Essays on the effectiveness of air pollution control policies in China

By

Bowen Liu

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Sciences
University of Birmingham
January 2022

STATEMNET OF PRIOR PUBLICATION

A version of Chapter three is publicy available as:

Cole, M. A., Elliott, R. J., Liu, B. (2020). The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. Environmental and Resource Economics, 76(4), 553-580.

DEDICATION

This thesis is dedicated to my parents

Qinghan Liu and Fenxia Shi

for their love and support

ACKNOWLEDGMENTS

I would like to express the most sincere gratitude and appreciation to my supervisors Professor Matthew Cole and Professor Robert Elliott, they have been providing consistent and insightful support to my PhD study, they've always been inspiring me to explore new research ideas, and encouraging me to learn new research methods, it would be extremely difficult for me to finish my PhD thesis without their precious supports. I would also like to thank Professor Eric Strobl, Professor Anindya Banerjee, Dr Liyun Zhang who provided valuable comments on my work at the departmental seminars.

Additionally, I would like to express my sincere appreciation to Professor Zongbo Shi, Tuan Van Vu and Congbo Song from environmental science who kindly guided me to explore the research methods from Atmospheric science, having the opportunity to talk to people from different research backgrounds really helped me a lot to shape the interdisciplinary research methods.

I would like to thank all the senior colleagues and PhD classmates in the Muirhead tower ofcice, including Dr Tong Yin, Dr Yanan Zhang, Dr Danying Li, Dr Yaopeng Wang, Dr Shixuan Wang, Dr Tong Zhu, Dr Saul Basurto Hernandez, Dr Viet Nguyen-Tien, Huanjia Ma, Kai Cheng, Junchen Li, Yu Pei, Jingmin Zhu, Gianluca Bortoletto, Hamideh Mohtashami Borzadaran, Jorge Soguero Escuer, who helped me a lot both in study and life throughout the four-year PhD period.

Contents

Page

1 Introduction **1**

2 Air Pollution and Environmental Policy in China: Testing the Effectiveness of the Environmental Inspection Policy **5**

2.1 Introduction . . . 5
2.2 The Central Environmental Inspection Policy . . . 8
2.3 Literature Review . . . 14
2.3.1 Long term policy evaluation in developing and developed regions . . . 14
2.3.2 Short term policy evaluation in developing and developed regions . . . 18
2.3.3 Environmental protection inspection and air quality . . . 19
2.3.4 Environmental policy and air pollution in China . . . 21
2.4 Data . . . 24
2.4.1 Air quality data . . . 24
2.4.2 Meteorological Data . . . 26
2.5 Methodology . . . 27
2.6 Results . . . 29
2.6.1 Descriptive statistics . . . 29
2.6.2 Estimation results . . . 32
2.6.3 Robustness checks . . . 37
2.7 Conclusions . . . 40

3 The Impact of the Wuhan Covid-19 Lockdown on Air Pollution and Health: A Machine Learning and Augmented Synthetic Control Approach 65

3.1 Introduction 65

3.2 Data and Methodology 70

3.2.1 Machine learning 72

3.2.2 The augmented synthetic control method 76

3.3 Results 81

3.3.1 Machine learning results 81

3.3.2 The Impact of the Wuhan lockdown on local air pollution using ridge ASCM 82

3.3.3 Placebo tests 85

3.4 The Health Implications of China's Falling Pollution 87

3.5 Discussion and Conclusions 90

4 Can city level environmental inspections reduce local air pollution? A machine learning and ASCM approach 106

4.1 Introduction 106

4.2 Background: the Central Environmental Inspection Policy (CEIP) 110

4.3 Data 113

4.4 Empirical strategy 116

4.4.1 DID and the pre-parallel trends 118

4.4.2 Weather conditions and air quality: Comparing DID and Machine learning 120

4.4.3 Weather normalisation using Machine learning 122

4.4.4 The augmented synthetic control method (ASCM) 125

4.5 Results 128

4.5.1 DID and parallel trend test results 128

4.5.2 Weather normalised results using machine learning 130
4.5.3 The Impact of the Hebei inspection on local air pollution using a Ridge ASCM approach 132
4.5.4 Placebo tests 135
4.6 A comparison of DID and ML+ASCM 138
4.7 The benefits of reducing pollution 140
4.8 Discussion and Conclusions 141

5 Winter heating and air pollution in China **160**
5.1 Introduction 160
5.2 Policy background 164
5.2.1 Winter heating policy in China 164
5.2.2 Clean winter heating policy in China 166
5.3 Data 168
5.3.1 Air pollution data 168
5.3.2 Meteorological data 169
5.3.3 Additional predictors 169
5.4 Methods 170
5.4.1 Random forest-based weather normalisation method 171
5.4.2 The Augmented Synthetic Control method 173
5.5 Results 175
5.5.1 Variable importance 175
5.5.2 Impact of winter heating on local air quality in Northern Chinese cities 176
5.5.3 The effectiveness of the Clean Winter Heating Plan for Northern China (2017-2021) on local air quality 179
5.6 Conclusion and Discussions 185

6 Conclusions **213**

A Chapter 2 Appendices **217**

B Chapter 3 Appendices **225**

C Chapter 4 Appendices **229**

D Chapter 5 Appendices **247**

References **249**

List of Figures

2.1 "CEIP" inspection process 43

2.2 Time series of AQI by city group and time periods. 44

2.3 T statistics distribution of 1,000 times '(fake) prehebei' 45

2.4 T statistics distribution of 1,000 times '(fake) during hebei'. 46

2.5 T statistics distribution of 1,000 times '(fake) afterhebei'. 47

3.1 The annual average observed concentrations of SO_2, NO_2, CO and PM10 in Wuhan and 29 control cities between 2013 and 2019. 94

3.2 Daily averages of observed and weather normalised concentrations of SO_2, NO_2, CO and PM10 in Wuhan between January 2013 and February 2020. . . 95

3.3 The comparison of daily observed and weather normalised concentrations of SO_2, NO_2, CO and PM10 in Wuhan between 21st December 2019 and 3rd February 2020. 96

3.4 Ridge ASCM results on weather normalised NO_2 concentrations in Wuhan. . 97

3.5 Ridge ASCM results on weather normalised SO_2 concentrations in Wuhan. . 97

3.6 Ridge ASCM results on weather normalised CO concentrations in Wuhan. . 98

3.7 Ridge ASCM results on weather normalised PM10 concentrations in Wuhan. 98

3.8 The in-time placebo test results of NO_2wn using 21st January 2019 (left) and 21st January 2018 (right) as Wuhan lockdown date. 99

3.9 The in-time placebo test results of PM10wn using 22nd January 2019 (left) and 22nd January 2018 (right) as Wuhan lockdown date. 99

3.10 The results of in-place placebo test on NO_2wn . . . 100

3.11 The results of in-place placebo test on PM10wn. . . . 101

3.12 The results of alternative control group tests on NO_2wn (left) and PM10wn (right). . . . 101

4.1 The plot of pre-inspection parallel trend tests . . . 143

4.2 The comparison of weekly observed and weather normalised concentrations of $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei between January 2015 to June 2016. . . . 144

4.3 The weekly concentration of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei and 27 control cities between January 2015 to June 2016. . . . 145

4.4 The variable importance of $PM_{2.5}$, PM10, NO_2 and SO_2 random forest model in Shijiazhuang . . . 146

4.5 Ridge ASCM results on weather normalised and observed $PM_{2.5}$ and PM10 in Hebei. . . . 147

4.6 Ridge ASCM results on weather normalised and observed NO_2 and SO_2 in Hebei. . . . 148

4.7 The heterogeneity effects of inspection on weather normalised pollutants in 8 Hebei cities. . . . 149

4.8 The in-time placebo test results of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei. . . . 150

4.9 The in-place placebo test results of weather normalised $PM_{2.5}$, PM10 and SO_2 in Hebei. . . . 151

4.10 The results of alternative control group tests on $PM_{2.5}$wn, PM10wn and SO_2wn in Hebei. . . . 152

5.1 Variable importance of SO_2 in Beijing . . . 187

5.2 Variable importance of $PM_{2.5}$ in Beijing . . . 188

5.3 Variable importance of PM10 in Beijing . . . 189

5.4 Weather normalised concentrations of daily SO_2 in 15 northern Chinese cities before and after the winter heating turn on between 2015 and 2021. 190

5.5 Weather normalised concentrations of daily SO_2 in 15 northern Chinese cities before and after the winter heating turn off between 2015 and 2021. 191

5.6 Weather normalised concentrations of daily SO_2 in 14 southern Chinese cities before and after the hypothetical winter heating turn on date between 2015 and 2021. 192

5.7 Weather normalised concentrations of daily SO_2 in 14 southern Chinese cities before and after the hypothetical winter heating turn off date between 2015 and 2021. 193

