cy	.90	.93	1																									
cz	.82	.85	.89	1																								
ds	.94	.99	.94	.85	1																							
hb	.94	.99	.95	.84	.96	1																						
hy	.93	.94	.91	.79	.94	.96	1																					
jb	.88	.94	.94	.82	.92	.96	.92	1																				
jh	.78	.84	.79	.92	.82	.74	.76	.77	1																			
jp	.94	.96	.94	.81	.96	.96	.94	.93	.77	1																		
jy	.94	.98	.96	.91	.96	.95	.92	.91	.79	.95	1																	
jz	.88	.93	.86	.84	.89	.91	.87	.87	.80	.88	.83	1																
ld	.93	.97	.93	.86	.96	.99	.94	.94	.82	.94	.95	.91	1															
lj	.94	.98	.97	.86	.97	.98	.94	.94	.78	.96	.96	.88	.99	1														
lq	.91	.93	.88	.79	.92	.90	.95	.87	.79	.91	.85	.90	.91	.88	1													
pp	.93	.96	.92	.81	.96	.98	.94	.94	.78	.99	.97	.91	.95	.98	.92	1												
pz	.81	.85	.84	.88	.85	.88	.81	.85	.87	.83	.87	.85	.86	.88	.80	.83	1											
qb	.91	.92	.92	.87	.91	.89	.87	.87	.79	.91	.93	.92	.91	.90	.86	.91	.85	1										
sl	.93	.94	.94	.89	.94	.96	.90	.93	.83	.90	.94	.88	.95	.96	.87	.90	.88	.90	1									
sp	.95	.96	.94	.84	.96	.95	.95	.92	.81	.96	.93	.89	.95	.94	.94	.96	.84	.92	.92	1								
sx	.93	.92	.93	.83	.91	.92	.93	.90	.78	.91	.91	.86	.91	.90	.91	.91	.81	.89	.91	.94	1							
us	.98	.86	.86	.75	.88	.95	.88	.84	.73	.88	.94	.83	.86	.94	.87	.85	.76	.85	.91	.89	.93	1						
wj	.89	.93	.93	.93	.93	.95	.89	.93	.88	.90	.94	.90	.94	.96	.87	.90	.91	.91	.94	.91	.90	.83	1					
wr	.89	.90	.86	.79	.89	.92	.92	.90	.77	.89	.84	.90	.92	.90	.92	.89	.82	.85	.88	.90	.87	.87	.87	1				
xj	.87	.88	.92	.92	.87	.92	.85	.89	.86	.85	.92	.89	.89	.92	.84	.86	.90	.88	.90	.89	.91	.87	.94	.84	1			
xm	.86	.92	.90	.80	.89	.90	.85	.92	.72	.87	.87	.88	.92	.90	.81	.90	.89	.87	.88	.88	.84	.86	.92	.85	.82	1		
yl	.94	.94	.93	.77	.94	.96	.96	.93	.72	.95	.93	.85	.94	.95	.92	.94	.80	.89	.89	.96	.93	.94	.88	.88	.81	.87	1	
yn	.87	.91	.92	.87	.88	.92	.85	.93	.76	.88	.90	.88	.92	.92	.81	.88	.92	.87	.95	.88	.87	.86	.96	.86	.89	.94	.87	1

Table 3: Pairwise correlation coefficients of daily station-average $PM_{2.5}$ concentrations at different monitoring stations during the three pollution seasons

1.5.3 Figures

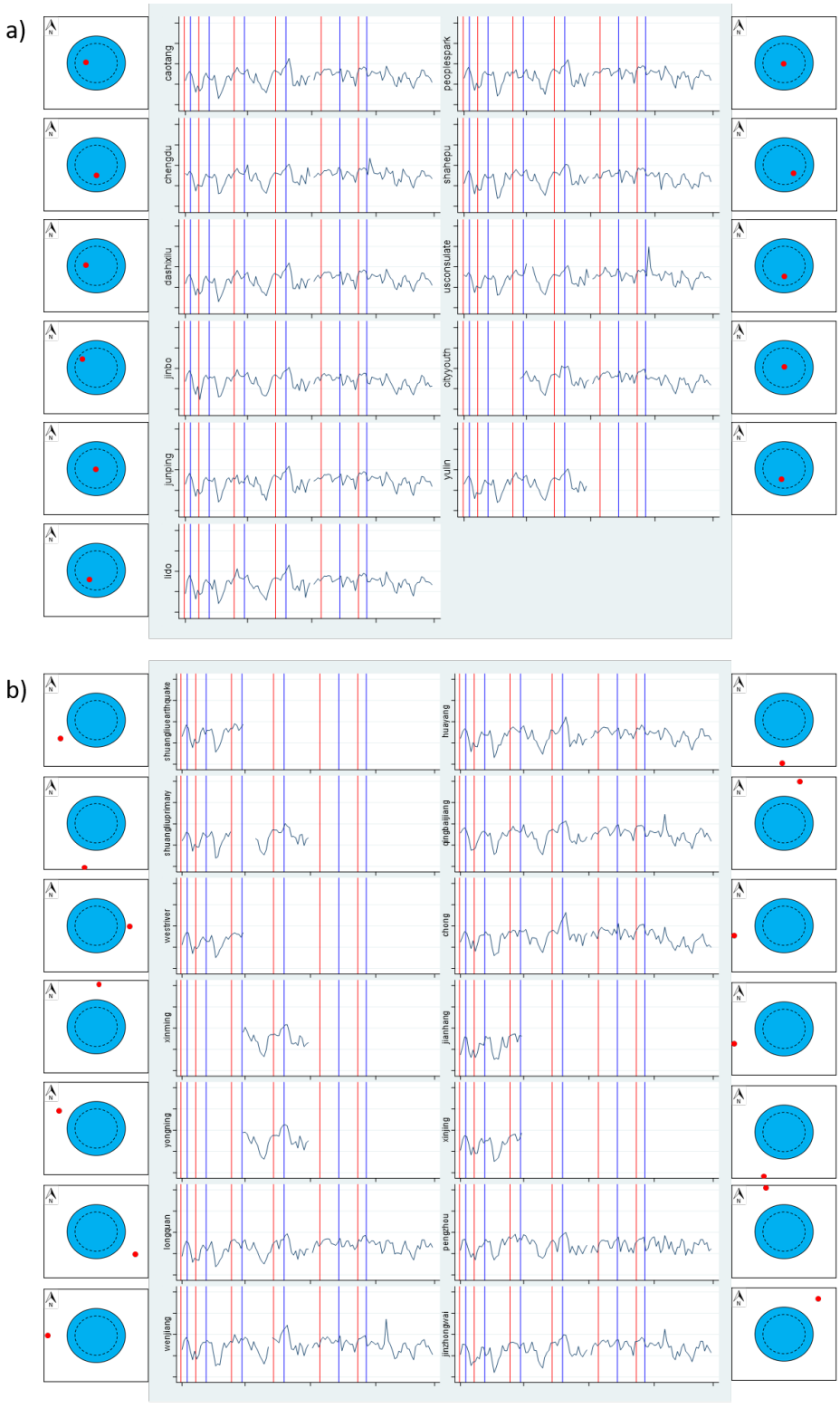

Figure 5: Daily station-average $PM_{2.5}$ concentrations at individual monitoring stations during the second pollution season

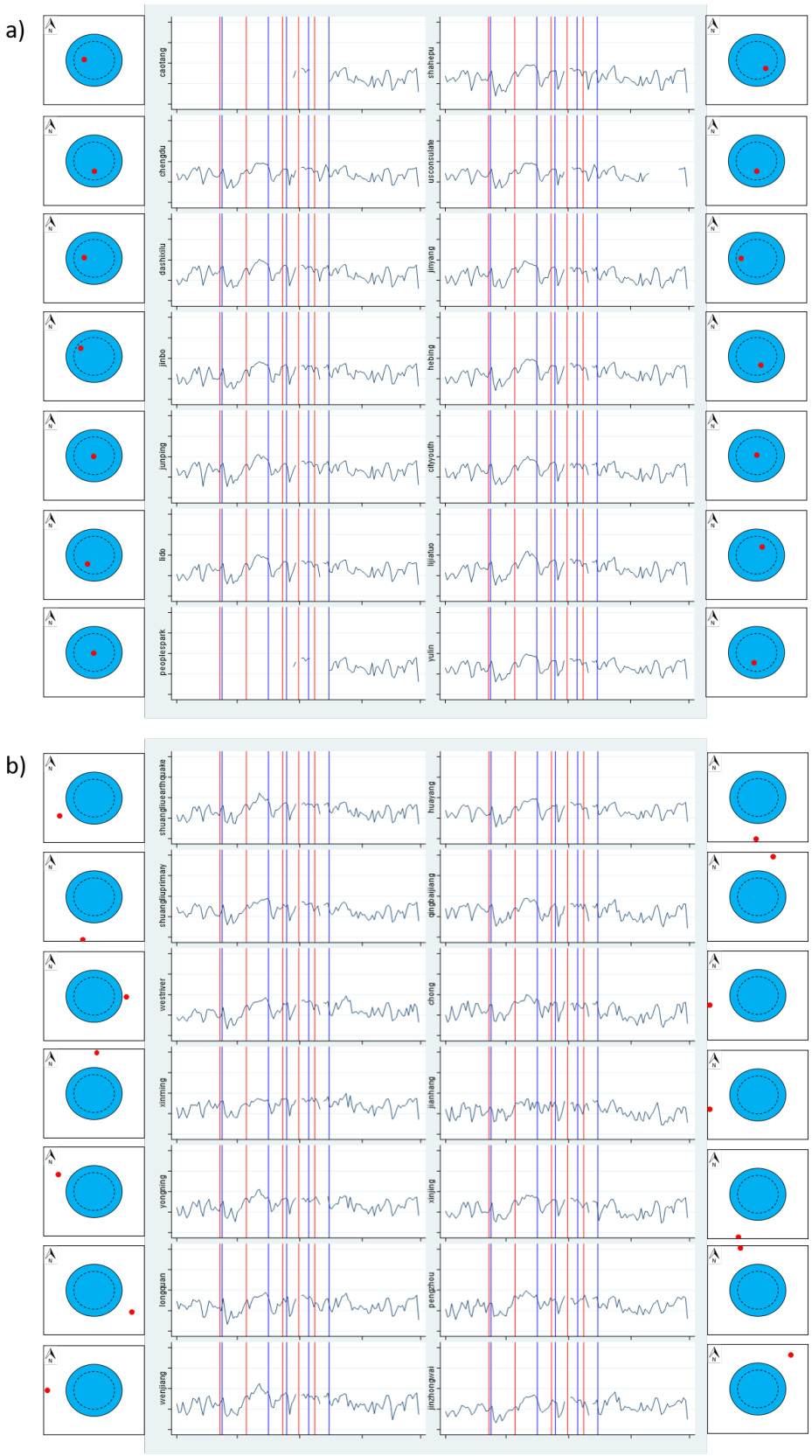

Figure 6: Daily station-average $PM_{2.5}$ concentrations at individual monitoring stations during the third pollution season

2 Chapter 2: Empirics

2.1 Introduction

In this chapter, I first test how well the historical air quality data can be fitted to the state transition equation, detailed in Section 2.2. This exercise helps assess the approach to modeling the pollution turnover process presented in Chapter 1. By estimating the state transition equation, historical dissipation factors can be back-calculated. Next, I empirically estimate the accuracy of weather forecasts in Chengdu by comparing the predicted and realized meteorological conditions from the same weather station. Since I assume the dissipation factor is a function of weather conditions, the accuracy of forecasted dissipation factors can also be estimated. If dissipation factors in the near future can be predicted relatively accurately, then the pollution control strategy based on both existing conditions and expected upcoming weather patterns is practical. This analysis is documented in Section 2.3. Lastly, I select a period in history and demonstrate the gains that could have been achieved by altering the timing of historical interventions, reported in Section 2.4. In Section 2.5, I summarize the findings of this chapter.

2.2 Estimation of the State Transition Equation

2.2.1 Data

The time span of the data in this section is from November 2017 to February 2020, pollution seasons only. Historical daily city-average $PM_{2.5}$ concentration was retrieved from *aqistudy.cn*, which is the simple average of hourly $PM_{2.5}$ concentration readings from the local EPA. Historical meteorological conditions from the Shuangliu and Wenjiang weather station were obtained from National Oceanic and Atmospheric Administration (NOAA). Ta-

ble 4 below shows the summary statistics of the data used in this section[6].

Variable	Obs	Mean	Std. Dev.	Min	Max
Daily city-average $PM_{2.5}$ concentration ($\mu g/m^3$)	352	68.6	33.6	9	201
Daily average wind speed (knots)	355	2.61	1.19	0.7	12.5
Daily maximum wind gust (knots)	355	5.64	2.01	2.2	16.5
Daily precipitation amount (inches)	346	.0144	.0369	0	0.22
Daily average temperature (Fahrenheit)	355	48.5	6.06	34.4	65.05
Daily high temperature (Fahrenheit)	355	54.9	7.14	37.05	75.4
Daily low temperature (Fahrenheit)	355	41.6	6.87	24.7	59.9
Daily average atmospheric pressure (mb)	355	958	4.69	946.4	971
Daily average dew point (Fahrenheit)	355	41.0	6.92	20.9	58.1

Table 4: Summary statistics of daily city-average $PM_{2.5}$ concentration and meteorological variables

2.2.2 Non-linear Least Square Estimation

Equation (22) shows the state transition equation again[7].

$$s_{t+1} = (1 - r_t)(c_t + s_t) \tag{22}$$

In Equation (22), daily city-average $PM_{2.5}$ concentration can serve as the proxy for s_{t+1} and s_t[8]. r_t is assumed to be a function of only contemporaneous meteorological conditions. Al-

[6]There are two weather stations within the city boundary, the Shuangliu station and Wenjiang station. Whenever possible, I use the average of readings at these two stations to represent the whole city. For wind direction, hourly wind speed, daily precipitation amount, and daily average atmospheric pressure, however, the Shuangliu station has either no observation or too many missing values. In these cases, I simply use the observations at the Wenjiang Station. For this reason, daily average wind speed, which includes information from the Shuangliu station as well, is not simply a linear combination of hourly wind speed.

[7]In reality, the pollution dissipation on day t does not fully operate on c_t as it is the pollution emitted during day t and thus has less than a full day of dissipation. On the other hand, dissipation is not necessarily constant throughout the day. To account for these two factors, we can add a multiplier γ and the state transition equation becomes $s_{t+1} = (1 - \frac{1}{\gamma}r_t)c_t + (1 - r_t)s_t$, where $\gamma \geq 1$. However, the exact value of γ is unknown. It can be explored but such analysis is beyond the scope of this chapter. We might expect γ to be different from one, which will bias the coefficients but will not affect the prediction in the empirical equations I explore in this chapter. Since prediction is the only concern of this chapter, the exact value of γ has no implications on the conclusions. Therefore, I keep using the original state transition equation (i.e. $\gamma = 1$) as in Equation (22) in the subsequent analysis.

[8]Note that this is an imperfect measure of the pollution stock at the beginning of a day for two reasons. First, c_t is distributed throughout day t. So the simple average of hourly $PM_{2.5}$ concentrations only incor-

though past meteorological conditions might also affect the current pollution stock turnover process, such effect is already captured by the contemporaneous pollution stock (s_t). c_t, however, is unobserved. Conceptually, c_t should be largely determined by factors related to human activity. It is convenient to treat c_t as a constant if it has little variation so that the estimation can be greatly simplified. To see if this is feasible, a simple analysis is performed to assess the variability of c_t. In Table 5 below, I present two pairs of comparisons as examples.

Date	Weekday/ Weekend	Treatment Status	Contemporaneous Temperature	Contemporaneous Wind	1-day Lagged Wind	2-day Lagged Wind	Contemporaneous Precipitation	1-day Lagged Precipitation	Incremental PM2.5
Dec. 15, 2018	weekend	1	43.0	1.7	1.1	1.5	0	0	-15
Dec. 16, 2018	weekend	1	44.0	1.6	1.7	1.1	0	0	+38
Nov. 22, 2017	weekday	0	50.0	2.9	2.9	2.5	0	0	-14
Nov. 23, 2017	weekday	0	45.1	2.6	2.9	2.9	0	0	+36

Table 5: Comparisons of similar days

Within each pair, I compare the incremental (daily city-average) $PM_{2.5}$ concentration which is defined as the difference of the $PM_{2.5}$ concentration between the next day and the current day. Note that the incremental $PM_{2.5}$ concentration is not equivalent to c_t because incremental $PM_{2.5}$ concentration is jointly determined by both c_t and pollution dissipation of day t. Therefore, I compare days with similar important meteorological conditions to somewhat control for the differences in dissipation. As human activity should be the main driver of c_t, I compare the days of the same month-of-year, in the same pollution season, with the same weekday/weekend and treatment (HAPEP alert) status of day t (=1 if treated, =0 otherwise). Even conditional on these factors, however, the comparisons of similar days show dramatically different incremental $PM_{2.5}$ concentrations, suggesting treating c_t as a constant is inappropriate. I thus treat c_t as a function of variables related to human activity. Below is the econometric model I estimate.

porates part of c_t. For example, if we assume c_t is uniformly distributed during day t, then only half of c_t is incorporated into this measure. Second, dissipation is not necessarily constant throughout the day, which also biases this measure.

$$s_{t+1} = g(\mathbf{weather_t}) * [h(\mathbf{activity_t}) + s_t] + \epsilon_{t+1} \tag{23}$$

In Equation (23), $(1 - r_t)$ is estimated by $g(\cdot)$ while c_t is estimated by $h(\cdot)$. $\mathbf{weather_t}$ is a vector of contemporaneous meteorological conditions. $\mathbf{activity_t}$ is a vector of time fixed effects and treatment status that are likely to be correlated with daily human activity. According to the definition of r_t, we have $r_t \in [0, 1]$, and thus $(1 - r_t) \in [0, 1]$. To confine the estimated r_t to be on the unit interval, it is necessary to choose a specific functional form for $g(\cdot)$. In Chapter 1, I showed that wind speed is the main driver of pollution dispersion. The distribution of wind speed which is illustrated in Figure 7.