5.8 The trend plot of observed and weather normalised SO_2, HDD5 and temperature in Beijing between 2015 to 2021 194

5.9 The HDD5-SO_2 response plot of Beijing in six heating seasons 195

5.10 The HDD5-SO_2 response plots in 15 northern cities between 2015 and 2021 . 196

5.11 The HDD5-SO_2 response plots in 14 southern cities between 2015 and 2021 . 197

5.12 The HDD5-SO_2 response difference plots in 15 northern cities between 2015 and 2021 198

5.13 The HDD5-SO_2 response difference plots in 14 southern cities between 2015 and 2021 199

5.14 The HDD5-NO_2 response difference plots in 15 northern cities between 2015 and 2021 200

5.15 The HDD5-NO_2 response difference plots in 14 southern cities between 2015 and 2021 201

5.16 The HDD5-$PM_{2.5}$ response difference plots in 15 northern cities between 2015 and 2021 202

5.17 The HDD5-$PM_{2.5}$ response difference plots in 14 southern cities between 2015 and 2021 203

5.18 The HDD5-PM10 response difference plots in 15 northern cities between 2015 and 2021 . . . 204
5.19 The HDD5-PM10 response difference plots in 14 southern cities between 2015 and 2021 . . . 205
5.20 The HDD5-CO response difference plots in 15 northern cities between 2015 and 2021 . . . 206
5.21 The HDD5-CO response difference plots in 14 southern cities between 2015 and 2021 . . . 207
5.22 The weekly observed and deweathered PM2.5, PM10, CO, SO2, NO_2 concentrations and SO2/NO_2 ratios from 2015 to 2021 in the three regions. . . . 208
5.23 Observed, deweathered and synthetic differences in weekly PM2.5 and SO2 for different regions . . . 209
5.24 Observed, deweathered and synthetic differences in weekly PM10, CO and NO_2 for different regions . . . 210

A.1 Calculating the AQI . . . 221

B.1 The location of Wuhan and the 29 control cities . . . 225

C.1 The location of Hebei cities and control cities . . . 233
C.2 The plots of 72-hour back-trajectory clusters in 4 selected cities . . . 236
C.3 The plot of pre-inspection parallel trend tests on observed pollutants . . . 237
C.4 The plot of pre-inspection parallel trend tests on weather normalised pollutants 238
C.5 The partial dependence of meteorological variable in $PM_{2.5}$ random forest model in Shijiazhuang . . . 239
C.6 The partial dependence of meteorological variable in PM10 random forest model in Shijiazhuang . . . 240

C.7 The partial dependence of meteorological variable in NO_2 random forest model in Shijiazhuang 241
C.8 The partial dependence of meteorological variable in SO_2 random forest model in Shijiazhuang 242
C.9 The heterogeneity effects of inspection on weather normalised pollutants in Shijiazhuang, Cangzhou, Xingtai and Baoding 243
C.10 The heterogeneity effects of inspection on weather normalised pollutants in Tangshan, Zhangjiakou, Qinhuangdao and Chengde 244
C.11 The plot of weekly weather normalised pollutants in 8 Hebei cities 245

D.1 The location of "2+26", "North" and "South" cities in China. 247

List of Tables

2.1 The health categories and meanings for AQI 48

2.2 Summary of daily average AQI by periods and city groups 49

2.3 Summary of percentage change of daily average AQI by periods and city groups 50

2.4 Summary of daily average $PM_{2.5}$ by periods and city groups 51

2.5 Summary of daily average PM10 by periods and city groups 52

2.6 Summary of daily average SO_2 by periods and city groups 53

2.7 Summary of daily average CO by periods and city groups 54

2.8 Summary of daily average NO_2 by periods and city groups 55

2.9 Summary of daily average O_3_8h by periods and city groups 56

2.10 Main results of inspection on air quality in Hebei cities (time setting 1) . . . 57

2.11 Main results of inspection on air quality in Hebei cities (time setting 2) . . . 58

2.12 Main results of inspection on air quality in Hebei cities (time setting 3) . . . 59

2.13 The results of only criticised cities in Hebei 60

2.14 The results of only un-criticised cities in Hebei 61

2.15 Results of inspection on air quality in Hebei using a 1 year benchmark . . . 62

2.16 Results of inspection on air quality in Hebei using a January benchmark . . 63

2.17 Results of inspection on air quality in Hebei after dropping bordering cities 64

3.1 Sources and health effects of our four pollutants 102

3.2 A list of input and output variables used in this study 103

3.3 Previous literature on the NO_2 mortality association 104

3.4 Previous literature on the NO_2 mortality association 105

4.1 A list of input variables (predictors) used in this study 153

4.2 Random forest model performance in Shijiazhuang 154

4.3 The impact of CEIP on air quality in Hebei (levels) 155

4.4 The estimated impact of CEIP on air quality in Hebei using different methods 156

4.5 Sources and health effects of our four pollutants 157

4.6 Summary of previous literature on the PM10 mortality association 158

4.7 The potential lives saved due to falling PM10 attributable to CEIP in Hebei and China 159

5.1 The causal impact of winter heating on SO2wn, $PM_{2.5}$wn and PM10wn in "2+26" and "North" in each heating season 211

5.2 The causal impact of winter heating on NO_2 wn, COwn and O3wn in "2+26" and "North" in each heating season 212

A.1 China five-year plan pollution reduction targets and achievements (10th Five-Year Plan for environmental protection: major targets vs. performance) . . . 222

A.2 Major targets in 11th Five-Year Plan for environmental protection 223

A.3 Major targets in 12th13th Five-Year Plan for environmental protection: targets vs. performance 224

B.1 Meteorological monitoring station information used in the research 226

B.2 Control groups used in the main analysis and sensitivity analysis 227

B.3 Selected socio-demographic data in Wuhan and 29 control cities in 2018 . . . 228

C.1 The impact of CGEP inspection on air quality in Hebei (logs) 234

C.2 The estimated impact of CEIP on air quality in Hebei using alternative weighting methods 235

C.3 Control groups used in the main analysis and sensitivity analysis 246

D.1 Winter heating start and end time in 29 cities 248

Chapter One

Introduction

Air pollution has become an increasingly major concern across the world and is considered by many to be the single largest environmental risk that humans face today. China's rapid economic development in recent decades was accompanied by a serious deterioration in air quality. In response to the quality of the air and its potentially damaging health impacts, the Chinese government has employed a range of different strategies and policies to mitigate air pollution problems. As a result, it is of great interest and importance to policymakers and the general public to understand the efficacy of such policies.

The primary objective of this thesis is to investigate the effect of selected Chinese government policies on air quality. The thesis contributes to the environmental economics literature by employing a set of cutting-edge interdisciplinary research methods from economics, atmospheric science, and computer science. By combining methods from different disciplines we are able to obtain results that are grounded in both science and social science so we are able to contribute to the literature across each of these disciplines. This thesis consists of six chapters. Chapter 1 provides a brief introduction, chapters 2, 3, 4 and 5 present the main body of the thesis and can be considered as four independent papers. Finally, chapter 6 concludes.

Chapter 2 investigates the effectiveness of China's Environmental Protection Inspection program that started in 2016 and ran through to 2017, a policy unparalleled in terms of its authority, scope and stringency. The implementation of the policy involved sending large teams of centrally appointed inspectors to a number of cities in a chosen province to inspect firms and to urge local government and party officials to take greater actions to protect the environment. Using an air quality index (AQI) and its constituent pollutants, we examine whether the inspection policy was effective in improving air quality in the targeted cities before, during and after the inspection period. We focus on the first inspection (January 2016) in Hebei province that acts as a quasi-natural experiment. Relative to a control group, our results show that there was a small decrease in the recorded AQI in the six months before the inspection period, a fairly large drop during the month of the inspection, and a longer-term effect of up to five months after the inspection team left Hebei. The results are robust to various sensitivity checks and placebo tests. Policy implications are discussed.

Chapter 3 quantifies the impact of the Wuhan Covid-19 lockdown on concentrations of four air pollutants using a two-step approach. First, we use machine learning to remove the confounding effects of weather conditions on pollution concentrations. Second, we use a new Augmented Synthetic Control Method (Ben-Michael, Feller and Rothstein, 2021) to estimate the impact of the lockdown on weather normalised pollution relative to a control group of cities that were not in lockdown. We find NO_2 concentrations fell by as much as 24 $\mu g/m^3$ during the lockdown (a reduction of 63% from the pre-lockdown level), while PM10 concentrations fell by a similar amount but for a shorter period. The lockdown had no discernible impact on concentrations of SO_2 or CO. We calculate that the reduction of NO_2 concentrations could have prevented as many as 496 deaths in Wuhan city, 3,368 deaths in Hubei province and 10,822 deaths in China as a whole.

Chapter 4 returns to the recent Central Environmental Inspection Policy (CEIP) and examine how effective it was in reducing local air pollution at the regional and city level.

Introduced in 2016, the CEIP is now one of the central pillars of China's environmental policy regime with the aim of encouraging local government officials to exert greater effort to protect the environment. Centrally-appointed inspection teams were dispatched to cities across China for about a month to inspect both the local government officials and the firms in the cities. Methodologically, we combine weather normalisation techniques (a random forest-based machine learning model) from atmospheric sciences and the Augmented synthetic control method (ASCM) to provide a causal estimate of the impact of a province having a CEIP inspection on local air quality. We find that the Hebei inspection led to a substantial, but short term, reduction in the emissions of PM2.5, PM10 and SO_2 in Hebei as a whole. The weather normalised value of PM2.5 and PM10 fell by as much as 25.89 μg/m^3 (20.81%) and 47.80 μg/m^3 (22.61%) immediately after the inspection, although pollution levels gradually returned to previous levels within 3 months. Simple back-of-the-envelope calculations suggest that, the 3-month reduction of PM10 that can be attributed to the CEIP inspection would have reduced the number of deaths in Hebei by between 565 and 1,048.

In the final empirical chapter, Chapter 5 estimates the impact of the winter heating policy and the effectiveness of the clean winter heating plan on air quality in Chinese capital cities. After decoupling the impact of meteorological conditions on observed air pollutant concentrations, this chapter finds that, during the winter period in 2015 and 2016, the turning on of the winter heating system immediately increase the air pollution level (SO_2) across northern Chinese cities while the turning off of the winter heating system led to the sudden drop of air pollution. Moreover, such immediate deterioration (/recover) on air quality as the winter heating system turn on (/off) was less evident after 2017, and gradually disappear in the winter period in 2019 and 2020. The results suggests that the clean heating plan which was implemented progressively since 2017 was effective in reducing air pollution. Finally, through the partial dependence relationship between winter heating (indicated by the HDD5 index) and air pollution, this chapter quantifies the contribution of winter heating on air pollution level across all cities in each winter period and shows substantial reductions in air pollution in 11 (out of 15) northern cities. Our results suggest that, following a series of regulations implemented under the 'Clean winter heating plan for Northern China (2017-2021)' since 2017, the air quality has largely improved in northern Chinese cities, demonstrating that the large scale clean heating plan was an effective measure to mitigate air pollution problems in northern China.

Chapter Two

Air Pollution and Environmental Policy in China: Testing the Effectiveness of the Environmental Inspection Policy

2.1 Introduction

China's rapid economic development in recent decades has come at a cost to the environment and to the health of the Chinese population. The World Health Organization (WHO) (2018) finds that ambient air pollution led to more than one million deaths in China in 2016 while annual average $PM_{2.5}$ levels in the Beijing-Tianjin-Hebei region have been over ten times WHO recommended limits (Wong and Karplus, 2017).[1] Similarly, Chen, Ebenstein et al. (2013) show that air pollution reduced the life expectancy of people living in northern China by 5.5 years. In response to these significant health impacts, the Chinese government has

[1] WHO Issues Latest Global Air Quality Report: Some Progress, but More Attention Needed to Avoid Dangerously High Levels of Air Pollution. Retrieved December 23, 2019, from (http://www.wpro.who.int/china/mediacentre/releases/2018/20180502-WHO-Issues-Latest-Global-Air-Quality-Report/en/).

made the reduction of air pollution a primary objective in recent years and has employed a range of different strategies and policies, both carrot and stick, at local and national levels. A recent major policy initiative has been the Central Environmental Inspection Policy (CEIP) which has involved large, centrally-appointed, inspection teams being dispatched to cities for a month at a time with the aim of achieving long-term reductions in local air pollution.

This paper provides the first analysis of the effectiveness of the CEIP, a policy with an unprecedented level of authority, scope and toughness relative to other Chinese environmental policies. Uniquely, we are also able to examine whether pollution fell in advance of the implementation of the CEIP (if firms responded to rumours of the impending inspections and took pre-emptive action), during the one-month inspections themselves and, perhaps more importantly, in the five month period after the inspections ended.

To examine the effectiveness of the CEIP policy we collect daily meteorological data and air pollution concentration data for $PM_{2.5}$, PM10, SO_2, NO_2, CO and O3 and calculate an Air Quality Index (AQI) for 334 Chinese cities from 2013 to 2017. This paper focuses on the first inspection (January 2016) that took place in Hebei province and compares the air quality in the eleven targeted cities in Hebei before, during and after the inspection with a number of different control groups. In further analysis we see whether air quality responded differently to the inspection across the targeted cities by dividing the cities into those that were criticized after the inspection (according to the official report subsequently published by the Ministry of Environmental Protection) and those that were not. Finally, as part of our robustness checks we estimate our results for different pre and post-inspection periods.

Our paper contributes to a broader literature on the effectiveness of environmental policy interventions. One of the first examples of a policy aimed at permanently reducing air pollution was the US Clean Air Act Amendments (CAAA) that led to a dramatic decline in concentrations of SO_2, total suspended particulates (TSPs) as well as particulate matter

(PM10) (Chay and Greenstone, 2003a; Greenstone, 2004; Auffhammer, Bento and Lowe, 2011). On a smaller scale, the 2008 Beijing Olympic Games provides an opportunity to identify the effects of temporary and strictly-enforced air pollution regulations (He, Fan and Zhou, 2016). Chen, Jin et al. (2013) found that the actions taken by the Chinese government at the time improved air quality in Beijing by 24.9%, although air pollution levels had returned to pre-games levels by the end of 2009. Similarly, Li, Qiao et al. (2017) examined the impact of the implementation of temporary environmental regulation during the 2014 APEC meeting and the 2015 Victory Day (the quasi-experimental design) in China and found a short term improvement of air quality in Beijing (a reduction in the AQI by 37.4% during the APEC Meeting and 35.9% during the Victory Day Parade).

However, not all environmental policies have been so successful, particularly in less developed countries where institutions may be weaker and less developed (Duflo et al., 2013; Greenstone, 2004; Sun, Zheng and Wang, 2014). For example, Davis (2008) investigated the effect of a driving restriction policy in Mexico City, and found no evidence that the restrictions improved ambient air quality. Likewise, Duflo et al. (2018) estimated the effects of environmental inspections on industrial plants in India and found that the treatment effect was weak and that there was no decrease in average pollution emissions. However, the centralised political system in China may mean that environmental policies are more strictly adhered to (Zheng and Kahn, 2013). This is particularly true of the CEIP which emanated from, and was managed by, senior governmental leaders who provided inspectors with considerable authority, scope and stringency in order to reduce air pollution.

To briefly summarise our results, we find that the CEIP inspection policy did reduce air pollution in the eleven Hebei cities that were inspected. Compared to the benchmark period (1st December 2013 to 31st June 2015) the policy reduced the AQI, $PM_{2.5}$, PM10 and O3 concentration levels before, during and after the inspection period. The largest impacts occurred during the five and a half months after inspectors left which saw a 47.3% fall in the

AQI level, a 59.6% drop in $PM_{2.5}$, a 55.3% drop in PM10, a 31.1% drop in SO_2, and a 31.7% drop of CO. A range of alternative benchmarks and placebo tests provide further support for the effectiveness of the inspection policy.

The remainder of the paper is structured as follows. Section 2.2 discusses briefly the institutional background of environmental regulations and introduces the Central Government Environmental Protection Inspection program. Section 2.3 presents the literature review. Section 2.4 describes relevant data resources and summary statistics. Section 2.5 discusses the specification approach. The main results and robustness checks are presented in section 2.6. Finally, section 2.7 concludes.

2.2 The Central Environmental Inspection Policy

The 14th meeting of the Central Committee Party and Central government overall reform leadership group was chaired by President Xi on July 1st of 2015 with the goal of promoting environmental protection and accelerating sustainable development. The meeting established an environmental inspection mechanism and on the same day the Environmental Protection Inspection Plan was announced. This inspection policy marked a notable departure from previous environmental policies and was a major institutional arrangement coming directly from central government.

The inspection teams were tasked with investigating environmental issues and providing the public with instant open information. The central governmental inspectors are authorized to schedule meetings or one-to-one talks with top local government officials, to investigate firms without prior warning and to hold any government officials to account for poor environmental performance. The motivation behind an inspection campaign is that is thought to be an effective mechanism by which central government is able to transmit pressure for environmental protection from the highest level to the local authorities at the city level.[2] The inspections also allowed the central government to understand how provincial-level, municipal-level and county-level CCP Committees and governments implement policies and regulations dealing with environmental protection, how they deal with challenging environmental problems, and who is given the main responsibility for protecting the local environment. The inspection policy specifically addresses the "Party-government equal responsibility" principle that means that local party committees and the central government are inspected together. The inspection results were seen as an important input into the hiring and firing of local party and government leaders and clearly spell out that role of local government and committee official in dealing with environmental protection problems.[3]

Hebei province was chosen as the first inspection target and was timed for the month of January, 2016.[4] Following the Hebei inspection, the next round of inspections took place from 15 July to 15 August 2016 and covered cities in eight different provinces with another round taking place between 25 November to 25 December 2016 aimed at seven cities and provinces. The next inspection round took from 25 April to 25 May 2017 targeting seven provinces while the final round of inspections was from 15 August to 15 September 2017 and covered the remaining eight provinces.[5]

At the end of this two-year long inspection campaign, the central inspection team had covered all of China. In each case the central government inspection teams were stationed

[2] http://usa.chinadaily.com.cn/epaper/2017-04/17/content$_2$8962789.*htm*.