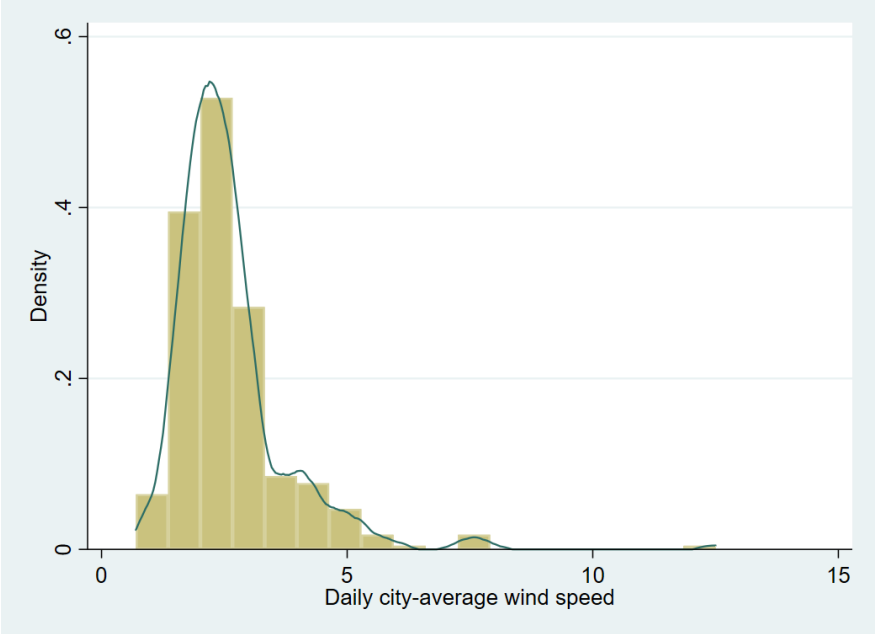

Figure 7: Distribution of daily city-average wind speed during three pollution seasons

For easier interpretation, I choose a logit functional form instead of the probit function. I treat $h(\cdot)$ as a linear function for simplicity. Below is the econometric model with explicit functional forms and variables.

$$PM2.5_{t+1} = \frac{exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})}{1 + exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})} * [(\alpha_2 + \beta_{\mathbf{f}}\mathbf{T_t} + \beta_s \cdot treat_t) + PM2.5_t] + \epsilon_{t+1} \quad (24)$$

In Equation (24), $PM2.5_t$ denotes the daily city-average $PM_{2.5}$ concentration of day t. $\mathbf{X_t}$ stands for a vector of meteorological conditions of day t, including daily average wind direction dummy variables[9], daily average wind speed, daily maximum wind gust, hourly wind speed, daily precipitation amount, daily average temperature, daily high and low temperature, daily average atmospheric pressure, and daily average dew point. $\mathbf{T_t}$ represents a vector of time fixed effects, including day-of-week dummy variables, workday dummy variable, month-of-year dummy variables, and pollution season dummy variables. $treat_t$ denotes the treatment status of day t. ϵ is the error term. I use the non-linear least square technique to estimate Equation (24). The regression returns an R-squared value of 0.964 and an adjusted R-squared value of 0.956. This suggests an excellent fit of the model and lends support to the setup of the state transition equation. Table 6 shows part of the regression results.

Variable	OLS Estimate
Maximum wind speed	-.10**
	(.04)
First pollution season	10.76**
	(4.61)
Second pollution season	.52
	(3.83)

Table 6: Partial results from OLS estimation of Equation (24)

In Table 6, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. Despite obvious multicollinearity between weather variables, the point estimation of daily maximum wind gust is both statistically significant at level 5 percent and has an expected negative sign. The regression results also show that the

[9]The detailed method of calculating daily average wind direction is described in the appendix (Section 2.6.1).

first pollution season dummy variable is significantly different from the third pollution season dummy variable, while the second pollution season dummy variable is not significantly different from the third pollution season dummy variable. This is consistent with previous observations in Chapter 1 that the pollution level in the first pollution season is higher than in the other two.

It is also important to learn how much variation in the next day's $PM_{2.5}$ concentration is explained by either weather variables or contemporaneous $PM_{2.5}$ concentration in the context of Equation (24). The central role of r_t in the design of the theoretical model will be challenged if the contemporaneous $PM_{2.5}$ concentration alone explains most of the variation in the next day's $PM_{2.5}$ concentration. In other words, if Equation (24) is simply an AR(1) model then the modeling approach in Chapter 1 should be revised. To this end, I first set $\beta_{\mathbf{r}} = \mathbf{0}$ in Equation (24) and estimate the restricted model as in Equation (25) with non-linear least squares. The purpose is to see how much variation in the next day's $PM_{2.5}$ concentration can be explained without meteorological conditions.

$$PM2.5_{t+1} = \frac{exp(\alpha_1)}{1 + exp(\alpha_1)} * [(\alpha_2 + \beta_{\mathbf{f}}\mathbf{T_t} + \beta_s \cdot treat_t) + PM2.5_t] + \epsilon_{t+1} \tag{25}$$

The regression of Equation (25) yields an R-squared value of 0.634 and an adjusted R-squared value of 0.618. Then I regress the residual of the above regression, representing the variation in the next day's $PM_{2.5}$ concentration that is not explained by the right-hand side variables (i.e. contemporaneous $PM_{2.5}$ concentration, time fixed effects, and treatment status), on meteorological variables with OLS, as in Equation (26). The point is to see how much of the remaining variation from the regression of Equation (25) can be explained by meteorological conditions.

$$Residual\_25_{t+1} = \beta_0 + \beta_{\mathbf{r}}\mathbf{X_t} + e_{t+1} \tag{26}$$

Residual_25 denotes the residual of the regression of Equation (25). e denotes the error term. The regression of Equation (26) returns an R-squared value of 0.294 and an adjusted R-squared value of 0.260. Next, I replace the $PM2.5_t$ of Equation (24) with the mean of daily city-average $PM_{2.5}$ concentration and estimate Equation (27) with non-linear least squares. The purpose is to see how much variation in the next day's $PM_{2.5}$ concentration can be explained without contemporaneous $PM_{2.5}$ concentration.

$$PM2.5_{t+1} = \frac{exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})}{1 + exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})} * [(\alpha_2 + \beta_{\mathbf{f}}\mathbf{T_t} + \beta_s \cdot treat_t) + \overline{PM2.5_t}] + \epsilon_{t+1} \tag{27}$$

The regression of Equation (27) yields an R-squared value of 0.936 and an adjusted R-squared value of 0.924. Compared to the R-squared value of 0.634 and adjusted R-squared value of 0.618 of the regression of Equation (25), this is already clear evidence that the dissipation factor explains more variation in the next day's $PM_{2.5}$ concentration than the contemporaneous $PM_{2.5}$ concentration does. To complete this exercise, I regress the residual of the above regression, representing the variation in the next day's $PM_{2.5}$ concentration that is not explained by the right-hand side variables (i.e. meteorological conditions, time fixed effects, and treatment status), on contemporaneous $PM_{2.5}$ concentration with OLS. The point is to see how much of the remaining variation from the regression of Equation (27) can be explained by contemporaneous $PM_{2.5}$ concentration.

$$Residual\_27_{t+1} = \beta_0 + \beta_p \cdot PM2.5_t + d_{t+1} \tag{28}$$

Residual_27 denotes the residual of the regression of Equation (27). d denotes the error term. The regression of Equation (28) returns an R-squared value of 0.131 and an adjusted R-squared value of 0.129. Table 7 summarizes the results of this exercise.

	Explanatory Variables	R^2	Adjusted R^2
Regression (26)	weather variables	0.294	0.260
Regression (28)	contemporaneous PM2.5 concentration	0.131	0.129

Table 7: Comparisons of the regression of Equation (26) and (28)

By the fact that the R-squared value and adjusted R-squared value of the regression of Equation (26) are higher than those of the regression of Equation (28), it is safe to conclude that the weather variables explain a significant amount of variation in leading $PM_{2.5}$ concentration, lending support to the central role of r_t in pollution stock turnover. As a robustness check, I test the hypothesis that $\beta_{\mathbf{r}}$ are jointly zero. I reject this null hypothesis at the $p \leq 5\%$ level.

2.2.3 Estimated Historical $\mathbf{r_t}$

With the estimation of Equation (24), I generate an estimated dissipation factor variable according to Equation (29) below.

$$\widehat{r_t} = 1 - \frac{exp(\widehat{\alpha_1} + \widehat{\beta}_{\mathbf{r}}\mathbf{X_t})}{1 + exp(\widehat{\alpha_1} + \widehat{\beta}_{\mathbf{r}}\mathbf{X_t})} \tag{29}$$

Figure 8 is a histogram of the estimated dissipation factors with a kernel density estimate.

The estimated dissipation factors are bounded between zero and one by construction. The

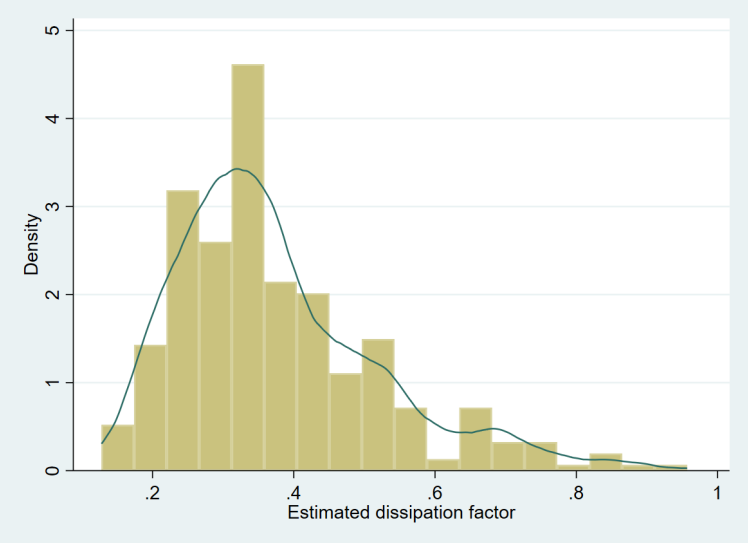

Figure 8: Histogram of estimated dissipation factors with a kernel density estimate

kernel density estimate shows a peak at the lower end and a long tail extending to the higher end, which is consistent with the pattern of daily average wind speed observed in Figure 7. I plot the estimated dissipation factors and city-average $PM_{2.5}$ concentration over time in Figure 9.

2.2.4 Simplifying the Estimation

In Section 2.2.2, I have shown that Equation (24) does well in explaining the next day's $PM_{2.5}$ concentration. In this section, I assess the performance of parsimonious models with fewer explanatory weather variables. As we have learned in Chapter 1, wind speed is the most important driver of the variation in $PM_{2.5}$ concentration, I first explore a model where only variables related to wind speed are included in $g(\cdot)$. Next, I investigate an even more parsimonious model which has daily average wind speed as the only explanatory variable for r_t. Equations (30) and (31) are simplified models.

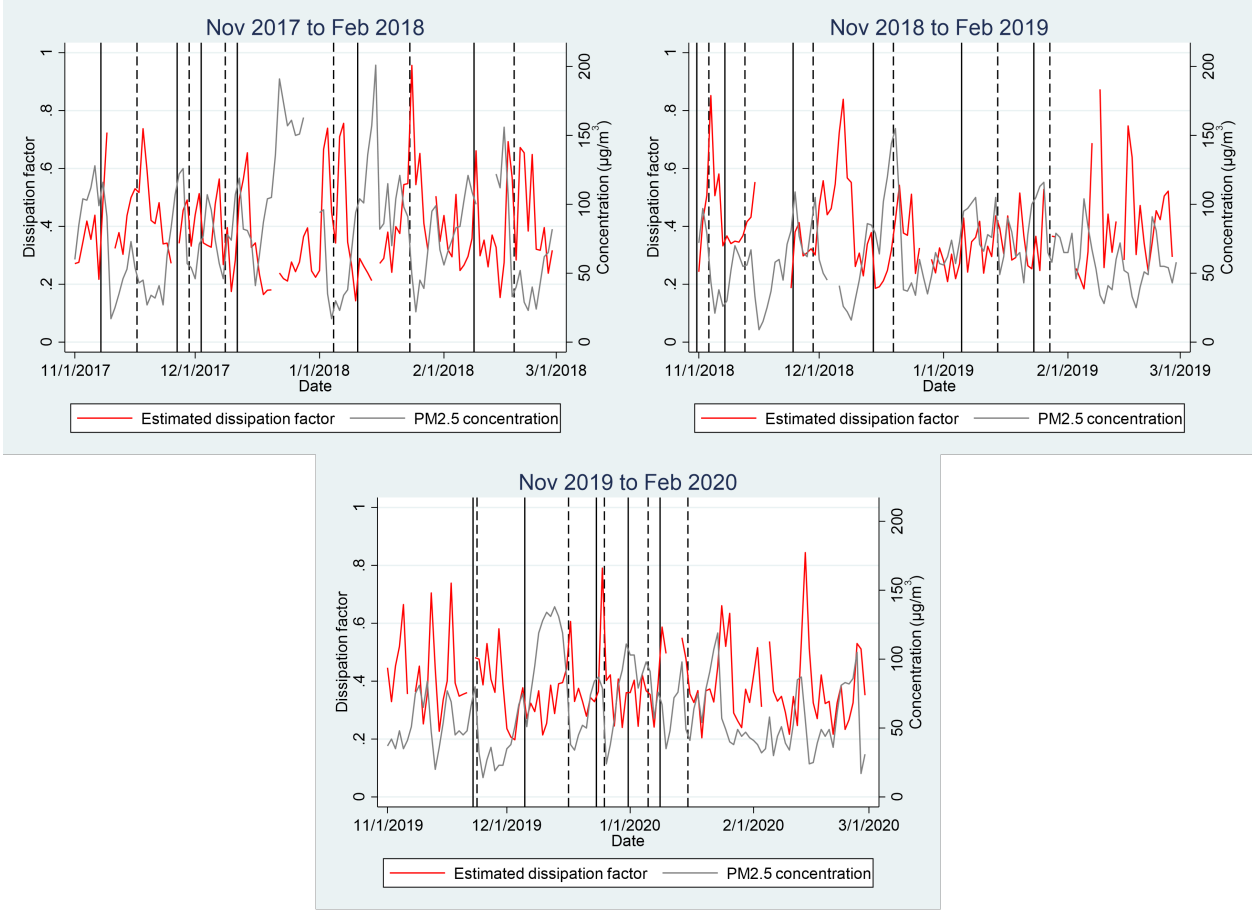

Figure 9: Estimated dissipation factors and city-average $PM_{2.5}$ concentration during three pollution seasons

$$PM2.5_{t+1} = \frac{exp(\alpha_1 + \beta_{wd} \cdot wdsp_t + \beta_{mx} \cdot mxspd_t + \beta_{\mathbf{h}} \cdot \mathbf{hwdsp_t})}{1 + exp(\alpha_1 + \beta_{wd} \cdot wdsp_t + \beta_{mx} \cdot mxspd_t + \beta_{\mathbf{h}} \cdot \mathbf{hwdsp_t})} * \\ [(\alpha_2 + \beta_{\mathbf{f}} \mathbf{T_t} + \beta_s \cdot treat_t) + PM2.5_t] + \epsilon_{t+1} \tag{30}$$

$$PM2.5_{t+1} = \frac{exp(\alpha_1 + \beta_{wd} \cdot wdsp_t)}{1 + exp(\alpha_1 + \beta_{wd} \cdot wdsp_t)} * [(\alpha_2 + \beta_{\mathbf{f}} \mathbf{T_t} + \beta_s \cdot treat_t) + PM2.5_t] + \epsilon_{t+1} \tag{31}$$

In Equation (30) and (31), $wdsp_t$ denotes daily average wind speed of day t. $mxspd_t$ denotes maximum wind gust of day t. $\mathbf{hwdsp_t}$ denotes hourly wind speeds of day t. I refer to Equation (24), (30), and (31) as the full, restricted, and simple model, respectively. Table 8

demonstrates a comparison of weather variables included in each model.

	Full	Restricted	Simple
Daily average temperature	yes	no	no
Daily high temperature	yes	no	no
Daily low temperature	yes	no	no
Daily average atmospheric pressure	yes	no	no
Daily average dew point	yes	no	no
Daily precipitation amount	yes	no	no
Daily average wind direction	yes	no	no
Hourly wind speed	yes	yes	no
Daily maximum wind gust	yes	yes	no
Daily average wind speed	yes	yes	yes

Table 8: Comparison of weather variables included in the full, restricted, and simple model

I use the non-linear least square technique to estimate the restricted and simple model. Table 9 shows a comparison of the goodness of fit of these three models.

	Full	Restricted	Simple
Obs	322	331	343
R^2	0.964	0.956	0.944
Adjusted R^2	0.956	0.949	0.941

Table 9: Comparison of goodness of fit of the full, restricted, and simple model

The goodness of fit of the model only deteriorates marginally as explanatory weather variables are removed - even the simple model has a very high R-squared value of 0.944. I then utilize the restricted and simple model to estimate historical r_t in the same way as I did in Section 2.2.3 with the full model. Figure 10 shows histograms of the estimated dissipation factors with a kernel density estimate for each model.

The distribution of estimated dissipation factors appears similar among the three models. Noticeably, as fewer weather variables are included, the distribution shifts to the right slightly. I then plot the estimated dissipation factors by these three models over time, shown

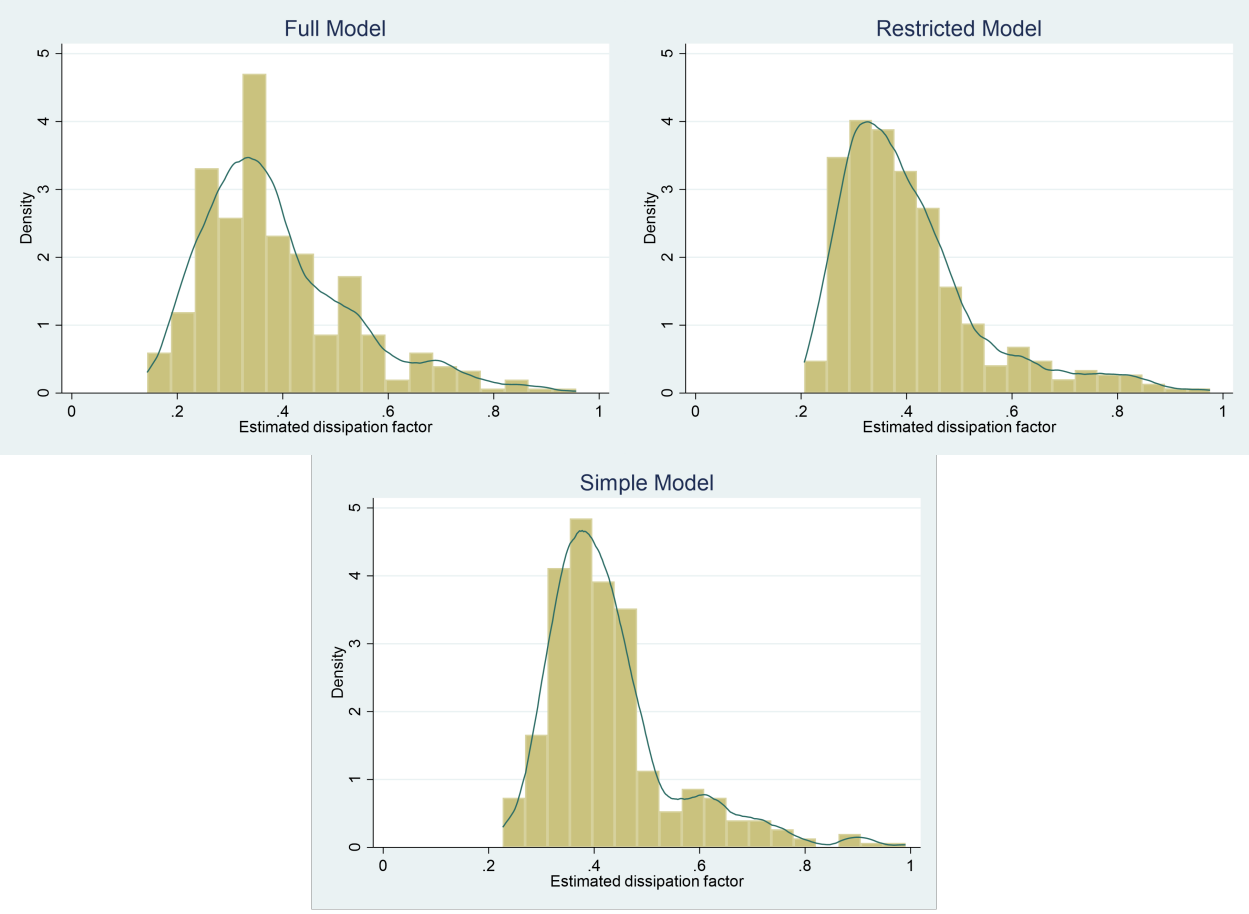

Figure 10: Histograms of estimated dissipation factors with a kernel density estimate for the full, restricted, and simple model

in Figure 11.

The vertical solid and dashed lines in Figure 11 represent the time when alert events began and ended, respectively. The red, blue, and green time series stand for the estimated dissipation factors by the full, restricted, and simple model, respectively. The three time series overlap relatively well except at the lower and higher ends. Figure 12 shows comparisons of dissipation factors estimated by the Full model and the alternative models in scatter plots.

In Figure 12, I plot the estimated dissipation factors by alternative models (vertical axis) against the estimated dissipation factors by the full model (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. Consistent with the

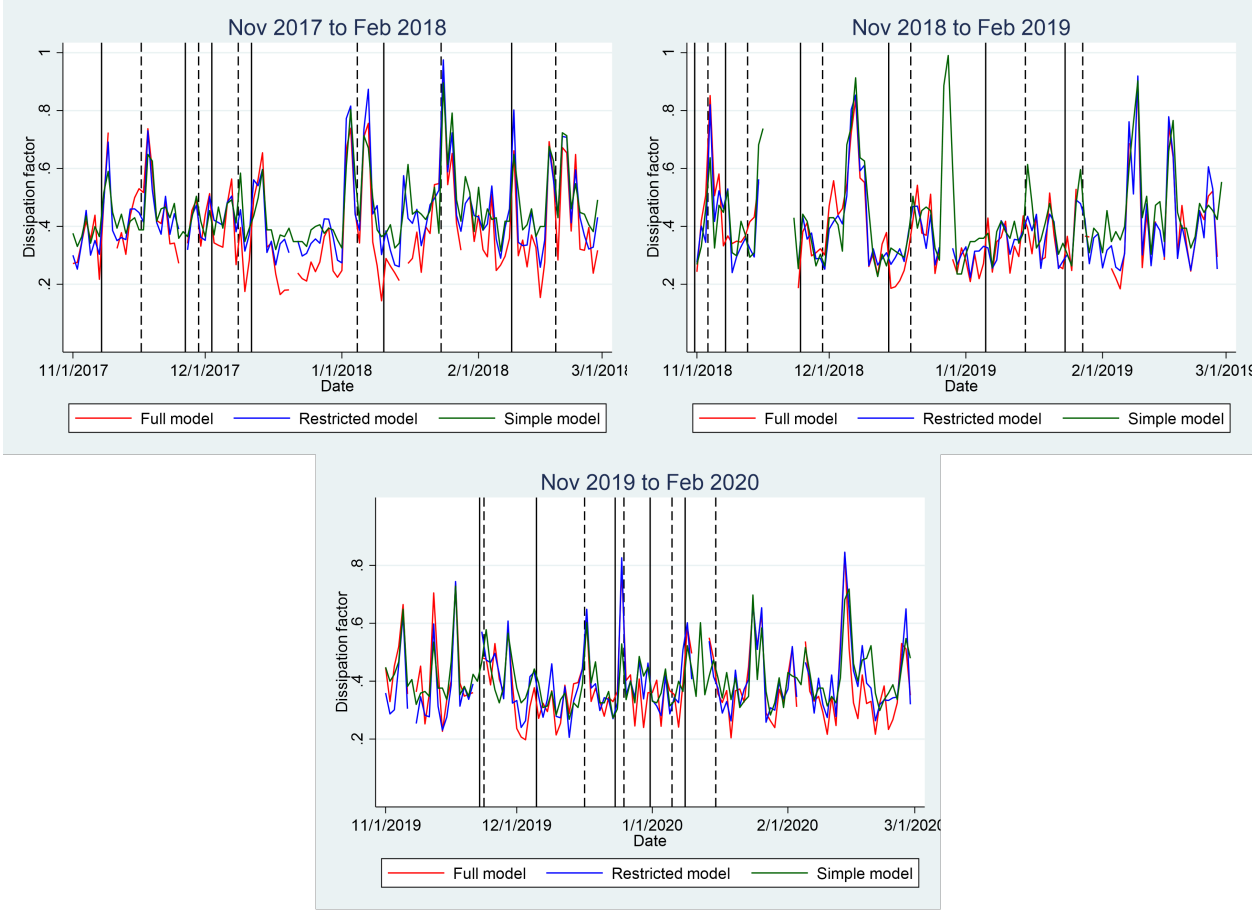

Figure 11: Estimated dissipation factors by the full, restricted, and simple model during three pollution seasons

observation in Figure 11, the estimated dissipation factors by alternative models are biased in the lower and higher ends: they are biased up in the lower end while biased down in the higher end. The biases are more pronounced for the simple model than for the restricted model. In conclusion, the restricted and simple models do a comparably good job of estimating dissipation factors at the middle range but predict less extreme values at both ends.

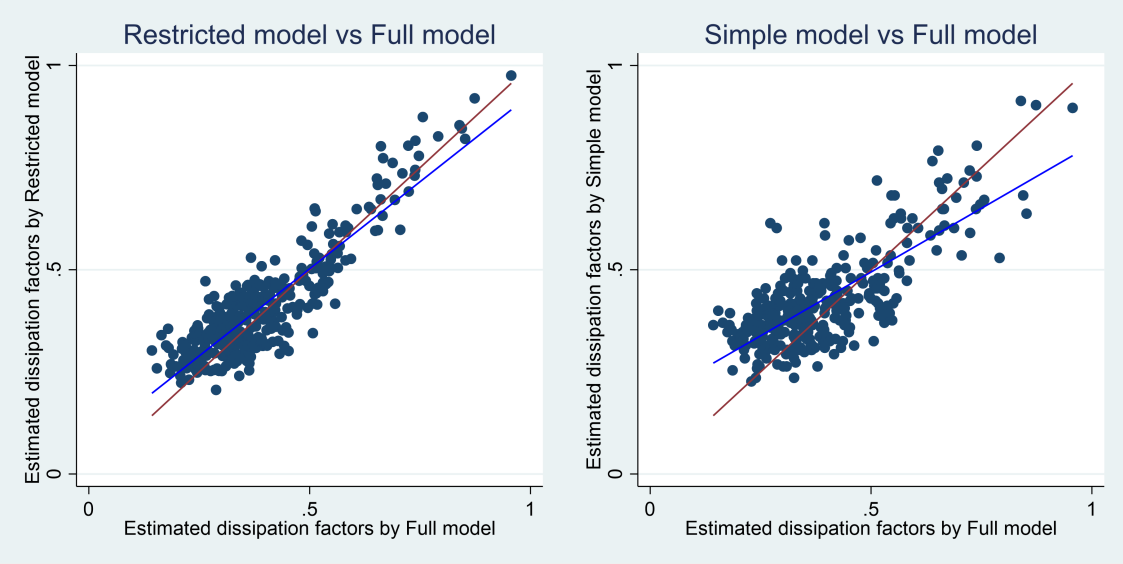

Figure 12: Comparisons of dissipation factors estimated by the full model and the alternative models in scatter plots

2.3 Forecastability of Dissipation Factors

2.3.1 Data

To assess the forecastability of dissipation factors, first I need to estimate the accuracy of predicted weather conditions in Chengdu. To this end, I obtained forecasted and observed weather data from the Wenjiang weather station of Chengdu. The data set includes wind speed, wind direction, and atmospheric pressure observed at 8pm local time. It also includes daily high and low temperature, and cumulative precipitation by 8pm local time. The prediction data consist of up to 7-day ahead forecasts of the same set of variables from November 1, 2021 to March 21, 2022. The time span of the observed weather data is from November 2, 2021 to March 28, 2022. Table 10 below shows the summary statistics of the data used in this section. Wind direction is reported by the direction from which the wind blows in the standard 16-direction format.

Variable	Obs	Mean	Std. Dev.	Min	Max
Predicted wind speed (m/s)	987	1.88	1.32	0.15	7.25
Observed wind speed (m/s)	147	1.33	0.50	0.70	3.15
Predicted atmospheric pressure (hPa)	987	956.5	6.0	934.7	970.4
Observed atmospheric pressure (hPa)	147	956.3	6.5	937.0	969.7
Predicted high temperature (Celsius)	987	13.5	5.3	3.9	29.8
Observed high temperature (Celsius)	147	13.6	5.5	3.9	29.4
Predicted low temperature (Celsius)	987	6.5	3.6	-1	19.1
Observed low temperature (Celsius)	147	6.5	4.0	-3.3	19.5
Predicted precipitation (mm)	987	0.61	1.08	0	7.9
Observed precipitation (mm)	147	0.46	1.25	0	9.5

Table 10: Summary statistics of predicted and observed weather conditions

2.3.2 Comparisons of Predicted and Observed Weather Conditions

I directly compare the predicted weather conditions to the corresponding observed weather conditions. High temperature, low temperature, and atmospheric pressure forecasts are accurate. As an example, Figure 13 illustrates the comparisons of the up to 7-day ahead predicted and observed high temperature.

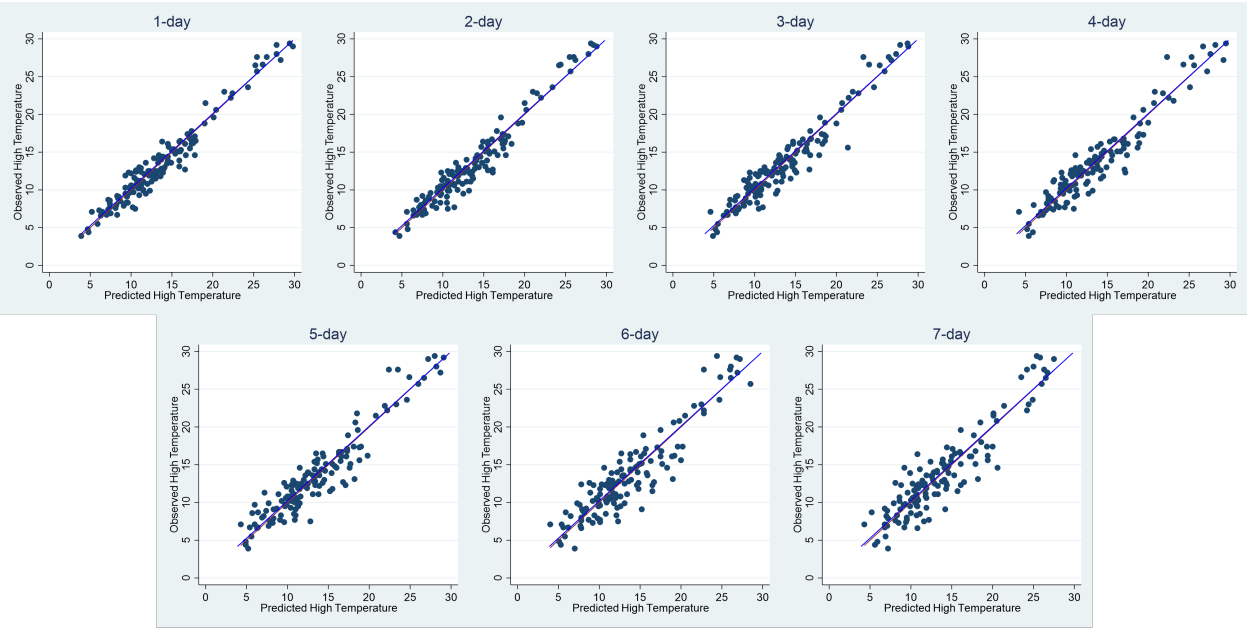

Figure 13: Comparison of the predicted and observed high temperature

In each graph of Figure 13, I plot the observed high temperature (vertical axis) against the

predicted high temperature (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines overlap very well in all graphs in Figure 13, suggesting the bias of prediction is minimal. Table 11 shows a quantitative assessment of the accuracy of the high temperature prediction.

Day	1	2	3	4	5	6	7
ρ	.97	.97	.96	.96	.94	.92	.92
P	.50	.48	.45	.47	.50	.45	.46

Table 11: Quantitative assessment of the accuracy of high temperature prediction

ρ denotes the correlation coefficient between the predicted and observed high temperature, P denotes the probability of the predicted high temperature being higher than the observed high temperature. As can be seen in Table 11, the correlation coefficient is close to one and the probability of over prediction is close to one-half, again indicating accurate predictions of the high temperature. Similar comparisons of the predicted and observed low temperature and atmospheric pressure can be found in the appendix (Section 2.6.2).

The prediction of wind speed and precipitation, however, is inaccurate. Figure 14 illustrates the comparisons of the up to 7-day ahead predicted and observed wind speed.

In each graph of Figure 14, I plot the observed wind speed (vertical axis) against the predicted wind speed (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. It is very obvious that the predicted wind speed is systematically larger than the observed wind speed. Table 12 shows a quantitative assessment of the accuracy of the wind speed prediction.

ρ denotes the correlation coefficient between the predicted and observed wind speed, P denotes the probability of the predicted wind speed being higher than the observed wind speed.

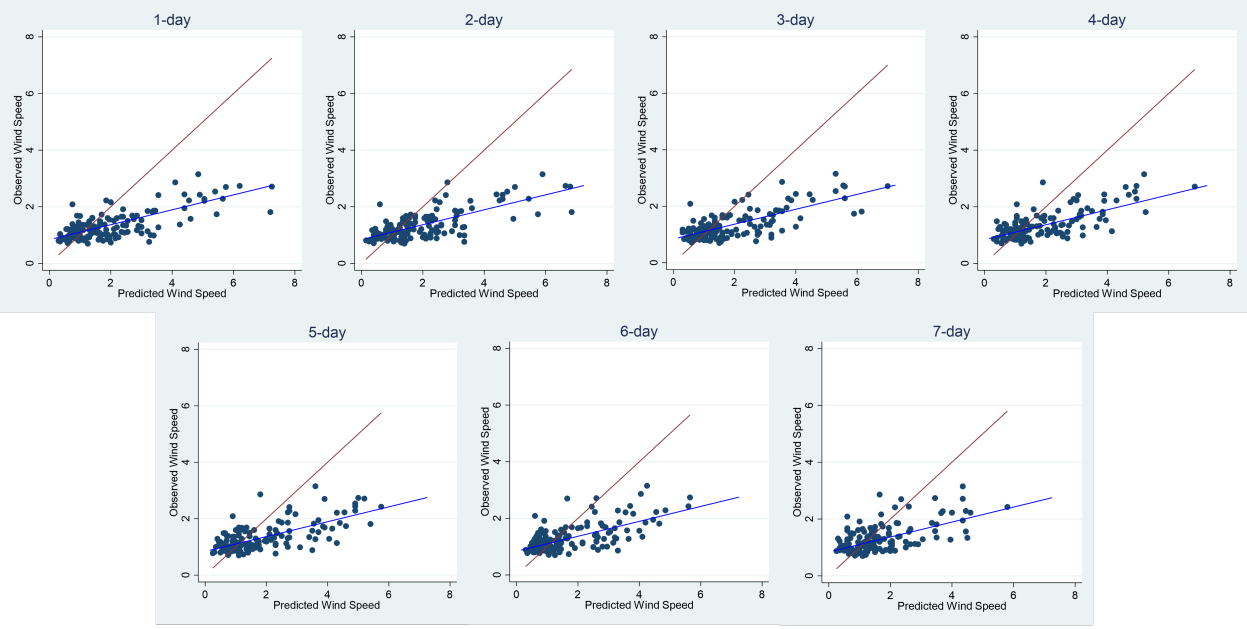

Figure 14: Comparison of the predicted and observed wind speed

Day	1	2	3	4	5	6	7
ρ	.72	.72	.72	.72	.67	.65	.60
P	.67	.67	.64	.63	.65	.62	.55

Table 12: Quantitative assessment of the accuracy of wind speed prediction

As expected, the correlation coefficient is lower than one. The probability of over prediction is also much higher than one-half, consistent with my observation from Figure 14.

Figure 15 illustrates the comparisons of the up to 7-day ahead predicted and observed precipitation.

In each graph of Figure 15, I plot the observed precipitation (vertical axis) against the predicted precipitation (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines diverge from each other markedly. Table 13 shows a quantitative assessment of the accuracy of the precipitation prediction.