[3] http://www.chinadaily.com.cn/china/2017-08/08/content$_3$0368011.*htm*.

[4] Hebei province has 11 district level and above cities: Shijiazhuang, Baoding, Cangzhou, Chengde, Handan, Hengshui, Langfang, Qinhuangdao, Tangshan, Xingtai and Zhangjiakou.

[5] Provinces inspected from 15 July to 15 August 2016 include Inner Mongolia, Heilongjiang, Jiangsu, Jiangxi, Henan, Guangxi, Yunnan and Ningxia; Inspected from 25 November to 25 December 2016 include Beijing, Shanghai, Hubei, Guangdong, Chongqing, Shaanxi, Gansu; Inspected from 25 April to 25 May 2017 include Tianjin, Shanxi, Liaoning, Anhui, Fujian, Hunan and Guizhou; Inspected from 15 August to 15 September 2017 include Jilin, Zhejiang, Shandong, Hainan, Sichuan, Tibet, Qinghai and Xinjiang.

in the target provinces/region for around one month. During this time, letters and phone calls from the public on the topic of environmental protection target provinces (regions) were accepted and processed by the inspection teams. [6]

Figure 2.1 illustrates the how the CEIP inspection programme operated which, generally speaking followed a seven-step process.

[Figure 2.1 about here]

(1). Inspection preparation period: The state council directly choses the inspection team leaders and members. In the case of Hebei, the team leader and vice leader were the former and current vice minister of the MEP, the other 20-30 members consisted of relevant inspection experts, including officials and staffs from the MEP, the Ministry of Natural Resources, and State Council).

(2). Inspection team is stationed in the targeted area: Inspection teams are usually located in the targeted area for one month after which they follow three distinct steps. The first step is an inspection at the province level. At this stage the team listen to reports on environmental issues, accept public tip-offs, hold one-to-one talks, get access to and extract pollution related materials,. The team meet with relevant province level government departments and party leaders, including the Development of Natural resources, Forestry, Housing, Agriculture, Urban management, Industry and Information Technology, Transportation, Tourism, Development and Reform and others. The individuals that are interviewed are expected to tell the truth regarding issues related to the environmental without evading questions, or exaggerating or giving misleading answers. The second step, based on information and material collected from the province level inspections, is to visit cities and

[6] It is worth noting that although this was not known at the time of the first inspections, the Chinese central government has recently implemented a new inspection series (CEIP Inspection "Revisit" program) that spanned the period 2018 and 2019 with a remit of improving the effectiveness of inspection programmes.

counties within that province to make spot checks, investigate whether the problems that have been suggested can be validated and who is responsible (see Appendix 2 for details). The third and final step is to sort, classify and file all of the relevant documents associated with the inspection. [7]

(3) & (4). Write reports and publish feedback: All the information, materials and problems are classified and analysed by the inspection team, and an official report and feedback published by the central government and available to both the general public and the government officials.

(5). Case transfer: The inspection results (including some cases that are relate to serious environmental issues) are transferred through one of three channels. The first is to transfer cases to the ministerial–level departments (including the Ministry of Organization, the Ministry of Supervision and the Ministry of Finance, which play key roles in the assessment, appointment and removal of cadres, and central financing arrangements, respectively (Zhao and Percival, 2017). The second is to transfer to the MEP for direct treatment. For those areas with extreme environmental problems, the central government promises to severely punish the relevant individuals and departments and even to dismiss them. Third, is to transfer cases to local party committees and local government officials.

(6). Rectification and implementation: It is a requirement that local government and party officials are dedicated to solving problems, e.g. they must put forward plans to rectify any problems, clearly define responsibilities, suggest time limits per rectification, and to strengthen future supervision.

(7). Filing and archives: Following the end of an inspection, all relevant materials and information is available for future research and to be analysed by central government with a view to guiding future policymaking.

[7] https://www.guancha.cn/local/2016$_0$5$_1$2$_3$60055.$shtml$.

Compared with previous environmental policies and regulations, the CEIP has three distinct features. First is the high political level of the inspected targets that includes province level party committee and government officials as well as leaders of the relevant departments. During previous environmental inspections, it was thought that province level leaders would often act as a protective umbrella for local officials and enterprises and 'turn a blind eye' to local environmental violations as long as local economic growth continued, given economic performance was and remains one of the main criteria by which officials can get promoted (Wang, 2013).[8] However, in the CEIP inspection campaign, province level officials have to worry about themselves first rather than attempting to protect lower level officials or companies within their regions. Given the inspection teams are led by ministerial-level officials they have the ability to impact the promotion prospects of thousands of officials.[9] For example, during the first official inspection round, 1,140 officials were held accountable including 130 departmental level (prefectures or city level equivalent) and 504 divisions level (county level equivalents). Under pressure from inspectors, local officials had to act quickly and potentially even act to shut down problematic factories before the inspection teams arrived. The inspection teams can hold local officials, including province or city level leaders, responsible for environmental problems and under certain circumstances local officials or different departmental leaders are publically exposed and hence open public criticism and in some cases the officials can be dismissed from their positions.[10] [11]

The second key feature of the CEIP is the level of authority and broad scope of the policy. The inspection team was approved by CCP Central Committee and State Council, and consists of current or retired provincial and ministerial officials and given the name

[8]http://www.scmp.com/news/china/policies-politics/article/1987159/china-sends-environmental-inspection-teams-8-more.

[9]http://www.straitstimes.com/asia/east-asia/china-punishes-18000-companies-for-pollution.

[10]https://www.nytimes.com/2017/06/13/world/asia/china-companies-air-pollution-paris-agreement.html.

[11]http://www.mee.gov.cn/xxgk/hjyw/201706/t20170614$_4$15973.*shtml*>.

Central Environmental Protection Inspection Team. The structure of the CEIP is in sharp contrast to previous policies where policy enforcement officers from the MEP were often treated with defiance and resistance and in severe cases were prohibited from visiting certain factories or even being illegally detained by local officials and firms. For example, local officials were thought to inform plants to shut down operations in advance to avoid inspection only for the closed factories to secretly reopen after the inspection team had departed (Wang, 2013). However, because the CEIP teams are considered to be ‘imperial envoys’ it was hoped (assumed) that they would be more effective.[12] Equally important is the engagement of two influential party agencies called the Central Commission for Discipline Inspection and the Central Organisation Department, both of which are arguably the most important in determining the promotion prospects of government officials. The nature of China’s centralised government is that they are also able to take radical actions quickly at a large scale (Chen, Ebenstein et al., 2013).

The third feature is that it was hoped that the initial inspections would help to establish a longer-term monitoring mechanism. Although investigating specific pollution cases was part of the remit a second primary objective was to clearly define who had the responsibility for environmental protection among local party and government officials (Chen, Jin et al., 2013). It was also promised that the inspections were not a one-off crackdown but would continue as a biennial event as evidenced by the new CEIP revisited programme.[13] The long term nature of the inspection process was hoped to provide an incentive on local officials and enterprises to properly comply with environmental regulations and make robust plans for environmental management of their city and province.

According to the reports from the inspection teams and other local level government officials, across all inspection rounds, more than 17,709 government officials were disciplined,

[12]http://www.straitstimes.com/asia/east-asia/china-punishes-18000-companies-for-pollution.

[13]https://thediplomat.com/2017/12/china-cleans-up-its-act-on-environmental-enforcement/.

26,837 companies were punished, 98,057 public tip-offs were received and total fines exceeded 1 billion Chinese yuan (around 150 million U.S. dollars). During the inspection process a range of severe environmental problems were uncovered including pollution data fabrication, weak and lax pollution control, insufficient enforcement of regulations and even the protection of companies by local governments. [14]

2.3 Literature Review

Environmental policy evaluations have always created a heated debate and become increasingly significant for economists and policy makers to adjust and design optimal environmental regulations in both developed and developing countries.

2.3.1 Long term policy evaluation in developing and developed regions

A large literature has studied the direct effect of environmental policy on ambient quality improvements in developed countries.

The 1970 Clean Air Act Amendments (CAAA) was one of the most significant environmental policy interventions from the U.S. federal government to regulate air pollution. The establishment of the Environmental Protection Agency (EPA) in 1970 further stressed the government efforts to protect the environment. The EPA set national ambient air quality standards for four pollutants, including carbon monoxide (CO), tropospheric ozone (O3), sulfur dioxide (SO_2), and total suspended particulates (TSPs). All counties are required to

[14] See for example, https://www.nytimes.com/2017/06/13/world/asia/china-companies-air-pollution-paris-agreement.html.

meet the minimum standard level (Greenstone, 2004).The results of this policy evaluation and its effect on ambient quality is mixed in past due to estimations on different pollutants.