ρ denotes the correlation coefficient between the predicted and observed precipitation, P

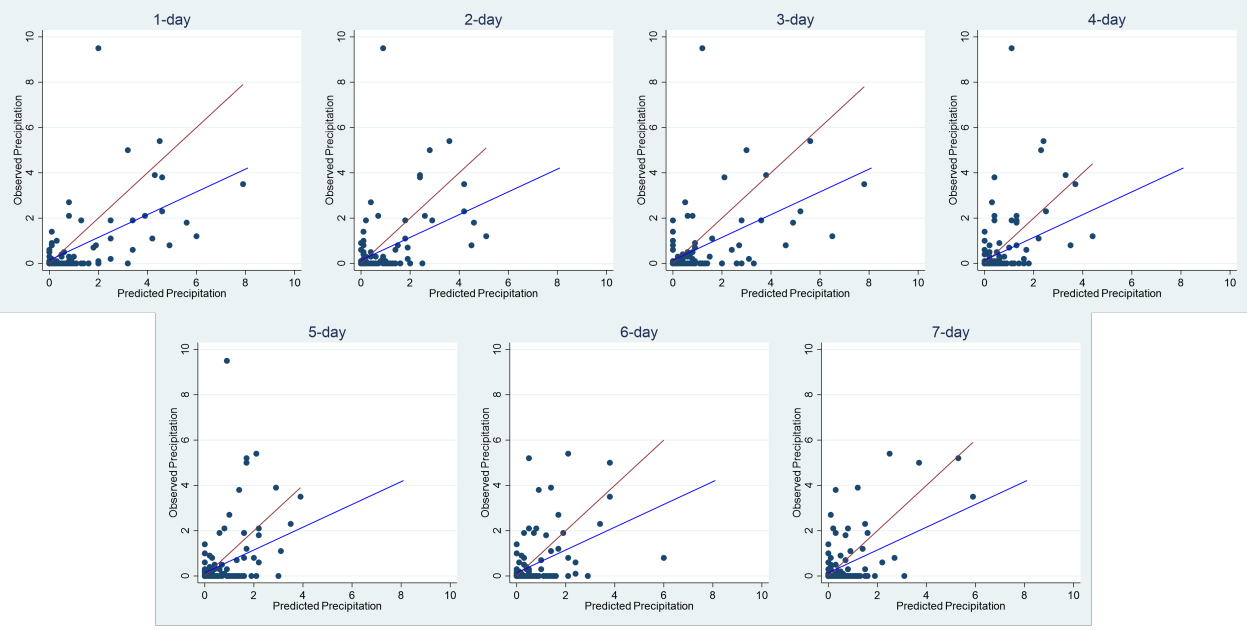

Figure 15: Comparison of the predicted and observed precipitation

Day	1	2	3	4	5	6	7
ρ	.58	.51	.50	.49	.49	.48	.60
P	.43	.38	.41	.35	.40	.43	.41

Table 13: Quantitative assessment of the accuracy of precipitation prediction

denotes the probability of the predicted precipitation being larger than the observed precipitation. As expected, the correlation coefficient is significantly lower than one. There is also a tendency to underestimate precipitation.

Although the wind speed prediction is biased, in Figure 14 the data points are tightly distributed along the fitted line. Once the bias is corrected, the predicted wind speed will be relatively reliable. Therefore, I attempt to correct the bias of wind speed prediction by making my own prediction.

2.3.3 Prediction of Wind Speed

In this section, I aim to use the available predicted weather conditions (i.e. wind speed, wind direction, atmospheric pressure, high temperature, low temperature, and precipitation) to forecast actual wind speed. First, I use a 5-fold cross validation approach for model selection. Table 14 lists the ten linear specifications with the lowest average mean squared error and two reference specifications. These models have a constant term unless stated otherwise.

Variables	**Average MSE**	**Std. Dev.**
wdsp (no constant)	.392	.047
wdsp	.152	.022
wdsp, *stp*	.144	.024
wdsp, *stp*, *high*	.144	.024
wdsp, *stp*, *low*	.144	.024
wdsp, *stp*, *prcp*	.144	.023
wdsp, *stp*, *high*, *low*	.144	.023
wdsp, *stp*, *high*, *prcp*	.143	.023
wdsp, *stp*, *low*, *prcp*	.144	.023
wdsp, *stp*, *high*, *low*, *prcp*	.143	.022
wdsp, *stp*, *high*, *prcp*, *direction*	.145	.023
wdsp, *stp*, *high*, *low*, *prcp*, *direction*	.145	.021

Table 14: Linear prediction models of wind speed with the lowest average mean squared error and two reference models

In Table 14, *wdsp*, *stp*, *high*, *low*, *prcp*, *direction* represent wind speed, atmospheric pressure, high temperature, low temperature, precipitation, and wind direction, respectively. The first two specifications have significantly larger average mean squared errors than the rest of the specifications. The others return very similar average mean squared errors so I decide to choose the simplest one with only *wdsp* and *stp*. The selection is consistent with physical science as to what determines wind speed. I also try adding higher orders of these two variables into the model. The results are shown in Table 15. All these models have a constant term.

Variables	Average MSE	Std. Dev.
$wdsp$, stp	.144	.024
$wdsp$, $wdsp^2$, stp	.143	.024
$wdsp$, $wdsp^2$, $wdsp^3$, stp	.142	.023
$wdsp$, $wdsp^2$, $wdsp^3$, $wdsp^4$, stp	.142	.022
$wdsp$, $wdsp^2$, $wdsp^3$, $wdsp^4$, $wdsp^5$, stp	.143	.022
$wdsp$, stp, stp^2	.143	.023
$wdsp$, stp, stp^2, stp^3	.143	.023
$wdsp$, stp, stp^2, stp^3, stp^4	.142	.023
$wdsp$, stp, stp^2, stp^3, stp^4, stp^5	.142	.023

Table 15: Prediction models of wind speed with higher orders

According to Table 15, the addition of higher orders only marginally improves the average mean squared error. Therefore, I choose the linear specification. Next, I assess the performance of my prediction model, again with a 5-fold cross validation approach. Figure 16 illustrates a comparison of the predicted (by my prediction model) and observed wind speed.

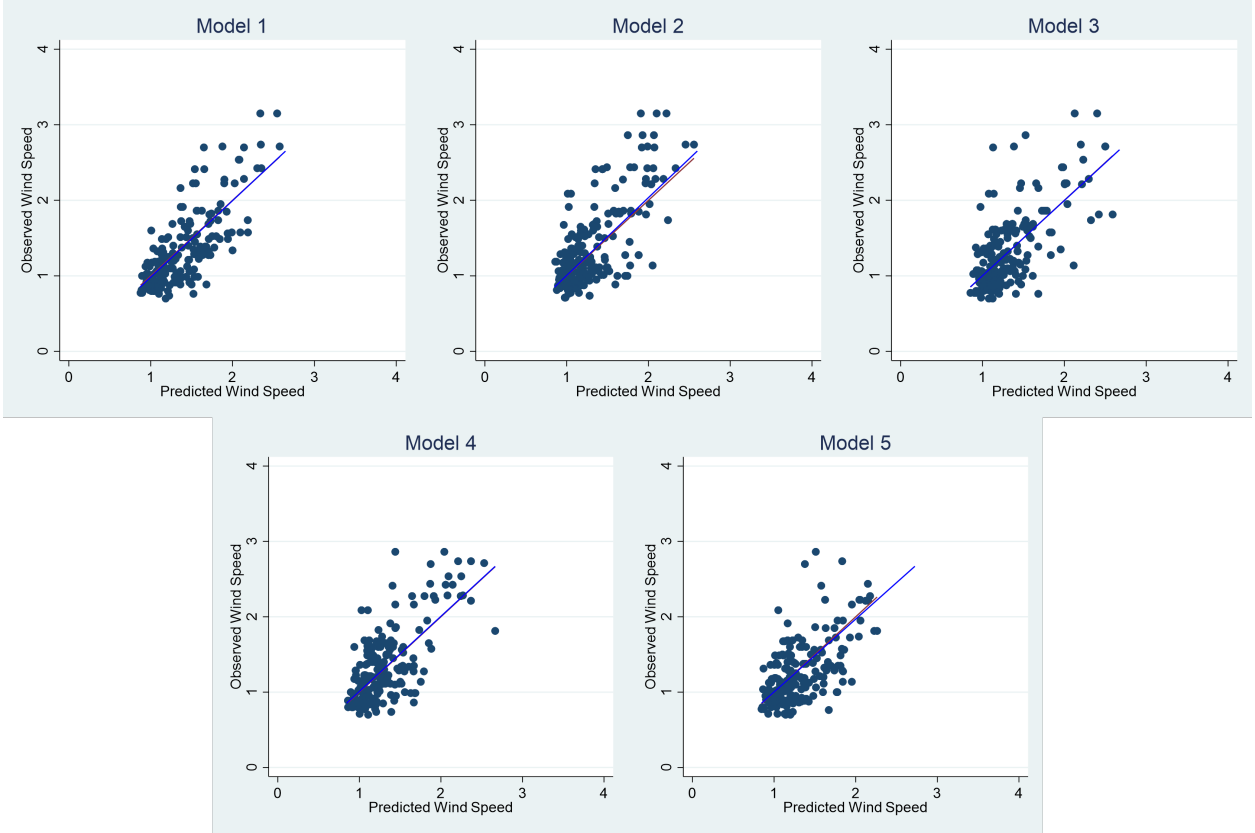

Figure 16: Comparison of predicted (by my prediction model) and observed wind speed

In each graph of Figure 16, I plot the observed wind speed (vertical axis) against the predicted wind speed (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines overlap very well in all graphs in Figure 16, suggesting that the bias of prediction is minimal. This prediction model provides much better wind speed forecasts than those from the weather station. Table 16 presents a quantitative assessment of the accuracy of the wind speed prediction (by my prediction model).

Model	1	2	3	4	5
ρ	.70	.70	.70	.70	.70
P	.55	.55	.55	.55	.55

Table 16: Quantitative assessment of the accuracy of wind speed prediction (by my prediction model)

ρ denotes the correlation coefficient between the predicted and observed wind speed, P denotes the probability of the predicted wind speed being higher than the observed wind speed. As can be seen in Table 16, although the correlation coefficient is only 0.7, the probability of over prediction is close to one-half, indicating that my prediction model is able to mostly correct the systematic error of the wind speed forecasts from the weather station. This is important for the forecasts of dissipation factors due to the crucial role of wind speed. Alternatively, I use a Lasso approach for model selection and prediction. The results are consistent with the above analysis.

2.3.4 Prediction of Dissipation Factors

In this section, I predict dissipation factors with the predicted weather conditions. Then I make a comparison of the predicted dissipation factors with realized dissipation factors to infer the forecastability of dissipation factors. To this end, I first estimate the following model.

$$PM2.5_{t+1} = \frac{exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})}{1 + exp(\alpha_1 + \beta_{\mathbf{r}}\mathbf{X_t})} * [(\alpha_2 + \beta_{\mathbf{f}}\mathbf{T_t} + \beta_s \cdot treat_t) + PM2.5_t] + \eta_{t+1} \tag{32}$$

In Equation (32), $PM2.5_t$ denotes the daily city-average $PM_{2.5}$ concentration of day t. $\mathbf{X_t}$ stands for a vector of meteorological conditions of day t, including daily average wind direction dummy variables, daily average wind speed, daily precipitation amount, daily high and low temperature, daily average atmospheric pressure, and daily average dew point. $\mathbf{T_t}$ represents a vector of time fixed effects, including day-of-week dummy variables, workday dummy variable, and month-of-year dummy variables. $treat_t$ denotes the treatment status of day t. η is the error term. I use non-linear least squares to estimate Equation (32). The regression returns an R-squared value of 0.938 and an adjusted R-squared value of 0.935. With the estimates of Equation (32), I calculate predicted dissipation factors with predicted weather conditions, and calculate realized dissipation factors with observed weather conditions according to Equation (33).

$$\widehat{r_t} = 1 - \frac{exp(\widehat{\alpha_1} + \widehat{\beta}_{\mathbf{r}}\mathbf{X_t})}{1 + exp(\widehat{\alpha_1} + \widehat{\beta}_{\mathbf{r}}\mathbf{X_t})} \tag{33}$$

Figure 17 illustrates the comparisons of the up to 7-day ahead predicted and realized dissipation factors.

In each graph of Figure 17, I plot the realized dissipation factor (vertical axis) against the predicted dissipation factor (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines overlap well in all graphs, suggesting the bias of prediction is minimal. Table 17 shows a quantitative assessment of the accuracy of the dissipation factor prediction.

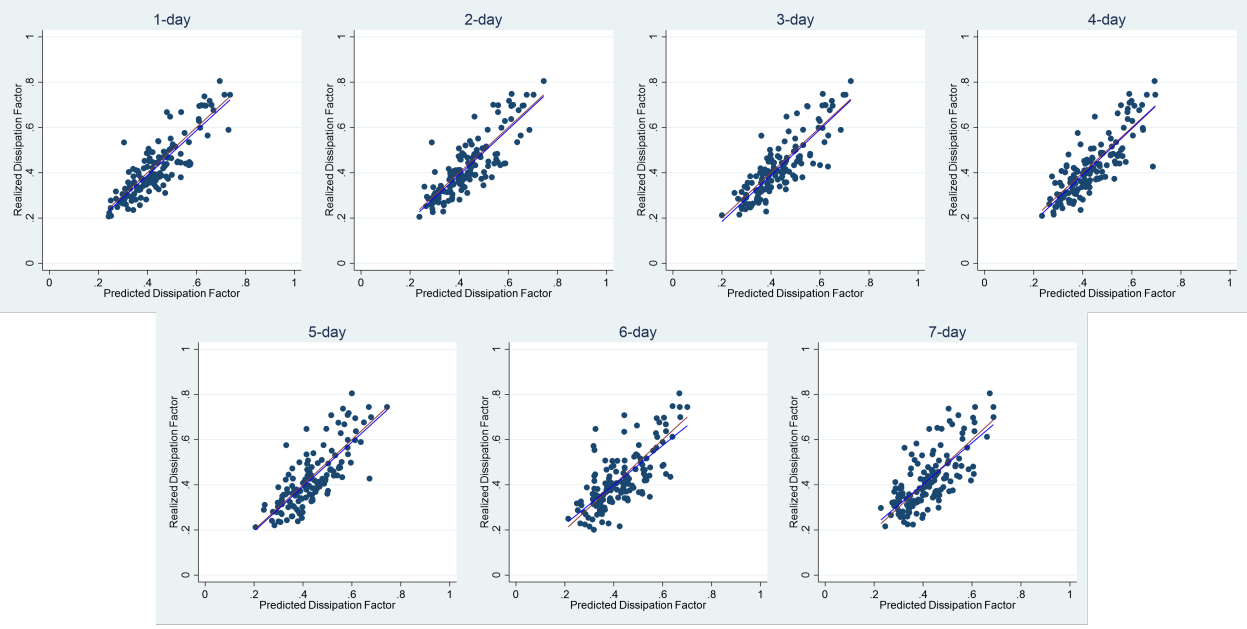

Figure 17: Comparison of predicted and realized dissipation factors

Day	1	2	3	4	5	6	7
ρ	.85	.84	.84	.83	.78	.72	.72
P	.57	.55	.59	.61	.56	.56	.55
MAE	.054	.055	.056	.058	.063	.067	.069

Table 17: Quantitative assessment of the accuracy of dissipation factor prediction

ρ denotes the correlation coefficient between the predicted and observed dissipation factor, **P** denotes the probability of the predicted dissipation factor being larger than the observed dissipation factor, **MAE** represents mean absolute error. As can be seen in Table 17, the correlation coefficient is relatively close to one and the probability of over prediction is close to one-half. The mean absolute error of prediction is on the level of 0.05 to 0.07. These results suggest that dissipation factors in the near future can be predicted with high accuracy.

2.4 Simulation

In this section, I estimate the gains that can be achieved by altering the timing of historical interventions according to both existing conditions and expected upcoming weather patterns. I select the period of November 23, 2018 to December 28, 2018 which ends with a very high dissipation factor (0.99) so that any scrap pollutant can be ignored. During this period, yellow alerts were issued from November 25 to 29, and December 15 to 19. I keep the total number of alert days the same and reassign them to days with the highest marginal health damage of pollution. In Equation (34) I define the marginal damage of emitting pollution on day t ($t < T$), assuming all pollution clears up on day T.

$$Marginal\,Damage = (ID_t + \sum_{n=t+1}^{t+6} DD_n) + (1-r_t)(ID_{t+1} + \sum_{n=t+2}^{t+7} DD_n) + ... + \prod_{s=t}^{T-1}(1-r_s)(ID_T + \sum_{n=T+1}^{T+6} DD_n) \tag{34}$$

In Equation (34), ID_t denotes the instantaneous damage on day t, DD_n represents the delayed damage on day n. In Chapter 3, I empirically examine the relationship between bronchitis hospital visits and (contemporaneous and up to the 7th lagging) $PM_{2.5}$ pollution levels. This analysis is explained in detail in Section 3.3. Briefly, I find $PM_{2.5}$ pollution levels up to 6 days ago have an effect on contemporaneous hospitalization that is statistically significant at level 5 percent. The estimated coefficients on the contemporaneous and lagging $PM_{2.5}$ pollution levels are used to specify the instantaneous damage (ID) and delayed damages (DD) in Equation (34), respectively. I assume both instantaneous damage and delayed damages are constant over time - delayed damages only depend on the length of the delay. Therefore, one unit of $PM_{2.5}$ pollution emitted on day t causes instantaneous damage on day t and delayed damages on day $t+1$ to $t+6$. Then, after dissipation, $(1-r_t)$ unit of $PM_{2.5}$ pollution transitions into day $t+1$, again causing instantaneous damage on day $t+1$ and delayed damages on day $t+2$ to $t+7$. The process goes on until day T. According to

Equation (34) and the estimated dissipation factors in Section 2.2.3, I calculate the marginal damage of emitting pollution for each day during this period. Figure 18 demonstrates the dissipation factors (in red), marginal damage (in blue), and $PM_{2.5}$ concentration (in grey) during this period with yellow arrows indicating the days with alert events. When an alert is issued, a set of measures to reduce pollution emissions are implemented including driving restrictions for vehicles and production suspension for firms.

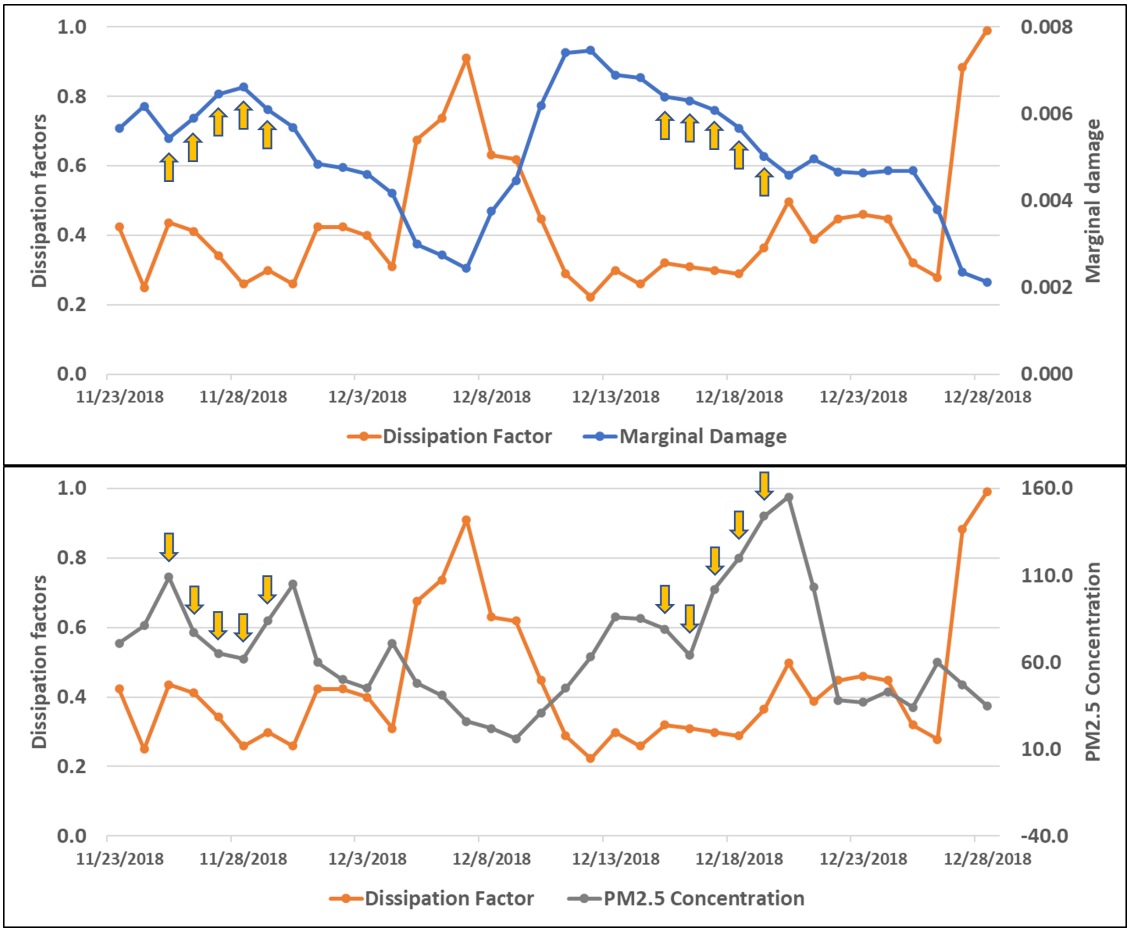

Figure 18: Dissipation factor, marginal damage, and $PM_{2.5}$ concentration with actual alert events

In Figure 18, the left vertical axis of both graphs is for dissipation factor. The right vertical axis of the upper and lower graph is for marginal damage and $PM_{2.5}$ concentration, respectively. From Figure 18 we can see the actual alert events happened generally, although not strictly, on days with the highest $PM_{2.5}$ concentration instead of the highest marginal

damage. Below I reassign the alert events to days with the highest marginal damage, shown in Figure 19.

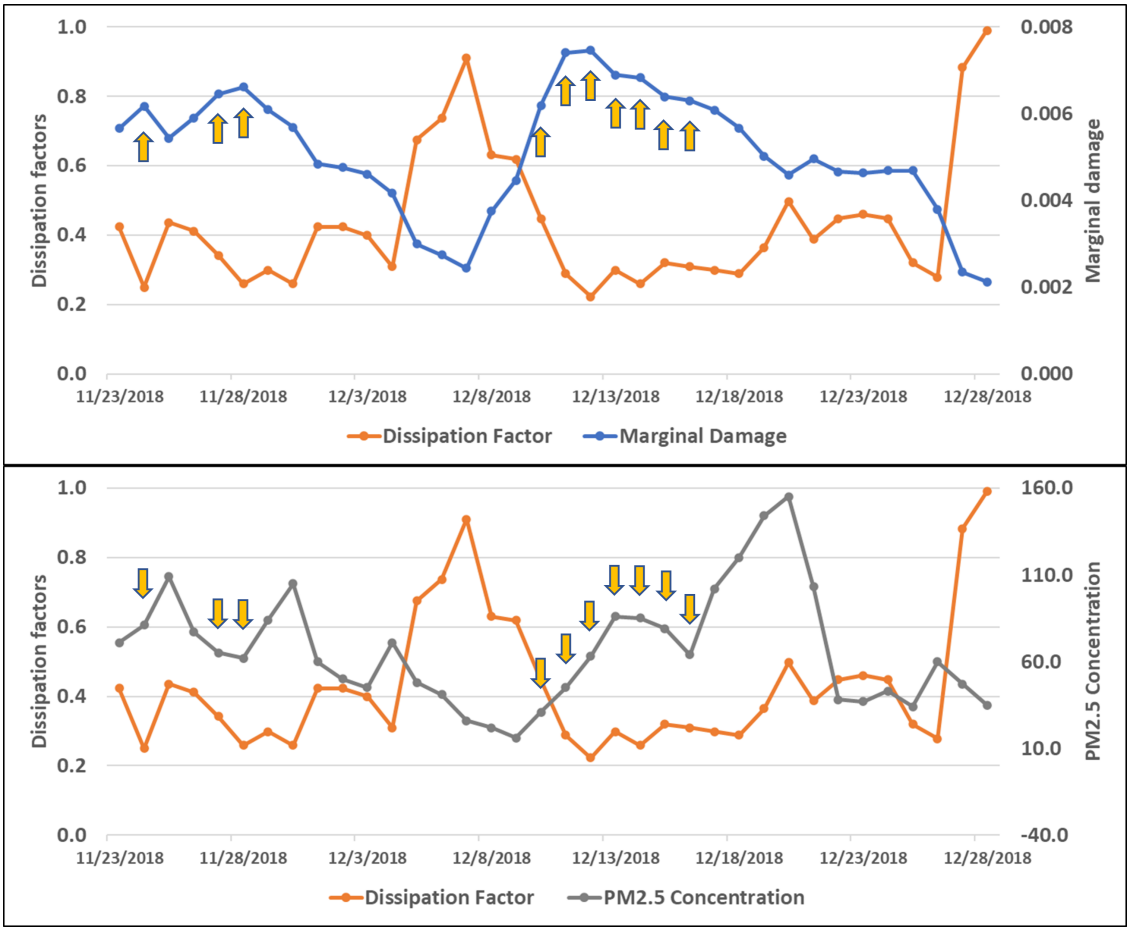

Figure 19: Dissipation factor, marginal damage, and $PM_{2.5}$ concentration with alternative alert events

Next, I calculate the $PM_{2.5}$ concentration path under the alternative alert events. I arbitrarily assume an alert event reduces the daily emission amount by a units. Therefore, if a day is actually untreated but alternatively treated, the daily emission amount is reduced by a units in the alternative world. If a day is actually treated but alternatively untreated, the daily emission amount is increased by a units in the alternative world. If a day is both actually and alternatively treated or untreated, the daily emission amount remains the same in the alternative world. Assuming $a = 10$, according to this rule and the estimated state transition equations in Section 2.2.3, I am able to calculate the $PM_{2.5}$ concentration path

under alternative treatments as well as if no interventions happened at all, shown in Figure 20.

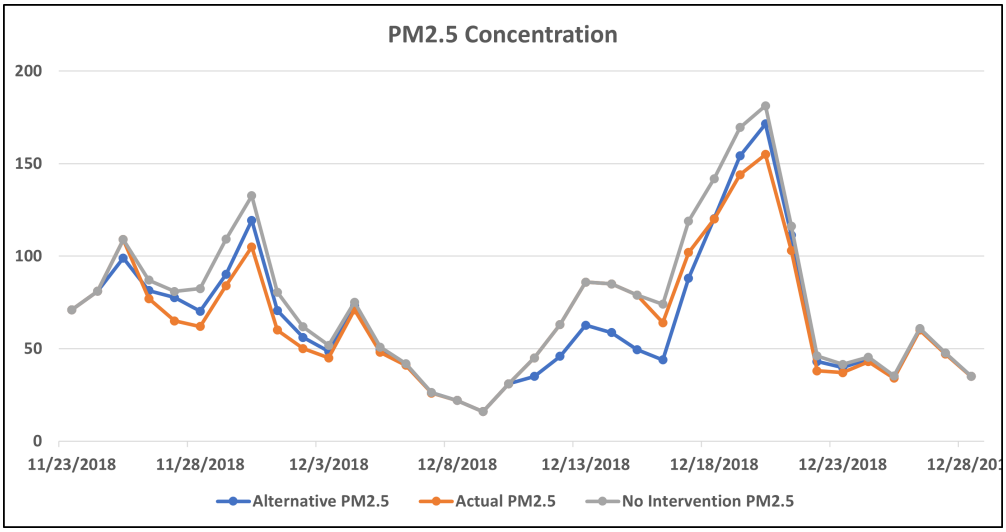

Figure 20: $PM_{2.5}$ concentration path under actual treatments, alternative treatments, and with no intervention, assuming $a = 10$

Although there are ups and downs, generally the average $PM_{2.5}$ concentration is the highest without interventions, second highest under actual treatments, and lowest under alternative treatments. Compared to no interventions, the average $PM_{2.5}$ concentration during this period under alternative and actual treatments are 13.0 and 11.6 units ($\mu g/m^3$) lower, respectively. I thus conclude that the alternative interventions are able to reduce 12.1% more $PM_{2.5}$ concentration compared to actual interventions during this period. Next, I multiply the differences in daily $PM_{2.5}$ concentration under alternative interventions and no interventions by the marginal damage day by day. Compared to no interventions, the average number of daily Bronchitis visits during this period under alternative and actual treatments are 6.8 and 5.4 lower, respectively. I thus conclude that the alternative interventions are able to reduce 25.5% more bronchitis hospital visits compared to actual interventions during this period. Lastly, I vary the assumption of a (from 1 to 50) and repeat the above calculations. The results show that the percentage reduction of $PM_{2.5}$ concentration and bronchitis hospital visits is invariant to the exact value of a. It means, regardless of the assumption

of pollution emission reduction effect of an alert event, alternative interventions on average always reduce 12.1% more $PM_{2.5}$ concentration and 25.5% more bronchitis hospital visits compared to actual interventions during this period.

2.5 Conclusions

In this chapter, I first empirically estimate the dynamic process of pollution dissipation by identifying how daily weather conditions drive the change in ambient PM levels from one day to the next. The estimation returns an excellent fit, substantiating the assumptions on the dissipation factor as well as the setup of the theoretical model. I also find that the weather variables explain a significant amount of variation in the next day's $PM_{2.5}$ concentration, which again corroborates the key role of the dissipation factor in pollution stock turnover. With the estimated state transition equation, I back-calculate the historical dissipation factors. Then I empirically estimate the accuracy of weather forecasts in Chengdu by comparing the predicted and observed meteorological conditions from the Wenjiang weather station. I correct the bias of wind speed prediction by making my own forecast. With the forecasted weather conditions, I am able to predict dissipation factors in the near future accurately, with mean absolute errors under 0.07. Lastly, I show by altering the timing of interventions, a 12.1% more $PM_{2.5}$ concentration reduction and a 25.5% more bronchitis hospital visit reduction can be achieved within a period in 2018.

2.6 Appendix

2.6.1 Daily Average Wind Direction

To generate daily average wind direction dummy variables, I use the hourly wind angle observations at the Wenjiang station from NOAA. The Shuangliu station has too many missing

observations and thus is not included. The raw data are the clockwise angles between the true north and the direction that the wind blows from. I use a standardized approach to calculate the average wind angle: First, I convert the hourly angles into coordinates. Secondly, I calculate weighted average coordinates for each day. The weights are determined by corresponding hourly wind speed observations. Then I convert average coordinates back to average angles. Next, I divide the 360 degrees of a circle into eight equal portions and generate eight corresponding dummy variables. Each portion then has a range of 45 degrees. The dummy variable of a specific portion equals one if an average angle falls into that portion, and equals zero otherwise. For example, if the average angle is 80 degrees, then the second wind direction dummy variable equals one, and all others equal zero.

2.6.2 Comparisons of Predicted and Observed Low Temperature and Atmospheric Pressure

Figure 21 illustrates the comparisons of the up to 7-day ahead predicted and observed low temperature.

In each graph of Figure 21, I plot the observed low temperature (vertical axis) against the predicted low temperature (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines overlap well in all graphs in Figure 21, suggesting the bias of prediction is minimal. Table 18 shows a quantitative assessment of the accuracy of the low temperature prediction.

Day	1	2	3	4	5	6	7
ρ	.89	.89	.89	.89	.87	.86	.85
P	.46	.52	.50	.49	.49	.47	.48

Table 18: Quantitative assessment of the accuracy of low temperature prediction

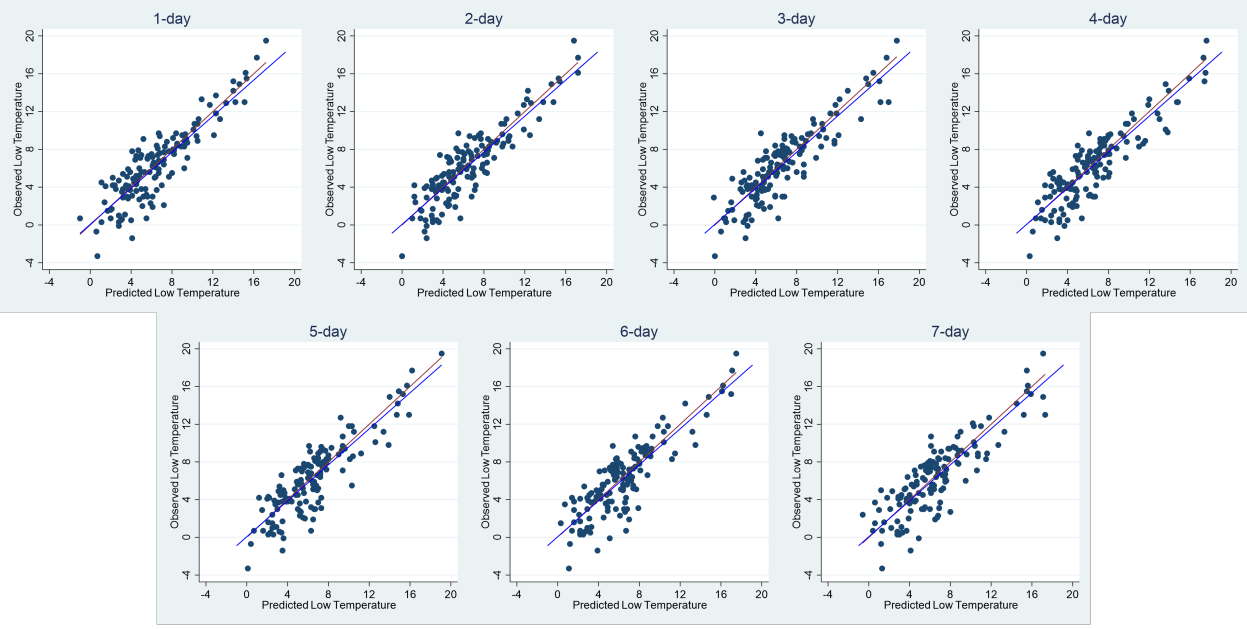

Figure 21: Comparison of the predicted and observed low temperature

ρ denotes the correlation coefficient between the predicted and observed low temperature, P denotes the probability of the predicted low temperature being higher than the observed high temperature. As can be seen in Table 18, the correlation coefficient is high and the probability of over prediction is close to one-half, again indicating accurate predictions of the high temperature.

Figure 22 illustrates the comparisons of the up to 7-day ahead predicted and observed atmospheric pressure.

In each graph of Figure 22, I plot the observed atmospheric pressure (vertical axis) against the predicted atmospheric pressure (horizontal axis). I add a fitted line of data points (blue), and a 45-degree line (red) in each graph. These two lines overlap well in all graphs in Figure 22, suggesting the bias of prediction is minimal. Table 19 shows a quantitative assessment of the accuracy of atmospheric pressure prediction.

ρ denotes the correlation coefficient between the predicted and observed atmospheric pres-

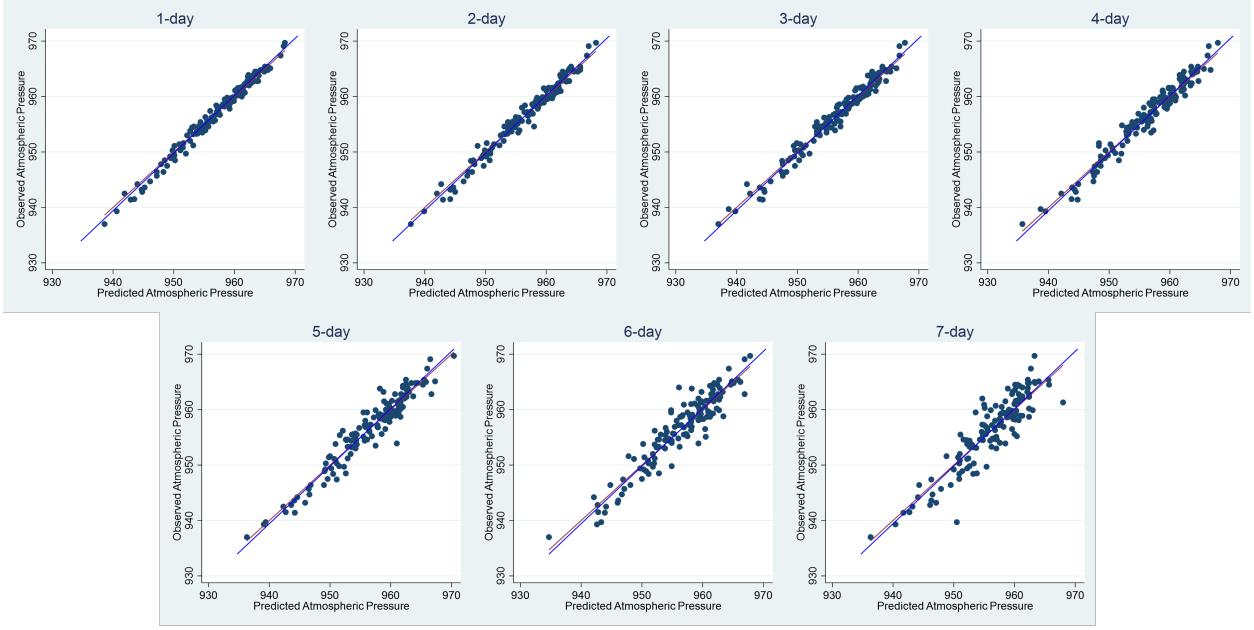

Figure 22: Comparison of the predicted and observed atmospheric pressure

Day	1	2	3	4	5	6	7
ρ	.99	.99	.99	.98	.96	.94	.91
P	.55	.48	.49	.45	.50	.48	.46

Table 19: Quantitative assessment of the accuracy of atmospheric pressure prediction

sure, P denotes the probability of the predicted atmospheric pressure being higher than the observed atmospheric pressure. As can be seen in Table 19, the correlation coefficient is very close to one and the probability of over prediction is close to one-half, again indicating highly accurate predictions of atmospheric pressure.