Chay, Dobkin and Greenstone (2003) found from 1971 to 1974, a sharp break appeared in the trend for TSPs concentrations in nonattainment counties after implementation of the CAAA. Newly regulated counties enjoyed a dramatic 20-µg/m3 reduction in total suspended particulates (TSPs), indicating that the policy intervention contribute to the total national decline in TSPs from 1971 to 1974. Consistent with this study, using the monitor level PM10 concentrations data obtained from US EPA from 1990 to 2005, Auffhammer, Bento and Lowe (2009) empirically examine whether the reduction of ambient PM10 concentrations across the US can be attributed to the CAAA interventions. After controlling for heterogeneous monitor and county treatment effect, they found that the CAAA significantly reduce PM10 concentrations by around11% to 14% for the monitors with concentrations exceeding the national annual standard.

Conversely, using county level SO_2 concentration data from 1969 to 1997, Greenstone (2004) empirically test whether the CAAA contributed to the 80% decrease of SO_2 during that period and found that nonattainment designation plays a minor role in the dramatic decline of SO_2 concentrations. [15] Consistent with this study, Auffhammer, Bento and Lowe (2011) tested the effects of environmental regulations on smaller regulated area by CAAA, employing city level PM10 concentrations data from 1068 California cities. They investigate the determinants of the fluctuations in PM10 concentrations during the period 1990 to 2000 and find that the direct impact of CAAA regulations is not significant in explaining the changes in PM10 concentrations, indicating that the CAAA's relatively had little or no impact on pollution improvement.

[15] The core of this legislation is the annual assignment of all counties to SO_2 nonattainment or attainment categories. Polluters face stricter regulations in nonattainment counties (Greenstone, 2004).

For developing countries, it is suggested that institutional differences may lead to different findings. Similarly to USA, the Indian central government published the Water Act of 1974 and Air Act of 1981 that provide comparable channels on the effects of policies. Since the air and water regulations were implemented and enforced in different manners, a comparison of their relative effectiveness can shed light on how to design policy successfully in weaker regulatory contexts (Gray and Shimshack, 2011).

Using pollution data from 572 air pollution monitors in 140 cities in India from 1987-2007, 489 water pollution monitors in 424 cities between 1986 and 2005, as well as annual city-level infant mortality data, Greenstone and Hanna (2014) tested the impact of key air and water pollution policies in India. Air pollution regulations are associated with substantial improvements in air quality whereas water regulations failed to cause improvements in three available measure of water pollution. Additionally, the most successful air pollution regulations caused a modest but insignificant reduction in infant mortality (Greenstone and Hanna, 2014). These findings demonstrate the possibility of environmental regulations' success in improving ambient quality in developing countries where institutional settings are relatively weak.

Other environmental policy evaluation papers look at the indirect effects of policy intervention by looking at the effect of pollution on public health, industries and even house prices and school attendance.

Using county level individual and infant birth record and micro economic data from USA during the period 1969 to 1974, Chay, Dobkin and Greenstone (2003) examined the reduction in total suspended particulates (TSPs) pollution induced by the 1970 CAAA to investigate the effects of TSPs on infant mortality. They found that the 1970 CAAA Air pollution regulations are associated with sharp decline in both TSPs pollution and infant mortality. A 1% decrease in TSPs is estimated to cause a decrease a 0.5% in infant mortality. Chay and Greenstone (2003b) evaluated the same Clean Air Act Amendments in 1970 and examined the impact of TSPs pollution reductions induced by aggressive air pollution regulations on adult health, using the county level mortality data from National Center for Health Statistics (NCHS) between 1969 and 1974. They found that on the contrary to the previous study (Chay, Dobkin and Greenstone, 2003), regulation-induced decline in TSPs pollution has little effect on either adult or elderly mortality.

To estimate the relationship between environmental regulations and Foreign Direct Investment (FDI) and to measure the response of US-based multinationals to the Clean Air Act Amendments (CAAA), Hanna (2010) used firm level panel data between 1966- 1999 to investigate the effect of regulation on firms' foreign production decisions. The evidence suggests that the CAAA caused US based multinationals to increase foreign assets by 5.3% and foreign output by 9% in response to stricter regulations, which is equivalent to about 0.6% of pollution industries' multinationals' domestic assets. By using detailed production data in 1.2 million plants from 1972-1993 in the U.S., Greenstone, List and Syverson (2012) estimate the effects of pollution regulations on manufacturing plant's total factor productivity (TFP) and found that a roughly 2.6 percent reduction in TFP is estimated to be attributed to stricter air quality regulations for surviving polluting plants. The productivity loss corresponds to approximately $11.0 billion dollars (in 2010 dollars). To provide empirical evidence confirming the pollution Heaven Hypothesis, Cai, Chen and Gong (2016) estimated the effect of TCZ policy on inbound FDI to China and found that an increase in pollution intensity decreases 8% points of FDI, indicating that tougher environmental regulation lead to less foreign direct investment. In a different study related to the estimation of effects for the TCZ, Tanaka, Yin and Jefferson (2014) showed that more stringent environmental regulations positively improved productivity in China.

2.3.2 Short term policy evaluation in developing and developed regions

Compared with those permanent and long-term environmental regulations like the U.S. Clear Air Act and the Indian Air Act and Water Act, there is also a rich literature that evaluates short-term or temporary environmental policies change and its effects.

Driving restrictions have been introduced in a number of cities all over the world with the main aim of reducing traffic congestion. However, driving restrictions may impact air quality in the restricted areas (Davis, 2008). In the US, Currie and Walker (2011) examine the impact of the "E-ZPass" which is an electronic toll collection system that enables vehicles with a special tag to drive through toll lanes without stopping to manually pay a toll and using health data from Pennsylvania from 1997 to 2002, and New Jersey from 1994 to 2003, they find that both traffic congestion and vehicle emissions near highway toll plazas were greatly reduced but more importantly, they find that premature and low birth weight among mothers within 2 km of a toll plaza were reduced by 10.8% and 11.8% respectively. Similarly, Venigalla and Krimmer (2006) studied E-ZPass adoption for the case of the George Washington Bridge toll plaza and found that NO_2 and CO emissions decreased by 5.8% and 23.5% respectively.

On the other hand, in developing countries, to deal with the high level of air pollution in Mexico City, the Mexico City government introduced a driving restriction policy (Hoy No Circular, HNC) that banned most drivers from using vehicles one weekday per week. Using air quality data from 1986 to 1993, Davis (2008) examined the effect of driving restrictions and found no evidence of an improvement in air quality although he did find a relative increase in air pollution at the weekends and non-peak weekdays that suggests that the HNC program changed the transportation choices of individuals but failed to improve air quality. In contrast, Viard and Fu (2015) used a daily Air Pollution Index (API) from monitoring stations from the State Environmental Protection Agency, in China to show that air pollution in Beijing decreased by 21% during the one-day-a-week restriction that was introduced into Beijing in 2008 as part of the odd-even policy. [16]

In a different policy, the Mexican Clean Industry Program was introduced by the Mexican government in 1997 and was a voluntary pollution reduction program where participating plants pay for a third party audit agency to detect whether the plant meets the pollution standards. Successful plants are granted a Clean Industry Certificate and are exempt from further government inspections for two years. Using city level satellite-based Aerosol Optical Depth (AOD) data measuring air quality, Mexican Industrial Census data from 2000 to 2005, and infant mortality data from Mexican Statistics Office in Mexican, Foster, Gutierrez and Kumar (2009) (2009) evaluated the voluntary air-quality regulations in Mexico and found that certification causes around a 3.6% improvement in air quality based on AOD and a corresponding 16% reduction in infant mortality due to respiratory illnesses.

2.3.3 Environmental protection inspection and air quality

The evidence on the impact of environmental protection inspections on air quality is mixed. Policymakers consider monitoring and inspection to be important tools to ensure compliance and enforcement of environmental regulations (Kagan, Gunningham and Thornton, 2003; Gray and Shimshack, 2011). Various studies examine the effect of environmental monitoring and inspection in developed countries (Shimshack and Ward (2008) for U.S; Telle (2013) for Norway). By exploiting plant-level monthly biological oxygen demand (BOD) and total suspended solids (TSS) pollution discharge data from Quebec Ministry of the Environment from

[16] Based on the last digit of vehicle license plate numbers, the Odd-Even policy requires odd-numbered license plates could drive only on odd-numbered dates and even-numbered only on even-numbered (Viard and Fu, 2015).

1985 to 1990, Laplante and Rilstone (1996) estimated the impact of government inspection on plant emission levels in the pulp and paper industry. The results show that inspection (and the threat of inspection) significantly reduces emissions. Eckert (2004) also provides empirical evidence showing that inspection deters future regulatory violations for pollution emitters although the effect is small.

Within the context of the U.S. Clean Air Act, the EPA implemented an inspection programme for manufacturing plants to evaluate the degree of compliance with the regulations. Hanna and Oliva (2010) used a panel of plant-level air, land and water emissions from 1987 to 2001 and show that air pollution was reduced on average by 15% for 5 years after the inspection without increasing other emissions.