3 Chapter 3: Hospitalization

3.1 Introduction

PM pollution is known to induce a myriad of serious health problems. $PM_{2.5}$, the PMs with a diameter less than 2.5 micrometers, pose the greatest health hazard (Brown et al., 2013). Due to their small size, $PM_{2.5}$ are able to reach and deposit on the respiratory tract (Londahl et al., 2007), and even get into the bloodstream (Fu et al., 2011). Common health effects of PM pollution include respiratory diseases, cardiovascular diseases, and even premature death (Anderson et al., 2012). The adverse health effect is the damage of pollution in my theoretical model. In this chapter, I perform a series of analyses on the health effects of PM pollution, which provides important policy implications. I first describe the hospitalization data I obtained in Section 3.2. Then, I seek to estimate the effects of PM pollution on contemporaneous and delayed hospital visits in Chengdu, presented in Section 3.3. Next, I investigate the possible non-linear effects of PM pollution on hospitalization, reported in Section 3.4. I also assess the existence and extent of autocorrelation in the hospitalization data, shown in Section 3.5. Finally, Section 3.6 summarizes the findings of this chapter.

3.2 Data

I requested and received hospital visit data from the No.2 People's Hospital, one of the largest hospitals in Chengdu. The data consist of the counts of daily outpatient department visits by condition from March 2017 to December 2019. The dataset includes three acute respiratory conditions, namely, bronchitis, pneumonia, and asthma. Broken bone visits, which should have no correlations with PM pollution, are also included as a control. For the same period, I also obtained daily city-average $PM_{2.5}$ concentration from *aqistudy.cn*, and meteorological conditions from National Oceanic and Atmospheric Administration (NOAA)[10]. However,

[10]There are two weather stations within the city boundary, the Shuangliu station and Wenjiang station. Whenever possible, I use the average of readings at these two stations to represent the whole city. For wind

there are missing values in city-average $PM_{2.5}$ concentration and atmospheric pressure during this period. For some analysis in this chapter, no gaps in time series data are allowed. Therefore, I describe how these gaps are filled in by imputation in the appendix (Section 3.7.1). Next, I plot the daily number of visits by condition from March 2017 to December 2019, shown in Figure 23.

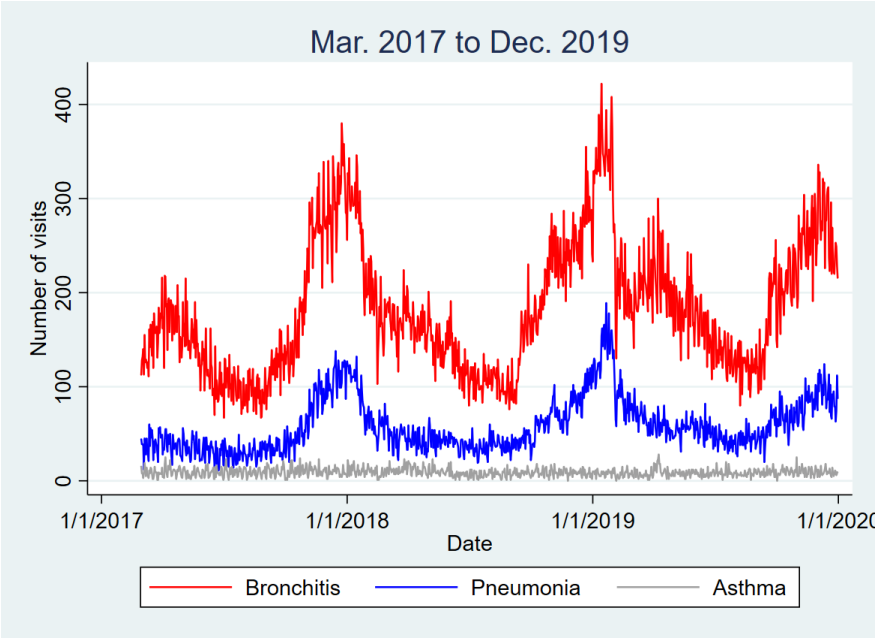

Figure 23: Daily counts of respiratory outpatient visit by condition, Mar. 2017 to Dec. 2019

In Figure 23, the red, blue, and grey time series represents bronchitis, pneumonia, and asthma, respectively. There is a clear seasonal and monthly pattern of hospital visits for bronchitis and pneumonia. Pollution seasons generally experience much more visits than other months, with November and December having the most visits. Figure 24 illustrates the daily number of visits by condition from November 2017 to February 2018 as a zoomed-in view.

direction, hourly wind speed, daily precipitation amount, and daily average atmospheric pressure, however, the Shuangliu station has either no observation or too many missing values. In these cases, I simply use the observations at the Wenjiang Station. For this reason, daily average wind speed, which includes information from the Shuangliu station as well, is not simply a linear combination of hourly wind speed.

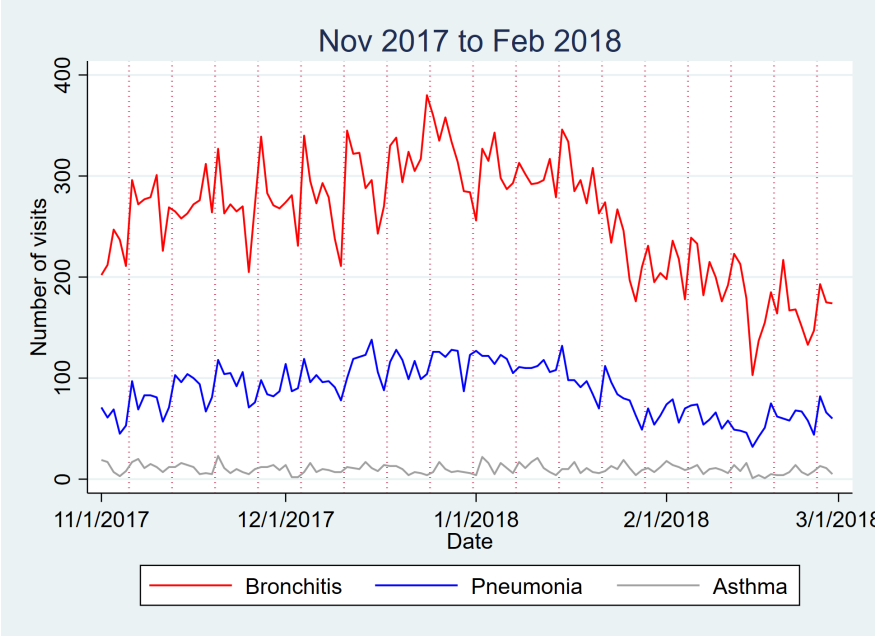

Figure 24: Daily counts of respiratory outpatient visit by condition, Nov. 2017 to Feb. 2018

In Figure 24, each vertical dashed line denotes the beginning of a week. There is a clear weekly pattern of hospital visits for bronchitis and pneumonia: the number of visits is higher during weekdays, especially the first several days, than during weekends. This pattern is mainly driven by the number of doctors serving the outpatient department. It is also noted that there is a significant drop in visit counts in mid-February 2018 during the Chinese New Year vacation when the hospital operated at a limited capacity. As an example of non-respiratory conditions, the number of daily broken bone visits from March 2017 to December 2019, and from November 2017 to February 2018 is shown in Figure 25 and Figure 26, respectively.

Each vertical dashed line in Figure 26 denotes the beginning of a week. According to Figure 25 and Figure 26, there is no clear seasonal or monthly pattern for broken bone visits. Rather, the daily number of visits is mostly driven by weekly patterns. Observations from these figures contribute greatly to the determination of the specifications in subsequent regression analysis. Table 20 below shows the summary statistics of the hospitalization dataset

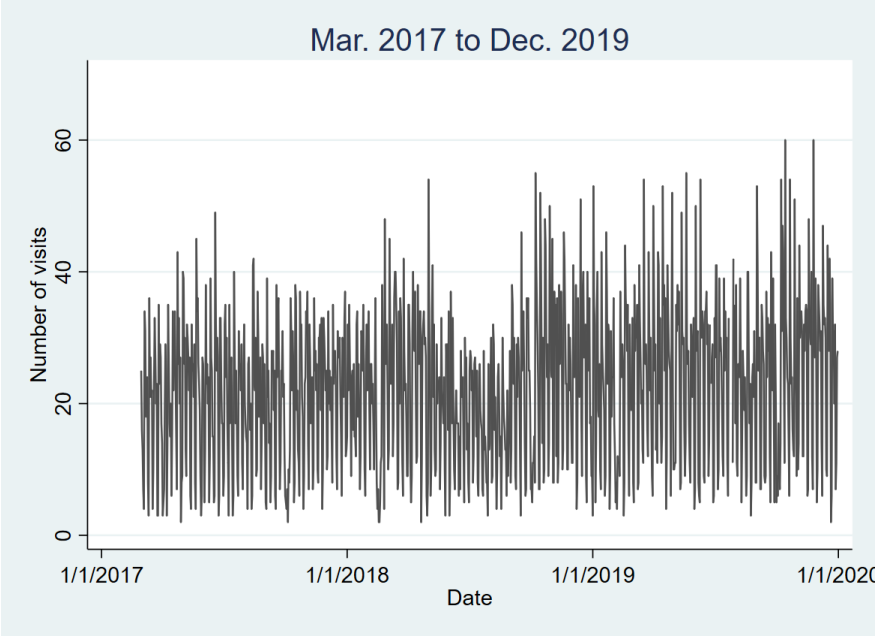

Figure 25: Daily counts of broken bone outpatient visit, Mar. 2017 to Dec. 2019

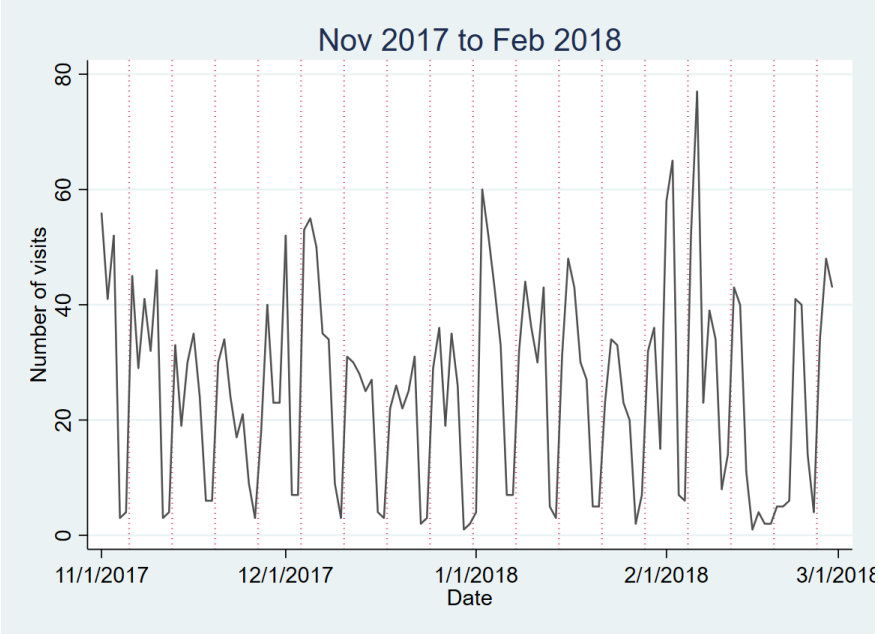

Figure 26: Daily counts of broken bone outpatient visit, Nov. 2017 to Feb. 2018

from March 2017 to December 2019. The daily average atmospheric pressure is represented by the sea level pressure.

Variable	Obs	Mean	Std. Dev.	Min	Max
Bronchitis	1,036	182.2	69.3	67	422
Pneumonia	1,036	58.1	28.7	11	189
Asthma	1,036	9.0	4.5	0	28
Broken bone	1,029	22.2	11.9	2	60
Daily city-average $PM_{2.5}$ concentration ($\mu g/m^3$)	1,036	44.8	29.9	4	201
Daily average wind speed (knots)	1,036	2.93	1.11	0.7	12.5
Daily maximum wind gust (knots)	1,036	6.37	2.30	2.2	19.5
Daily precipitation amount (inches)	1,036	.14	.43	0	5.51
Daily average temperature (Fahrenheit)	1,036	64.3	13.1	34.4	87.6
Daily high temperature (Fahrenheit)	1,036	71.3	14.3	37.1	98.8
Daily low temperature (Fahrenheit)	1,036	57.4	13.1	24.7	78.9
Daily average atmospheric pressure (mb)	1,036	1014.1	9.3	995.1	1038.7
Daily average dew point (Fahrenheit)	1,036	41.0	6.92	20.9	58.1

Table 20: Summary statistics of the hospitalization dataset

3.3 Effect of $PM_{2.5}$ Concentrations on Hospital Visits

To estimate the effect of $PM_{2.5}$ concentration on hospitalization, I have to control for confounding factors that are correlated to $PM_{2.5}$ concentration. According to the observed patterns in Section 3.2, including time fixed effects in the regression analysis is important. There is also a clear connection between respiratory diseases and weather conditions. For example, it is well known that breathing cold, dry air exacerbates asthma. Therefore, I also control for meteorological conditions when estimating the effect of $PM_{2.5}$ concentration on hospital visits. Equation (35) shows the baseline specification of the model I estimate.

$$Visit_t = \beta_0 + \beta_1 \cdot PM2.5_t + \beta_{\mathbf{f}}\mathbf{T_t} + \beta_{\mathbf{r}}\mathbf{X_t} + \eta_t \tag{35}$$

In Equation (35), $Visit_t$ denotes the number of hospital visits of one condition on day t. I use the number of bronchitis, pneumonia, asthma, and broken bone visits as the outcome variable, respectively. I perform the regression of broken bone visits as a comparison. $\mathbf{T_t}$ represents a series of time fixed effects, including day-of-week dummy variables, year-month dummy variables, workday dummy variables, and holiday dummy variables. $\mathbf{X_t}$ denotes a

vector of meteorological conditions of day t, including daily average wind speed, daily maximum wind gust, daily precipitation amount, daily average temperature, daily maximum and minimum temperature, daily average atmospheric pressure, and daily average dew point. η_t is the error term. Table 21 summarizes the results of OLS estimations of Equation (35).

	(1) Bronchitis	(2) Pneumonia	(3) Asthma	(4) Broken Bone
$PM2.5_t$	.243***	.047***	.007	-.004
	(.041)	(.017)	(.006)	(.010)
Observations	1,036	1,036	1,036	1,029
R-squared	0.870	0.862	0.371	0.711

Table 21: Regression results from OLS estimations of Equation (35)

In Table 21, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. The contemporaneous $PM_{2.5}$ concentration is statistically significant at level 1 percent in the bronchitis and pneumonia regression. The contemporaneous $PM_{2.5}$ concentration in the asthma regression, however, is not statistically significant even at level 10 percent. At 0.371, the R-squared value is also much lower than the other regressions. The signs of estimated coefficients of the contemporaneous $PM_{2.5}$ concentration in these three regressions are all positive, indicating the visits of respiratory conditions increase with the pollution level, matching our intuition. The contemporaneous $PM_{2.5}$ concentration in the broken bone regression is not statistically significant even at level 10 percent which is expected. The R-squared value is higher than that of the asthma regression, probably due to the explanatory power of time fixed effects. The sign of the estimated coefficient of the contemporaneous $PM_{2.5}$ concentration is negative. A possible explanation of the negative sign is that people perform fewer outdoor activities, and thus fewer broken bone incidents when the ambient pollution level is high. In the subsequent analysis, I use the number of bronchitis visits as the respiratory hospitalization outcome.

Next, I attempt to quantify the contemporaneous and delayed effects of $PM_{2.5}$ exposure on bronchitis hospitalization. First, I use a simple linear specification similar to Equation (35), as shown in Equation (36).

$$Br_t = \beta_0 + \sum_{\tau=0}^{7} \beta_{1,\tau} \cdot PM2.5_{t-\tau} + \beta_{\mathbf{r}}\mathbf{X_t} + \beta_{\mathbf{f}}\mathbf{T_t} + \eta_t \tag{36}$$

In Equation (36), Br_t denotes the number of bronchitis visits on day t. $\mathbf{T_t}$ and $\mathbf{X_t}$ are the same as in Equation (35). With 1,036 observations, the OLS estimation of Equation (36) returns an R-squared value of 0.876. The OLS estimation results are shown in Table 22.

Variable	OLS Estimate
$PM2.5_t$	.168***
	(.054)
$PM2.5_{t-1}$	.060
	(.062)
$PM2.5_{t-2}$	.117*
	(.061)
$PM2.5_{t-3}$	.016
	(.061)
$PM2.5_{t-4}$	.048
	(.061)
$PM2.5_{t-5}$	.088
	(.061)
$PM2.5_{t-6}$	.045
	(.061)
$PM2.5_{t-7}$	.033
	(.049))

Table 22: The effects of $PM_{2.5}$ concentrations on bronchitis hospitalization from OLS estimation of Equation (36)

In Table 22, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. According to Table 22, only the contemporaneous $PM_{2.5}$ concentration is statistically significant at level 5 percent. This is likely a result of the high correlation between $PM_{2.5}$ concentrations of adjacent days.

To better pin down the effects of each $PM_{2.5}$ concentration, I introduce a new approach where $\beta_{1,\tau}$'s in Equation (36) are modeled as cubic B-spline functions of time (Barwick et al., 2018). This approach is based on the Weierstrass Approximation Theorem that any continuous function defined on a closed interval can be uniformly approximated closely by a polynomial function. Equation (37) demonstrates the specification for $\beta_{1,\tau}$'s.

$$\beta_{1,\tau} = \gamma_0 + \gamma_1\tau + \gamma_2\tau^2 + \gamma_3\tau^3 \tag{37}$$

Then, by plugging Equation (37) into Equation (36), we have

$$\begin{aligned} Br_t = \beta_0 + \gamma_0 \cdot PM2.5_t + (\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3) \cdot PM2.5_{t-1} + ... \\ + (\gamma_0 + \gamma_1 7 + \gamma_2 7^2 + \gamma_3 7^3) \cdot PM2.5_{t-7} + \beta_{\mathbf{r}}\mathbf{X_t} + \beta_{\mathbf{f}}\mathbf{T_t} + \eta_t \end{aligned} \tag{38}$$

Rearrange Equation (38), we have

$$
\begin{aligned}
Br_t = \beta_0 &+ \gamma_0 \cdot (PM2.5_t + PM2.5_{t-1} + ... + PM2.5_{t-7}) \\
&+ \gamma_1 \cdot (PM2.5_{t-1} + 2^1 \cdot PM2.5_{t-2} + ... + 7^1 \cdot PM2.5_{t-7}) \\
&+ \gamma_2 \cdot (PM2.5_{t-1} + 2^2 \cdot PM2.5_{t-2} + ... + 7^2 \cdot PM2.5_{t-7}) \\
&+ \gamma_3 \cdot (PM2.5_{t-1} + 2^3 \cdot PM2.5_{t-2} + ... + 7^3 \cdot PM2.5_{t-7}) \\
&+ \beta_{\mathbf{r}} \mathbf{X_t} + \beta_{\mathbf{f}} \mathbf{T_t} + \eta_t
\end{aligned}
$$

Let $v1_t = PM2.5_t + PM2.5_{t-1} + ... + PM2.5_{t-7}$, $v2_t = PM2.5_{t-1} + 2^1 \cdot PM2.5_{t-2} + ... + 7^1 \cdot PM2.5_{t-7}$, $v3_t = PM2.5_{t-1} + 2^2 \cdot PM2.5_{t-2} + ... + 7^2 \cdot PM2.5_{t-7}$, and $v4_t = PM2.5_{t-1} + 2^3 \cdot PM2.5_{t-2} + ... + 7^3 \cdot PM2.5_{t-7}$. The above equation becomes

$$
Br_t = \beta_0 + \gamma_0 \cdot v1_t + \gamma_1 \cdot v2_t + \gamma_2 \cdot v3_t + \gamma_3 \cdot v4_t + \beta_{\mathbf{r}} \mathbf{X_t} + \beta_{\mathbf{f}} \mathbf{T_t} + \eta_t \tag{39}
$$

Equation (39) represents the contemporaneous and lagging $PM_{2.5}$ concentrations with only four variables as opposed to eight variables in Equation (36), thus alleviating some of the concern for multicollinearity.

From Equation (38), we can derive the marginal effects of $PM_{2.5}$ concentration variables in this new specification by taking partial derivative with respect to each $PM_{2.5}$ concentration variable. The marginal effects are listed in Table 23.

I then perform an OLS estimation of Equation (39). With 1,036 observations, the R-squared value is 0.876. Table 24 below lists the OLS estimates of the marginal effect of each $PM_{2.5}$ concentration variable.

Variable	Marginal Effect
$PM2.5_t$	γ_0
$PM2.5_{t-1}$	$\gamma_0 + 1^1 \cdot \gamma_1 + 1^2 \cdot \gamma_2 + 1^3 \cdot \gamma_3$
$PM2.5_{t-2}$	$\gamma_0 + 2^1 \cdot \gamma_1 + 2^2 \cdot \gamma_2 + 2^3 \cdot \gamma_3$
$PM2.5_{t-3}$	$\gamma_0 + 3^1 \cdot \gamma_1 + 3^2 \cdot \gamma_2 + 3^3 \cdot \gamma_3$
$PM2.5_{t-4}$	$\gamma_0 + 4^1 \cdot \gamma_1 + 4^2 \cdot \gamma_2 + 4^3 \cdot \gamma_3$
$PM2.5_{t-5}$	$\gamma_0 + 5^1 \cdot \gamma_1 + 5^2 \cdot \gamma_2 + 5^3 \cdot \gamma_3$
$PM2.5_{t-6}$	$\gamma_0 + 6^1 \cdot \gamma_1 + 6^2 \cdot \gamma_2 + 6^3 \cdot \gamma_3$
$PM2.5_{t-7}$	$\gamma_0 + 7^1 \cdot \gamma_1 + 7^2 \cdot \gamma_2 + 7^3 \cdot \gamma_3$

Table 23: The marginal effects of $PM_{2.5}$ concentration variables on bronchitis hospitalization in Equation (39)

Variable	OLS Estimate
$PM2.5_t$	.158***
	(.039)
$PM2.5_{t-1}$	.094***
	(.017)
$PM2.5_{t-2}$	.062***
	(.021)
$PM2.5_{t-3}$	.054***
	(.016)
$PM2.5_{t-4}$	.057***
	(.015)
$PM2.5_{t-5}$	.061***
	(.021)
$PM2.5_{t-6}$	.057***
	(.017)
$PM2.5_{t-7}$	.032
	(.036))

Table 24: The effects of $PM_{2.5}$ concentration variables on bronchitis hospitalization from OLS estimation of Equation (39)

In Table 24, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses, which are calculated by the square root of the variance of the linear combinations of coefficients as shown in Table 23. According to the results shown in Table 24, $PM_{2.5}$ exposure up to 6 days ago has an effect on contemporaneous bronchitis hospitalization that is statistically significant at level 1 percent. The statistical

significance of lagging $PM_{2.5}$ concentrations is probably a result of the delayed health effects of $PM_{2.5}$ exposure. These results are consistent with our intuition as well as literature that find delayed health hazard of PM exposure (Pope et al., 1992; Schwartz, 2000; Kim et al., 2003; Min et al., 2008). These estimates also serve as the basis of the simulation analysis in Chapter 2. In the appendix (Section 3.7.2), I test the effect of future $PM_{2.5}$ concentration on contemporaneous hospital visits for this new approach.

3.4 Non-linear Effect of $PM_{2.5}$

In this section, I search for evidence that the effect of $PM_{2.5}$ on hospitalization is non-linear. First, I perform a kernel-weighted local-mean smoothing of hospital visits on contemporaneous $PM_{2.5}$ concentration. This method is selected because it allows an assessment of potential non-linear effects in a non-parametric way. It generates a graph of smoothed values as a visual depiction of the relationship between hospital visits and contemporaneous $PM_{2.5}$ concentration. I use the Epanechnikov kernel and set the kernel bandwidth to 20, as shown in Figure 27.

As can be seen in Figure 27, it is clear that the smoothed values in (A) and (B) are upward-sloping, while (C) displays an almost flat line, illustrating the responsiveness of respiratory conditions and the unresponsiveness of broken bone to $PM_{2.5}$ concentration, respectively. However, in (A) and (B), we do not observe any obvious curvature of the smoothed values which are almost linear. Next, I attempt to quantitatively determine if any non-linear effect of $PM_{2.5}$ on hospitalization exists in a few ways. First, I use a quadratic specification to test for the non-linear effect, shown in Equation (40)[11].

[11] In the specifications of this section, I only include the non-linear terms of the contemporaneous $PM_{2.5}$ concentration by assuming the lagged effects are linear. Admittedly, the non-linear effects of lagging $PM_{2.5}$ concentrations might also exist. However, incorporating both contemporaneous and lagged non-linear effects will lead to a convoluted specification. Because the $PM_{2.5}$ concentration is highly correlated over time, it is difficult to individually identify the contemporaneous and lagged non-linear effects by estimating such a

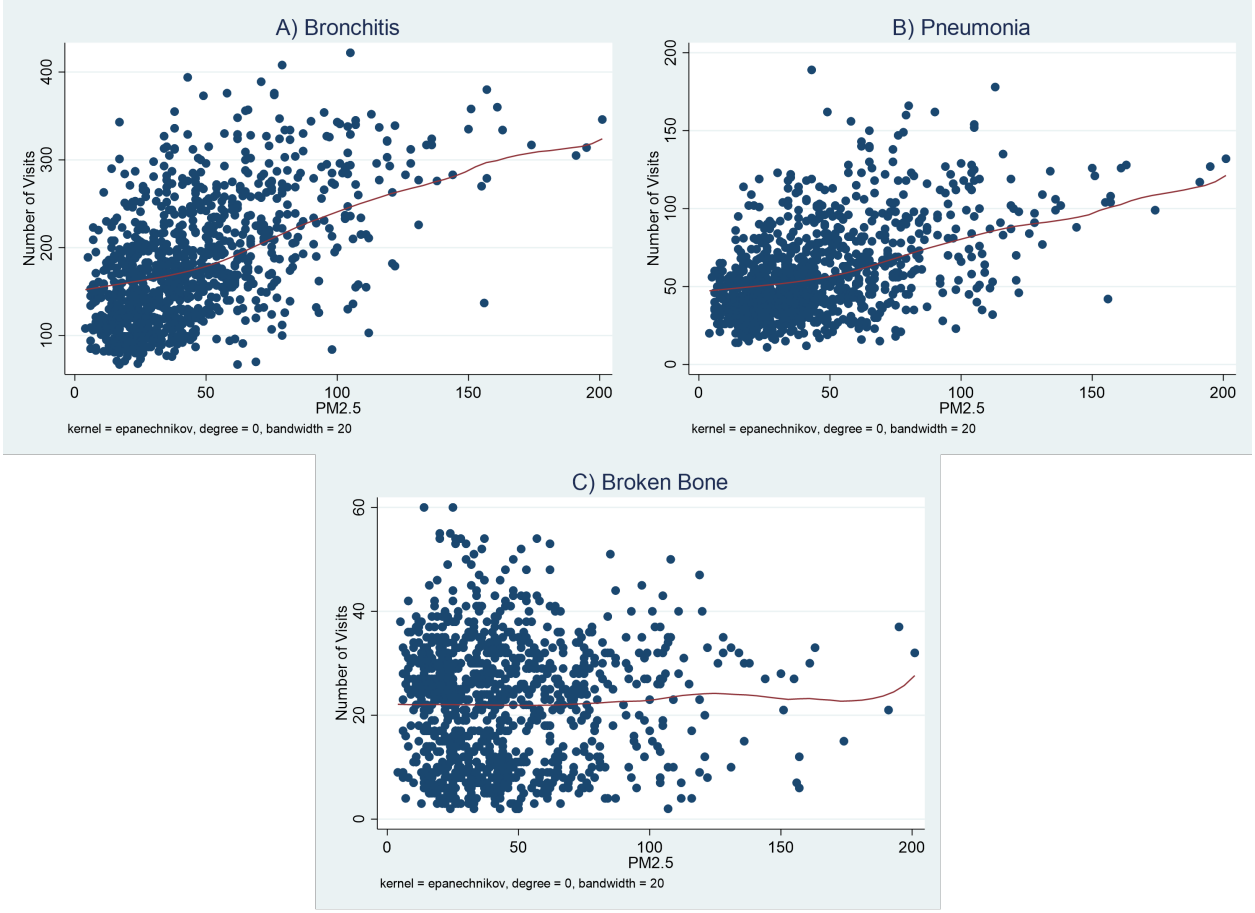

Figure 27: Kernel-weighted local-mean smoothing of bronchitis (A), pneumonia (B), and broken bone (C) visits on contemporaneous $PM_{2.5}$ concentration

$$Visit_t = \beta_0 + \beta_1 \cdot PM2.5_t + \beta_2 \cdot PM2.5_t^2 + \sum_{\tau=1}^{7} \beta_{3,\tau} \cdot PM2.5_{t-\tau} + \beta_{\mathbf{r}}\mathbf{X_t} + \beta_{\mathbf{f}}\mathbf{T_t} + \eta_t \quad (40)$$

In Equation (40), $Visit_t$ is the number of hospital visits for one disease on day t. $\mathbf{T_t}$ and $\mathbf{X_t}$ are the same as in Equation (35). I use the number of bronchitis, pneumonia, and broken bone visits as the dependent variable, respectively. I perform the regression of broken bone visits as a comparison. I control for the seven lagging $PM_{2.5}$ concentrations. Table 25 sum-

specification. It is unclear how this assumption will bias the estimates if both contemporaneous and lagged non-linear effects exist. Since I am interested in prediction rather than the exact point estimates in this analysis, the presence of any potential bias is of less concern.

marizes the results of an OLS estimation of Equation (40).

	(1) Bronchitis	(2) Pneumonia	(3) Broken Bone
$PM_{2.5}$	.1690*	.0153	-.0407
	(.1006)	(.0435)	(.0263)
$PM_{2.5}^2$	-7.95e-06	.0002	.0003
	(.0006)	(.0003)	(.0002)
Observations	1,036	1,036	1,036
R-squared	0.876	0.865	0.713

Table 25: Regression results from OLS estimation of Equation (40)

In Table 25, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. For bronchitis visits, the estimated coefficient of the quadratic term is negative while that of the linear term is positive, corresponding to an inverted U-shaped parabola with the vertex at 10,628.9. However, the estimated coefficient of the quadratic term is very small and not statistically significant even at level 10 percent, suggesting the curvature is at best trivial. For pneumonia visits, the estimated coefficients of the quadratic and linear terms are both positive, corresponding to a U-shaped parabola with the vertex at -32.5. The estimated coefficient of the quadratic term is also very small and not statistically significant even at level 10 percent. I fail to find any evidence of non-linearity for bronchitis and pneumonia visits in this exercise. For broken bone visits, as a comparison, the estimated coefficient of the quadratic term is positive while that of the linear term is negative, corresponding to a U-shaped parabola with the vertex at 68.1. The estimated coefficient of the quadratic term is also very small and not statistically significant even at level 10 percent. As a control, the results of broken bone visits match our expectation.

Next, I investigate the existence of non-linearity following the approach of Reyes (2007). I create four variables containing a linear spline of $PM_{2.5}$ concentration. These variables are constructed such that their coefficients represent the slopes for the corresponding interval.

The knots are at the borders of Chinese air quality classification, namely, 35, 75, 115. Equation (41) shows the specification.

$$Visit_t = \beta_0 + \sum_{n=1}^{4} \beta_{1,n} \cdot s\_n_t + \sum_{\tau=1}^{7} \beta_{3,\tau} \cdot PM2.5_{t-\tau} + \beta_{\mathbf{r}} \mathbf{X_t} + \beta_{\mathbf{f}} \mathbf{T_t} + \eta_t \tag{41}$$

In Equation (41), $Visit_t$ is the number of hospital visits of one condition on day t. s_1_t, s_2_t, s_3_t, and s_4_t denote the variables containing a linear spline of $PM_{2.5}$ concentration. $\beta_{1,n}$ are the slopes of the different segments of the linear spline. $\mathbf{T_t}$ and $\mathbf{X_t}$ are the same as in Equation (35). I use the number of bronchitis, pneumonia, and broken bone visits as the dependent variable, respectively. I perform the regression of broken bone visits as a comparison. I control for the seven lagging $PM_{2.5}$ concentrations. Table 26 summarizes the results.

	(1) Bronchitis	(2) Pneumonia	(3) Broken Bone
s_1_t	.27**	.11*	-.01
	(.13)	(.06)	(.04)
s_2_t	.17*	.02	-.03
	(.10)	(.04)	(.02)
s_3_t	-.04	-.04	.06
	(.14)	(.06)	(.04)
s_4_t	.41**	.22***	.02
	(.16)	(.07)	(.04)
Observations	1,036	1,036	1,029
R-squared	0.877	0.867	0.714

Table 26: Regression results from OLS estimation of Equation (41). Knots for linear splines are at borders of Chinese air quality classification.

In Table 26, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. For bronchitis visits, a marginal increase in $PM_{2.5}$ concentration is statistically significant at level 5 percent in the first and fourth intervals. The point estimation of the coefficient on s_4_t is also the largest. This seems to suggest

that there is a non-linear response of bronchitis visits once the air quality becomes moderately polluted. The results of pneumonia visits are similar. A marginal increase in $PM_{2.5}$ concentration is statistically significant at level 5 percent only in the fourth interval. The point estimation of the coefficient on s_4_t is also the largest. This seems to suggest that there is a non-linear response of pneumonia visits once the air quality becomes moderately polluted. For broken bone visits, as a comparison, none of the variables containing a linear spline is statistically significant even at level 10 percent, consistent with the expectation. I then repeat this exercise but let the knots be at the quartiles of $PM_{2.5}$ concentration instead of the borders of the Chinese air quality classification, namely, 23, 38, and 59. With the new set of variables containing a linear spline, I perform an OLS estimation of Equation (41) again. Table 27 summarizes the results.

	(1) Bronchitis	(2) Pneumonia	(3) Broken Bone
s_1_t	.08	.13	.08
	(.29)	(.13)	(.08)
s_2_t	.39*	.11	-.07
	(.21)	(.09)	(.06)
s_3_t	.08	-.04	-.03
	(.16)	(.07)	(.04)
s_4_t	.15**	.06*	.03
	(.07)	(.03)	(.02)
Observations	1,036	1,036	1,029
R-squared	0.877	0.866	0.714

Table 27: Regression results from OLS estimation of Equation (41). Knots for linear splines are at the quartiles of $PM_{2.5}$ concentration.

In Table 27, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. For bronchitis visits, a marginal increase in $PM_{2.5}$ concentration is statistically significant at level 5 percent only in the fourth interval. However, the point estimation of the coefficient on s_4_t is not the largest. For pneumonia visits a marginal increase in $PM_{2.5}$ concentration is not statistically significant at level 5 percent in

any interval. The point estimation of the coefficient on s_4_t is also not the largest. Therefore, I fail to find any evidence of non-linearity for bronchitis and pneumonia visits. For broken bone visits, as a comparison, none of the variables containing a linear spline is statistically significant even at level 10 percent, consistent with the expectation. Therefore, the evidence of non-linear effects from this approach is also unconvincing as it is sensitive to the choice of knots.

3.5 Autocorrelation

In the regression analyses I perform so far in this chapter, I have assumed that the error terms are independent and identically distributed. However, due to the time series nature of the hospitalization data, the error terms may be autocorrelated. If this is true, there might be a serious downward bias of the estimated standard errors, which can invalidate statistical inferences. For the test of autocorrelation, I focus on the regression of Equation (39) as it provides the key outputs of this chapter. I first plot the residuals of the OLS estimation of Equation (39) against its first lagging residuals, illustrated in Figure 28.