However, the effect of environmental inspections in developing countries does not appear to be as successful. For example, Duflo et al. (2018), in conjunction with the Indian government, conducted a two-year (2009-2011) field experiment which assigned half of 960 industrial plants in the Indian state of Gujarat, to a treatment group to have more frequent inspections relative to a control group. Using pollution data and detailed internal inspection reports collected from government administrative records, they found that although the treatment groups were twice as likely to be inspected and cited for violations, compliance with standards were only marginally higher in the treatment plants and average emissions were not significantly lower (Duflo et al., 2018). A possible explanation is that environmental inspection may not be effective for pollution abatement in developing countries that reply mainly on command-and-control regulations without incentives for compliance.

2.3.4 Environmental policy and air pollution in China

Air pollution and health

Air pollution in China has posed a serious threat to public health for many years now. As a result, a large literature has emerged examining the adverse impacts of air pollution and the impact of different government policies. For example, the “Huai River Policy” was introduced by the Central Chinese government to build a free winter heating system with provision of free coal fuel for boilers only for Northern China (the border being the Qinling Mountains and Huai River). Using air pollution data from World Bank and China Environmental yearbooks, weather data from China Meteorological Administration for 76 Chinese cities from 1981 to 1993, Almond et al. (2009) where the first to demonstrate the impact of China’s Huai River policy on TSP, SO_2 and NOx concentrations. The results showed that the winter heating system led to dramatically higher TSP concentration levels in the north but no impact on SO_2 and NOx concentrations. Similarly, using mortality data from China’s Disease Surveillance Points (DSPs) system, Chen, Ebenstein et al. (2013) evaluate the effect of the Huai River Policy on life expectancy and found that the Huai River policy had a dramatic impact on pollution and human health. TSP concentrations were 55% higher and life expectancy was 5.5 years lower in north of the Huai River as a result of the winter heating system. The research suggests that the Huai River policy led to a loss of more than 2.5 billion years of life for the 500 million residents living in Northern China. The impact changed a little when using PM10 as the main variable and using same mortality and life expectancy data from Chinese Center for Disease Control and Prevention (Chinese CDC) DSP survey. Ebenstein, Fan, Greenstone, He and Zhou (2017) found that the quasi-experimental variation in PM10 induced by China’s Huai River policy were 46% higher and had a causal impact on reducing life expectancy by 3.1 years in the north. The results also indicate that approximately 3.7 billion life-years would be saved if the whole China meets its Class I standards for PM10 concentrations.

Political economy and environmental protection

China's political system plays a key role in dealing with environmental pollution issues. The official leaders of each province, city and county were directly appointed by upper level CCP Committee and governmental officials based on performance evaluation (Zheng and Kahn, 2017).[17] Therefore, each local government acts as a "competing enterprise" when it comes to promotion (Zheng, Kahn, Sun et al., 2014). For a long time, economic growth was the main performance indicator evaluated by upper-level officials to decide political promotions (Li and Zhou, 2005; Chen, Li and Zhou, 2005; Kostka and Xiaofan, 2015) where the so-called "Political tournament" mechanism provides a strong incentive for local government officials to achieve developmental goals.

More recently, China has acknowledged that it is important for the country to improve its environmental performance in terms of domestic resource shortages but also the increasingly severe pollution problems that threaten public health. Reducing emissions will also help China to embrace the global sustainable and low-carbon agenda to present an international leading image to the rest of the world (Wang, 2013). As a result, the central government has adjusted the promotion criteria of local officials from pure economic based to include a greater emphasis on environmental goals (Zheng and Kahn, 2013).

There is a large literature that examines the effectiveness of Chinese central government policies as a way to incentivise local government officials. Using city-level pollution data from Chinese statistical yearbooks and U.S. satellite data, together with characteristics

[17] The Chinese governance structure consists of five vertical hierarchies: the nation, Province (or Municipality, autonomous region), prefecture-level city (or prefecture-level district), county (or county level district, county-level city), and township (Zheng and Kahn, 2013)

of city level government officials such as age, education level and length of service at current position, Chen, Li and Lu (2018) tested the effect of target based performance evaluation system for SO_2 emission reductions under the TCZ policy in China. Using a DID estimation approach, they found that local government officials put more effort into pollution reduction in TCZ cities after central the government announced that SO_2 emissions reductions were now one of the main criteria for performance evaluation of city level CCP committees and mayors in 2006. The result of the performance evaluation system reduced SO_2 emissions by 14.6% in TCZ cities (compared with non-TCZ cities) and demonstrates that adjustments to the political evaluation system could be an effective channel through which local bureaucrats can be incentivised to pay more attention to environmental protection (Chen, Li and Lu, 2018).

Kahn, Li and Zhao (2015) examine a similar regime shift but focus on water pollution indicators.[18] Using water pollution data from 499 water quality monitoring stations from 2004 to 2010, using a DID estimation, they found that chemical oxygen demand (COD) fell significantly after the policy regime shift in 2006 and COD levels fell faster at the province boundaries rather than inside provinces, which provides empirical evidence in support of the effectiveness of the political promotion inventive mechanism. Similarly, using a city mayor and CCP secretary dataset collected from internal sources, Zheng, Kahn, Sun et al. (2014) studied the promotion chances of leaders in 86 cities from 2004 to 2010 and found that although GDP growth maintained a dominant role in determining promotion rates, the declines in local air pollution levels and industrial energy intensities were statistically significant determinants of the probability of promotion. [19]

[18]In 2006, together with air pollution indicators, the Chinese central government also included water pollutant as a key promotion evaluation criterion for local officials (Kahn, Li and Zhao, 2015).

[19]The local official dataset includes information on the officials' name, age, gender, education, positions, etc. (Zheng, Kahn, Sun et al., 2014).

Li, Qiao et al. (2017) employed a quasi-experimental design (the air pollution regulations implemented during the 2014 APEC Meeting and 2015 Victory Day Parade in China), and used a time series regression discontinuity approach to detect the causal impact of these regulations. Using city-level daily AQI and 6 kinds of pollutant concentration data, they found these regulations improve air quality in Beijing (a reduction of AQI by 37.4% during the APEC Meeting and a decline of AQI by 35.9% during the Victory Day Parade). There was also an effect in several other Chinese cities although the effect is in relatively short term. Wu and Hu (2019) also use a regression discontinuity approach to investigate the impact of the CEIP inspection policy on provincial air quality levels across China and found that the inspection policy improved local air quality levels, but the effect was uneven across different pollutants and provinces. However, their focus on province air quality changes cannot reveal how different cities react to the CEIP inspection differently, and the limitation of RD design restricts the estimating window around the policy (they only use five months of data).

2.4 Data

2.4.1 Air quality data

The city level, daily air quality data that we use to construct our Air Quality Index (AQI) are from the Ministry of Environmental Protection of China and National Environmental Monitoring Centre (CNEMC). The six components of the AQI, PM2.5, PM10, $S0_2$, $N0_2$, O_3 and CO, are from "https://www.aqistudy.cn/", which calculates the emissions of each pollutant using a method provided by MEP.

The MEP adopted the US EPA AQI approach in 2012 to develop the Chinese AQI system. All the measurements stem from the national air quality monitoring sites located in each city. At each site, automated monitoring systems have been installed and used to measure the ambient concentrations of SO_2, NO_2, O_3, CO, PM2.5, and PM10 according to China Environmental Protection Standards HJ 193–2013 and HJ 655–2013. The values from the monitoring sites at each city were automatically reported to the China National Environmental Monitoring Centre, and published after being validated based on Technical Guideline on Environmental Monitoring Quality Management HJ 630–2011. The city wide average concentrations were calculated by averaging the concentrations at all sites in each city. This is the same method that the Chinese MEP uses to report daily concentrations of air pollutants to the public (see appendix 3 for details). The AQI approach converts the concentrations to a number on a scale of 0 to 500, with a sub-AQI of 100 corresponding to the upper limit of the CAAQS 24-hour Grade II standards.[20] The air quality level at a location on a given day is defined as unhealthy if the overall AQI level is greater than 100 based on daily-average concentrations (Hu et al., 2015). Table 2.1 presents a description of the AQI index and how it can impact health.

[Table 2.1 about here]

To answer research questions using the data described it is important that the data are accurate and reliable. For example, Chen, Jin et al. (2012) employed official API data from 37 Chinese cities between 2000 and 2009 and found significant discontinuity around the threshold of 100 (if the level of API in a city is equal or less than 100, the city is counted a "blue sky day" by the government and therefore rewarded the status of model city), suggesting the existence of data manipulation in some local Chinese cities. However, the accuracy and reliability of Chinese official air pollution data improved considerably since the adoption of the AQI system by the China National Environment Monitoring Centre in January 2013. To investigate the quality of air pollution data from Chinese surface monitoring stations,

[20]China National Ambient Air Quality Standards. GB3095-2012 (http://kjs.mep.gov.cn/hjbhbz/bzwb/jcffbz/201203/t20120302$_2$24166$_w$*ap.shtml*).

Wu, Tang et al. (2018) used hourly concentration data of 1,436 monitoring stations (for six pollutants including SO_2, NO_2, O_3, CO, PM2.5, and PM10) from the China National Environmental Monitoring Network (CNEMN) and employ four different types of outlier detection methods on the raw hourly concentration data and they found that the ratio of outliers is only 0% to 5.68% with the majority below 1.00% for most pollutants. These results confirm that there has been considerable improvements in air pollution data quality from the Chinese official monitoring stations between 2014 to 2016.

2.4.2 Meteorological Data

Our weather data are from the China Meteorological Administration (CMA). Since air pollution concentrations are also affected by meteorological conditions within each city (Rost et al., 2009; Jones, Harrison and Baker, 2010), to deal with the potential endogenous problem in the air pollution control evaluation literature, we control for city-level latitude, longitude, elevation, temperature, air pressure, precipitation, relative humidity, wind speed and wind direction (Chen, Jin et al., 2013; He, Fan and Zhou, 2016; Fu and Gu, 2017).

Since the following round of inspections started on 16th July 2016, we drop the air quality and weather data after this date, to make sure that our estimated effects of the Hebei inspection are not affected by later rounds of inspections in other regions. Our final data covers the period between December 2013 and the 15th July 2016 for Hebei and another 334 Chinese cities.

2.5 Methodology

To investigate the effect of the CEIP on air quality, we focus on the inspection of cities in Hebei province that took place in January 2016. We concentrate on Hebei for a number of reasons. First, within the CEIP process, Hebei was the only time that a single province was inspected. In other cases up to eight provinces were inspected during any one month period. Hebei has the other advantage of being the first province to be inspected. This helps us to allay possible endogeneity concerns. If we had chosen the next inspection round when eight provinces were inspected (from the 16th July 2016) it could be argued that it would be difficult to identify the true effect of the inspection as air pollution in any given city may be impacted by the pollution levels in neighbouring cities. In the case of these eight provinces, some of them are also neighbours of Hebei and hence their air pollution levels may have been impacted by the results from the Hebei inspection. For example, if the Hebei inspection had a dramatic effect on reducing pollution then this could also reduce pollution levels in neighbouring, but not inspected cities, in those eight provinces and will bias our results. The inspection of these cities may appear to be more successful than they really were. Similar arguments can be made for the other four rounds within the CEIP. In addition, since Hebei was the first province to be inspected, the local officials had no prior experience or clear idea of what to expect which means it would have been harder to take pre-emptive action ahead of the visit. Hence, our focus on Hebei provides us with the cleanest form of natural experimental setting with which to assess the effectiveness of the CEGP.

Our empirical strategy is to measure air pollution before, during and after the inspection period. First, we set a ‘pre-period’ defined as the 1st of July 2015 to the 31st of December 2015. The start date was chosen as it was the date of the announcement that in the future there was to be city-level inspections that would concentrate on environmental issues and hence this is the earliest date by which cities would have been able to react and begin to take remedial action to reduce air pollution. Although official reports state that local cities will not know in advance when and which cities will be targeted for inspection, we are still interested in investigating whether cities in Hebei did know in advance that they were to be inspected, and if they did know, how they responded. Second, we set the 'during period' to be from the 1st of January to the 31st January of 2016, which were the dates that the inspection team was physically based in Hebei province. Finally, we set an 'after period' to go from the 1st February to the 15 July 2016 which covers the period the inspection team left until the beginning of the next inspection round of a further eight provinces starting on 16th July 2016.

Using these three periods allows us to identify the inspection effects prior, during and after the inspection. After the inspection of Hebei, the Chinese MEP published an official report called the Central Government Environmental protection (CEIP) inspection feedback on the trial inspection in Hebei. Although there are eleven prefectural-level cities in Hebei province, eight out of those elven were criticized in the official feedback from MEP. This post-inspection categorization allows us to also investigate whether cities within Hebei responded differently to being inspected.

Our baseline identification strategy is to investigate the treatment effect of the Hebei inspection on local air quality by comparing cities in Hebei with cities outside of Hebei and is similar in spirit to Chen, Jin et al. (2013).

$$Y_{it} = \alpha_i + \delta_t + \sum_{z=0}^{n} \beta_{hebei,z} \times hebei\_period_z + \lambda W_{it} + f(t) + \epsilon_{it} \tag{2.1}$$

where,Y_{it} represents the log of AQI or one of the six 6 components (PM2.5, PM10, SO_2, NO_2, CO and O_3_8h) in city i and date t. α_i denotes city fixed effect, controlling for unobserved, time-unvarying city characteristics that affect air quality (Fu and Gu, 2017), δ_t represents a date fixed effect, controlling for all national wide policies that aim to reduce air pollution levels over the period (Chen, Jin et al., 2013).

Our main variable of interest, $\beta_{hebei,z}$, in its most basic form, represents the interaction of the Hebei dummy (cities in Hebei province equal to 1 and 0 otherwise) and three corresponding time periods, pre (01/07/2015 to 31/12/2015), during (01/01/2016 to 31/01/2016) and after (01/02/2016 to 15/07/2016). This allows us to detect how air pollution in inspected cities responds to the central governmental inspection prior, during and after the actual inspection. We further decompose the pre and post periods into shorter lengths of time to allow us to get a deeper understanding of the underlying dynamics. In further sensitivity checks we also distinguish between the criticised and non-criticised cities in Hebei.

To control for underlying meteorological conditions reported by CMA we include W_{it} that includes latitude, longitude, elevation, average temperature, average air pressure, average relative humidity, maximum wind speed and direction, total precipitation of city i and date t. $f(t)$ includes a city-specific linear trend and a city-specific quadratic trend, to control for unobserved confounding factors or trends that may affect daily air quality, e.g., the city-specific linear trend is the interaction of a city dummy and day count (between our start date (02/12/2013) and date t), while city-specific quadratic trend is the interaction of a city dummy and day count squared (Fu and Gu, 2017). Finally, ϵ_{it} is the error term and clustered at individual city-level.

2.6 Results

2.6.1 Descriptive statistics

Table 2.2 summarizes the daily average AQI and Tables 2.4 to 2.9 summarize the daily average of six pollutant concentrations by time periods and city groups separately. From

Table 2.2, during the benchmark period (before the inspection policy was announced), the daily average AQI in Hebei was 40 points higher than that of non-Hebei cities. As for the cities in Hebei, the AQI decreased slightly before and during the inspection with a much larger fall in the AQI after the inspection (around 29.49 point decline compared with the benchmark period). In contrast, non-Hebei cities experienced a slight increase in the AQI during the inspection period and no substantial improvement in the post inspection periods.

[Table 2.2 about here]

[Table 2.3 about here]

Based on the values in Table 2.2, Table 2.3 calculates the percentage change in the AQI in each period compared with the benchmark period. For example, before the inspection, for Hebei cities the percentage change of AQI compared with the benchmark period is calculated as (111.76-130.71)/130.71=-14%). Table 2.3 suggests that, compared with the benchmark period, cities in Hebei experienced larger reductions in the AQI in periods of pre3, pre2, during, after 1, and after 2. Tables 2.2 and 2.3 also show that air quality levels between criticized and un-criticized cities within Hebei differed substantially. Tables 2.2 and 2.3 illustrate that, compared with benchmark AQI levels in each city group, criticized Hebei cities experienced a larger AQI reduction in pre3, pre2, after 1, after 2 and after 3 periods than those of un-criticized Hebei cities, where un-criticized cities saw the largest AQI decline (26%) whereas the AQI increased by a small 1% for criticized Hebei cities during the month of the inspection. For the period after the inspection, compared with the benchmark group of cities, the AQI dropped substantially in the criticised cities perhaps in response to the critical report.

[Table 2.4 about here]

[Table 2.5 about here]

[Table 2.6 about here]

[Table 2.7 about here]

[Table 2.8 about here]

[Table 2.9 about here]

As for PM2.5 , Table 2.4 shows that the PM2.5 level in Hebei dropped sharply (a 33.87 point decline compared with the benchmark period) after the inspection whereas the drop of PM2.5 in non-Hebei cities was not as dramatic. Similar trends were also found in PM10, SO_2, CO, NO_2 and O_3_8h for the Hebei cities. Similar patterns can be also shown for PM10 and SO_2, shown in Tables 2.4 and 2.5, respectively. Observations from Table 2.4 to 2.9 suggest that air pollution of cities within Hebei province responded to the inspection policy differently and raises the question as to what is driving these differences. The large jump in pollution levels in the one month before the inspection is a reflection of the time of year of the inspection as the province moves into winter.

[Figure 2.2 about here]

To make comparisons across cities easier to interpret, in Figure 2.2 we plot the time series distribution of AQI (31-day moving average AQI level at specific date t) by different city groups and time periods. Compared with non-Hebei cities (black line), the AQI level was much higher in the Hebei cities (red line) throughout the time period and especially so in winter. Within Hebei province, the AQI level in criticized Hebei cities (blue line) was dramatically higher than that of un-criticized Hebei cities (yellow line), as well as the whole Hebei (red line). It is clear that the four city groups share similar trends and there is a very strong seasonal effect, with high AQI concentration levels in winter which is consistent with observations from (Chen, Jin et al., 2013). Since the Hebei inspection was set for January 2016 (which is the middle of winter), the extremely high level of AQI in January 2016 is something that would be expected but that our treatment effect could be underestimated if we use a simple before-after comparison in Hebei cities. To deal with this issue, we use different benchmark settings in a series of robustness checks. What Figure 2.2 tells us is that the fluctuations in AQI vary substantially by city groups, time periods and seasons. The implication is that a simple city and date fixed effect may not fully capture the trend and means that we may need to include city-specific trends if we want to estimate the causal effect of the inspection policy (Chen, Jin et al., 2012). Hence, to estimate the effect on the policy on air pollution we need to run our estimating equation.