In Figure 28, the data points appear random without any particular patterns, suggesting autocorrelation is minor. To investigate how the autocorrelation affects the estimated effects in Table 24, I rerun the regression of Equation (39) with Newey-West standard errors. I set the maximum lag order of correlation to 7. Table 28 summarizes the results.

In Table 28, */**/*** indicates statistical significance at the 10/5/1 percent levels. Standard errors are shown in parentheses. Compared to the results in Table 24, the standard errors are marginally larger indicating very little positive autocorrelation. Assuming the error terms to be autocorrelated does not change the statistical significance of $PM_{2.5}$ concentration variables at level 5 percent. Therefore, the conclusions from this regression remain unchanged

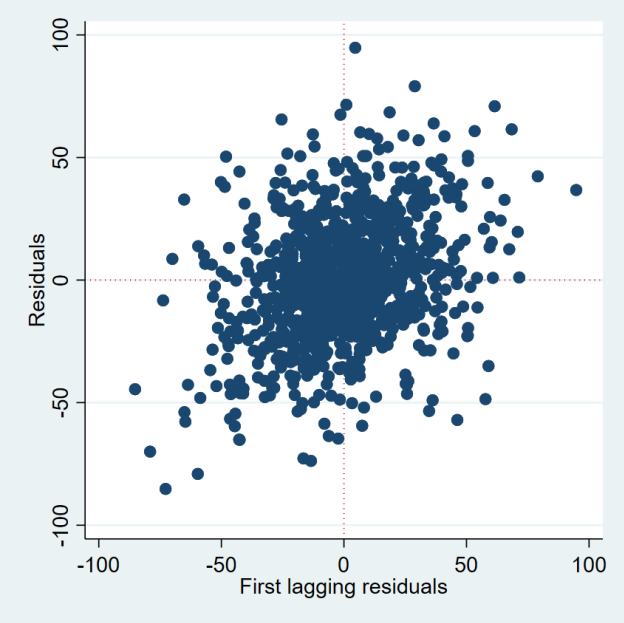

Figure 28: The residuals and the first lagging residuals of the OLS estimation of Equation (39)

Variable	OLS Estimate
$PM2.5_t$	.158***
	(.044)
$PM2.5_{t-1}$	.094***
	(.025)
$PM2.5_{t-2}$	.062**
	(.027)
$PM2.5_{t-3}$	.054***
	(.020)
$PM2.5_{t-4}$	.057***
	(.020)
$PM2.5_{t-5}$	.061**
	(.025)
$PM2.5_{t-6}$	.057***
	(.021)
$PM2.5_{t-7}$	.032
	(.046))

Table 28: The effects of $PM_{2.5}$ concentrations on bronchitis hospitalization from OLS estimation of Equation (39), with Newey-West standard errors

even allowing the error terms to be serially correlated.

3.6 Conclusions

In this chapter, I first estimate the health effects of PM pollution. I find that exposures to $PM_{2.5}$ up to six days ago are associated with contemporaneous bronchitis hospitalization. These results are in accordance with the literature that concludes delayed health effects of PM exposure. The estimates from this analysis provide very important policy implications by enabling the simulation in Chapter 2 which demonstrates the gains that could have been achieved by altering the timing of historical interventions according to both existing conditions and expected upcoming weather patterns. Then, I assess the curvature of the response function of respiratory hospitalization to PM pollution. I fail to find any convincing evidence that supports the presence of nonlinearity. One caveat of this conclusion is that I only consider possible non-linear effects of the contemporaneous pollution concentration but ignore the possible non-linear effects of the lagged pollution concentration, due to the complexity of the specification and the difficulty in identification if both of them are incorporated. Although this simplification might introduce a bias in an unclear way, it is of less concern since prediction instead of specific point estimates is the focus of this analysis. The lack of non-linear responses can be a result of people exercising avoidance behaviors when the ambient pollution level is high. However, I do not have access to data that can help to verify this hypothesis. Finally, I investigate the existence and extent of autocorrelation in bronchitis hospitalization data. The analysis shows evidence of minor positive autocorrelation, although it does not alter the conclusions of the effects of $PM_{2.5}$ exposures on bronchitis hospitalization.

3.7 Appendix

3.7.1 Imputation of Missing Values

In this section, I show how missing values in daily city-average $PM_{2.5}$ concentrations and atmospheric pressure are imputed. From March 2017 to December 2019, there are four short periods without daily city-average $PM_{2.5}$ concentration observations: December 30, 2017 to January 1, 2018; February 10, 2018 to February 13, 2018; May 29, 2018 to May 30, 2018; and December 5, 2018 to December 6, 2018. To fill in these gaps, I calculate the mean of daily station-average $PM_{2.5}$ concentrations from 24 monitoring stations in Chengdu. These monitoring stations are scattered all over the city area of Chengdu. Therefore, they are representative of the whole city. Figure 29 shows a comparison of the calculated mean of daily station-average $PM_{2.5}$ concentrations and daily city-average $PM_{2.5}$ concentration.

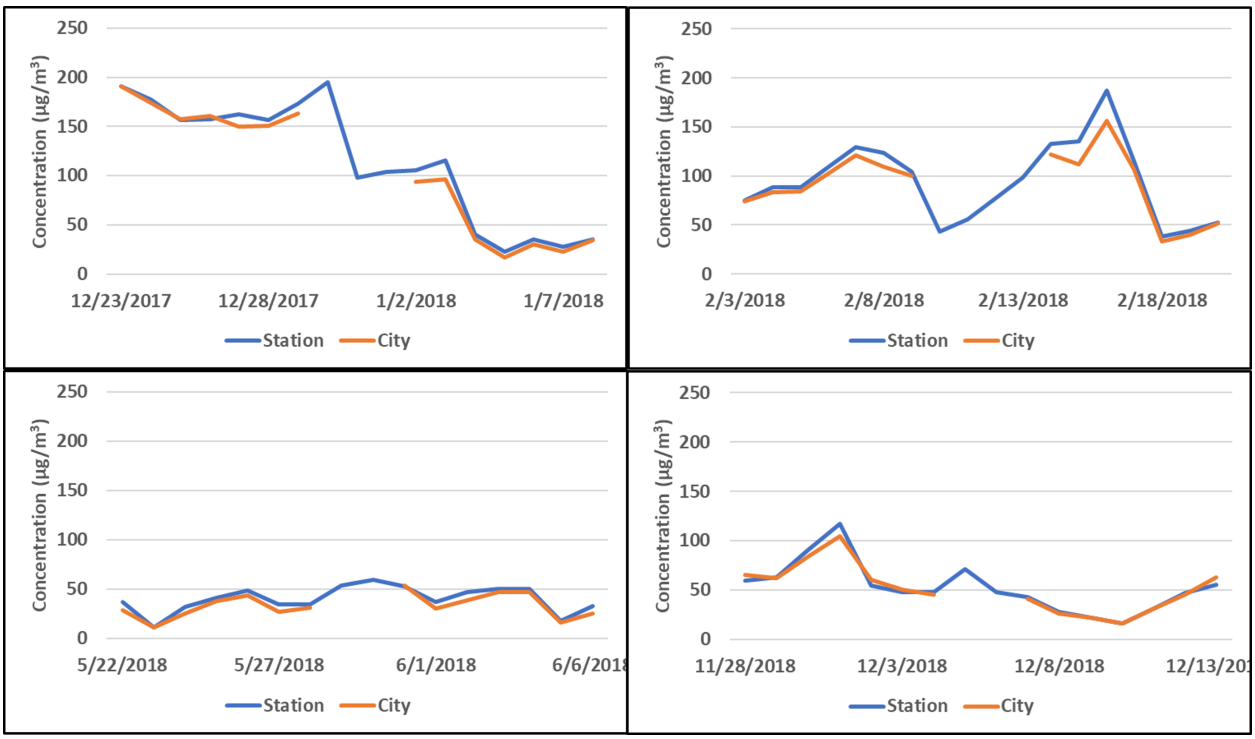

Figure 29: Comparison of the calculated mean of daily station-average $PM_{2.5}$ concentrations and daily city-average $PM_{2.5}$ concentration

In Figure 29, the blue and red time series represent the calculated mean of daily station-

average $PM_{2.5}$ concentrations and daily city-average $PM_{2.5}$ concentration, respectively. In all four graphs, I include seven days before and after the periods when daily city-average $PM_{2.5}$ concentrations are missing. It is clear that both the trend and level of the two time series before and after these periods are very similar. Therefore, I use the calculated mean of daily station-average $PM_{2.5}$ concentrations when daily station-average $PM_{2.5}$ concentration is missing.

From November 17, 2018 to November 22, 2018, the daily average atmospheric pressure is missing. To fill in this gap, I use the daily atmospheric pressure of a nearby city, Ya'an. Ya'an is located in the same basin area and is only less than 100 miles away from Chengdu. The two cities also share similar elevations (1,900 *vs* 1,240 ft). Figure 30 shows a comparison of the daily average atmospheric pressure in Ya'an and Chengdu.

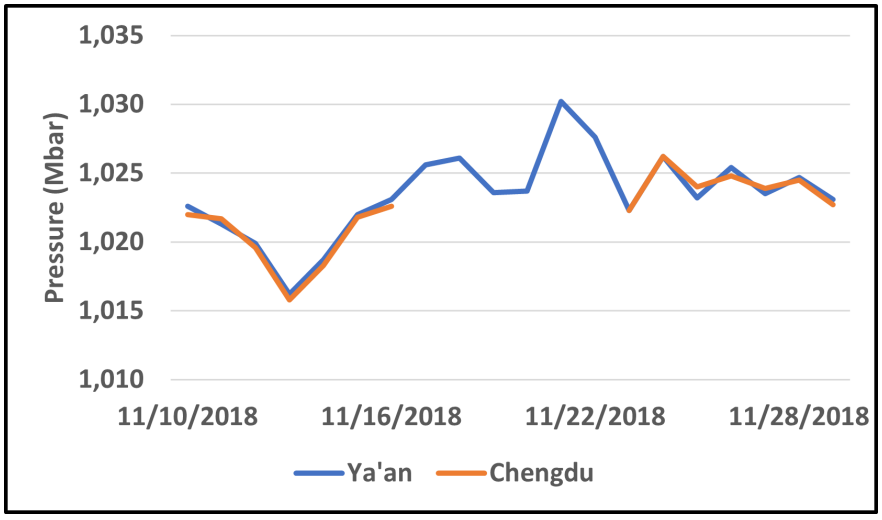

Figure 30: Comparison of the daily average atmospheric pressure in Ya'an and Chengdu

In Figure 30, the blue and red time series represent the daily average atmospheric pressure in Ya'an and Chengdu, respectively. I also include seven days before and after the periods when the daily average atmospheric pressure in Chengdu is missing. It is clear that both the trend and level of the two time series before and after these periods are very similar.

Therefore, I use the daily average atmospheric pressure in Ya'an when the daily average atmospheric pressure in Chengdu is missing.

3.7.2 A Test of the Effect of Future $PM_{2.5}$ Concentration on Contemporaneous Hospitalization

In this section, I test the effect of future $PM_{2.5}$ concentration on contemporaneous hospital visits when $\beta_{1,\tau}$'s are modeled as cubic B-spline functions of time. Future $PM_{2.5}$ exposures should not affect contemporaneous hospitalization. Therefore, I expect the estimated effect to be indistinguishable from zero. First, I include the first leading $PM_{2.5}$ concentration into Equation (36), as shown in Equation (42).

$$Br_t = \beta_0 + \sum_{\tau=-1}^{7} \beta_{1,\tau} \cdot PM2.5_{t-\tau} + \beta_{\mathbf{r}}\mathbf{X_t} + \beta_{\mathbf{f}}\mathbf{T_t} + \eta_t \tag{42}$$

Then, by plugging Equation (37) into Equation (42), we have

$$\begin{aligned} Br_t = \beta_0 &+ (\gamma_0 - \gamma_1 + \gamma_2 - \gamma_3) \cdot PM2.5_{t+1} \\ &+ \gamma_0 \cdot PM2.5_t + (\gamma_0 + \gamma_1 + \gamma_2 + \gamma_3) \cdot PM2.5_{t-1} + ... \\ &+ (\gamma_0 + \gamma_1 7 + \gamma_2 7^2 + \gamma_3 7^3) \cdot PM2.5_{t-7} + \beta_{\mathbf{r}}\mathbf{X_t} + \beta_{\mathbf{f}}\mathbf{T_t} + \eta_t \end{aligned} \tag{43}$$

Rearrange Equation (43), we have

$$
\begin{aligned}
Br_t = \beta_0 &+ \gamma_0 \cdot (PM2.5_{t+1} + PM2.5_t + PM2.5_{t-1} + ... + PM2.5_{t-7}) \\
&+ \gamma_1 \cdot (-PM2.5_{t+1} + PM2.5_{t-1} + 2^1 \cdot PM2.5_{t-2} + ... + 7^1 \cdot PM2.5_{t-7}) \\
&+ \gamma_2 \cdot (PM2.5_{t+1} + PM2.5_{t-1} + 2^2 \cdot PM2.5_{t-2} + ... + 7^2 \cdot PM2.5_{t-7}) \\
&+ \gamma_3 \cdot (-PM2.5_{t+1} + PM2.5_{t-1} + 2^3 \cdot PM2.5_{t-2} + ... + 7^3 \cdot PM2.5_{t-7}) \\
&+ \beta_{\mathbf{r}} \mathbf{X_t} + \beta_{\mathbf{f}} \mathbf{T_t} + \eta_t
\end{aligned}
$$

Let $w1_t = PM2.5_{t+1}+PM2.5_t+PM2.5_{t-1}+...+PM2.5_{t-7}$, $w2_t = -PM2.5_{t+1}+PM2.5_{t-1}+2^1 \cdot PM2.5_{t-2} + ... + 7^1 \cdot PM2.5_{t-7}$, $w3_t = PM2.5_{t+1} + PM2.5_{t-1} + 2^2 \cdot PM2.5_{t-2} + ... + 7^2 \cdot PM2.5_{t-7}$, and $w4_t = -PM2.5_{t+1} + PM2.5_{t-1} + 2^3 \cdot PM2.5_{t-2} + ... + 7^3 \cdot PM2.5_{t-7}$. The above equation becomes

$$
Br_t = \beta_0 + \gamma_0 \cdot w1_t + \gamma_1 \cdot w2_t + \gamma_2 \cdot w3_t + \gamma_3 \cdot w4_t + \beta_{\mathbf{r}} \mathbf{X_t} + \beta_{\mathbf{f}} \mathbf{T_t} + \eta_t \tag{44}
$$

In this specification, the marginal effect of $PM2.5_{t+1}$ is represented by $\gamma_0 - \gamma_1 + \gamma_2 - \gamma_3$ according to Equation (43). With 1,036 observations, the OLS estimation of Equation (44) returns an R-squared value of 0.876. OLS estimate of the marginal effect of $PM2.5_{t+1}$ is .077 with a standard error of .039, not statistically significant at level 5 percent. This result is consistent with the fact that future $PM_{2.5}$ concentration does not affect contemporaneous hospitalization.

4 Reference

1. Anderson, J.O., et al. 2012. Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *J. Med. Toxicol.* 8: 166–175.

2. Barwick, P.J., et al. 2018. The Healthcare Cost of Air Pollution: Evidence from the World's Largest Payment Network. *NBER Working Paper 24688.*

3. Brown, J.S., et al. 2013. Thoracic and respirable particle definitions for human health risk assessment. *Part. Fibre Toxicol.* 10:12.

4. Cohen, Aaron J., et al. 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *The Lancet* 2017 (10082), 389.

5. Davis, Lucas W. 2008. The effect of driving restrictions on air quality in Mexico City. *Journal of Political Economy* 116 (1): 38-81.

6. Fu, M., et al. 2011. Advances of study on monitoring and evaluation of PM2.5 pollution. *Meteorol. Disaster Reduc. Res.* 34:1–6.

7. Greenstone, Michael, and Rema Hanna. 2014. Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review* 104, 3038-3072.

8. Greenstone, Michael, et al. 2017. Clearing the air on Delhi's odd-even program. Working Paper. India Policy Forum Vol. 14 (2017-18).

9. Kim, H., et al. 2003. The lag-effect pattern in the relationship of particulate air pollution to daily mortality in Seoul, Korea. *Int. J. Biometeorol.* 48, 25–30.

10. Li, Lulu, et al. 2017. Characteristics and source apportionment of PM2.5 during persistent extreme haze events in Chengdu, southwest China. *Environmental Pollution* 230: 718-729.

11. Liao, Tingting, et al. 2017. Heavy pollution episodes, transport pathways and potential sources of PM_2.5 during the winter of 2013 in Chengdu (China). *Science of the Total Environment* 584-585, 1056-1065.

12. Londahl, J., et al. 2007. Size-resolved respiratory-tract deposition of fine and ultrafine hydrophobic and hygroscopic aerosol particles during rest and exercise. *Inhal. Toxicol.* 19(2): 109–16.

13. Lv, Baolei, et al. 2015. Characterizations of PM2.5 Pollution Pathways and Sources Analysis in Four Large Cities in China. *Aerosol and Air Quality Research* 15: 1836–1843.

14. Matus, K., et al. 2012. Health damages from air pollution in China. *Global Environmental Change* 22, 55-66.

15. Min, Jin-Young, et al. 2008. Lag Effect of Particulate Air Pollution on Lung Function in Children. *Pediatric Pulmonology* 43: 476–480.

16. Pope, C. Arden, et al. 1992. Daily mortality and PM10 pollution in Utah Valley. *Arch. Environ. Health* 42 (3): 211–217.

17. Reyes, Jessica W. 2007. Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime. *The B.E. Journal of Economic Analysis & Policy* 7 (1): 1-43.

18. Salas, Christian. 2010. Evaluating public policies with high frequency data: Evidence for driving restrictions in Mexico City revisited. Working Paper no. 374, Institute of Economics, Pontifical Catholic University of Chile.

19. Schwartz, Joel. 2000. The Distributed Lag between Air Pollution and Daily Deaths. *Epidemiology*: 11 (3): 320-326.

20. Simeonova, Emilia, et al. 2019. Congestion Pricing, Air Pollution, and Children's Health *J. Human Resources* 0218-9363R2.

21. Sun, Cong, et al. 2014. Restricting driving for better traffic and clearer skies: Did it work in Beijing? *Transport Policy* 32: 34-41.

22. Viard, V. Brian, and Shihe Fu. 2015. The effect of Beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics* 125: 98-115.

23. Zhong, Nan, et al. 2017. Traffic Congestion, Ambient Air Pollution, and Health: Evidence from Driving Restrictions in Beijing. *Journal of Association of Environmental and Resource Economists* 4:3, 821-856.