2.6.2 Estimation results

Results of time setting 1

Table 2.10 reports the estimated effect of the CEIP on air quality levels for those cities in Hebei that were inspected, controlling for weather conditions, city and date fixed effects and city specific time trends (linear and quadratic). Columns (1) to (7) contrast Hebei cities with non-Hebei cities, compared to the benchmark period. Column (1) suggests that AQI levels in Hebei decreased by 19.3% (significant at 99% confidence level) in the pre period, whereas Columns (2) to (7) show there was a 23% decrease in PM2.5, a 23.8% decrease in PM10, a 26.6% decrease in SO_2 and a 15% decrease in O_3_8h, respectively. No decrease was observed for the level of CO and or NO_2 There are a number of reasons why the air quality in Hebei fell in the pre period, one of which is potentially an "information leak". Although it is a requirement of the central inspectors to secretly investigate firms without prior warning, the improvement in the AQI (and PM2.5, PM10, SO_2, O_3_8h pollution levels) before the month of the inspection suggests some sort of pollution control actions may have been taken by local officials and firms in advance of the upcoming (expected) inspection. This finding is consistent with Wang (2013) who argues that local officials may inform firms to shut down certain factories in advance of the inspection team to avoid being inspected and hence any resulting punishment.

[Table 2.10 about here]

Looking at the actual month of the inspection we find a 40.9% decline in the AQI, a 47.6% decline in PM2.5, 48.4% decline in PM10 and a 21% decline in O_3_8h. For SO_2 and NO_2 we find falls of 0.9% and 8.7% respectively which are not statistically significant. Likewise, there was no significant improvement in CO levels. If we look at the post-inspection period we find the largest reduction is the AQI (the AQI fell 47.3%). Other pollutants also experienced substantial improvements with PM2.5 and PM10 levels falling by 59.6% and 55.3% respectively, and SO_2 and CO levels falling by 31.1% and 31.7% respectively. Finally, NO_2 dropped by 12.3% but only at the 90% confidence level. Column (7) shows that O_3 increased by 30.5% in the post inspection period. The results of rising O_3 is consistent with the results of Vu et al. (2019), where they identified the level of Ozone increased slightly in Beijing between 2013 and 2017, while the levels of PM2.5, PM10, SO_2, NO_2 and CO were all decreasing. This seemingly surprising result can be explained by the chemistry of O_3 formation Wang, Li et al. (2015). The increase in Ozone is reassuring given the chemistry and gives us greater confidence in our other results. A back-of-the-envelope calculation shows that a 40.9% decrease in the AQI during the inspection is equivalent to a 51.3 point decline of AQI concentration levels (since we know from Table 2.2 that daily average AQI level in Hebei during the inspection was 125.38 points), and a 47.3% AQI decline in the post-inspection period is equivalent to a 47.9 point decrease in the AQI level (since daily average AQI level of Hebei in the post inspection period was 101.22 points). Based on the estimation from China's Huai River Policy, Ebenstein, Fan, Greenstone, He and Zhou (2017) demonstrated that a 10 g/m^3 increase of PM10 decreases life expectancy by 0.64 years. Similarly, He, Fan and Zhou (2016) found that a 10% decrease in PM10 during the 2008 Beijing Olympic Games resulted in an all-cause mortality rate decline of 8%, and more than 285,000 premature deaths could be avoided annually if there was a 10% decline in PM10 concentrations. Our finding of dramatic falls in the AQI, $PM_{2.5}$, PM10 and SO_2 indicate the potential health benefits from the introduction of the CEIP inspection policy. Although it is unclear which specific mechanism is driving the reduction in air pollution (ideally we would have data on the emissions from individual firms and the actions taken in light of the inspection), the substantial and sharp improvements in air quality for the different time periods suggests that the CEIP inspection programme was effective.

Results for time setting 2

To understand how air pollution changed within our time period we further decompose the post-inspection period into three sub-periods, post1 (2016.02.01 to 2016.03.31), post2 (2016.04.01 to 2016.05.31) and post3 (2016.06.01 to 2016.07.15). Table 2.11 illustrates the results using time setting 2 using the same specification except that we now include post1hebei, post2hebei and post3hebei. In contrast with non-Hebei cities and compared to the benchmark period, Column (1) of Table 2.11 demonstrates a slight improvement in the AQI in Hebei in the pre period (7.5%), a fairly large improvement during the inspection period (20.1%), and a larger decline in the post1 period (30.1%). The effect then fades away gradually from post1 to post3. The AQI results confirm that air quality improved considerably during 1-month inspection with the greatest improvements happening in the 2-months after the inspection team left Hebei. The concentration levels of PM2.5, PM10 and SO_2 experienced similar trends, with an improvement in the pre period, a fairly large effect during the month of inspection and larger declines after the inspection team left.

[Table 2.11 about here]

It is useful to compare our findings with Chen, Jin et al. (2013) who also found that radical environmental regulations can have an immediate influence on the air quality in

China, although this effect faded away gradually. It is worth noting the difference between two studies. Air quality levels improved in both studies. However, the mechanisms were different. For the 2008 Beijing Olympic Games, the highest air pollution decline occurred during the games but the effect had disappeared by 2010. However, in our case, the highest air quality improvement occurred after the inspection team left, although there was also a fairly large effect during the month of the inspection. One possible explanation is the different motives of the central government. The motivation for the Beijing 2008 Games was to ensure good air quality during the Olympics so local government officials and firms adopted temporary pollution control actions (He, Fan and Zhou, 2016). Due to the temporary nature of these closures, it is not surprising that air pollution increased again after the games has finished. However, for the CEIP, reducing air pollution during the inspection period was not the key task. Instead it was to build an effective and long-term air pollution improvement mechanism.

Results for time setting 3

Table 2.12 presents the estimation results of time setting 3, where we decompose both pre and post periods into 3 sub periods (pre3, 2015.07.01 to 2015.08.31; pre2, 2015.09.01 to 2015.10.31; pre1, 2015.11.01 to 2015.12.31; post1, 2016.02.01 to 2016.03.31; post2, 2016.04.01 to 2016.05.31; post3, 2016.06.01 to 2016.07.15). We use the same specifications as in Table 1.11 except we now include the variables pre3hebei, pre2hebei, pre1hebei, post1hebei, post2hebei, post3hebei.

[Table 2.12 about here]

According to Table 2.12, PM10, SO_2 and O_3_8h levels fell significantly during the inspection period, although we fail to find significant declines in AQI and PM2.5, suggesting the results for these pollutants are sensitive to the time settings. However, the general trends in AQI, PM2.5, PM10 and SO_2 were relatively similar across the 3 different settings, all showing a stable decrease during the one month inspection with the largest effects in the post-inspection periods.

Results of sub-sample analysis

After the Hebei inspection, the Chinese MEP published an official report called Central Government Environmental protection (CEIP) inspection feedback on the trial inspection in Hebei. Altogether there are eleven Prefectural-Level cities in Hebei, eight of the eleven were criticized in the official feedback. Therefore, instead of including all eleven cities as the treatment group, within this sub-sample analysis, we set the eight criticized cities and the three un-criticized cities separately, and then used the same specification to examine the different policy effects.

We first set the eight criticized cities in Hebei as the Hebei dummy and interact this dummy with our pre, during and post periods (dropping the three un-criticized cities from the dataset). We then do the same for the three un-criticized cities.

[Table 2.13 about here]

[Table 2.14 about here]

Tables 2.13 and 2.14 show the results for the two different sub-samples. The first observation is that we find that the inspection policy had a large and statistically negative impact on air quality levels across the different cities within Hebei and in the different time periods consistent with our main results from Table 2.2 Second, by comparing the coefficient on duringhebei in the first column between Tables 2.13 and 2.14, we observe in the pre-period a decline in the AQI in the eight criticised cities (18.2%) that was smaller than the decline in the three un-criticized cities (22.8%). We find a similar pattern during the inspection

(31.4% for the eight criticised cities and 66.3% for the three uncriticised cities). In contrast, following the inspection the decline of AQI in afterhebei period (53.1%) was much larger than that of three un-criticized cities (31.3%).

One possible explanation for these differences is that the three un-criticized cities cleaned up their pollution (or had lower initial air pollution levels) before and during the inspection and hence were not 'criticised'. However, the eight criticised cities did not respond immediately (or they had higher initial air pollution levels), were then criticised and hence had to clean up their pollution later. The results for PM2.5 in Column (2) and PM10 in Column (3) are also supportive of this argument. To summarise, this sub-sample analysis provides evidence to suggest that air pollutants of different cities within Hebei reacted to the inspection policy differently although the overall result remains a decline in air pollution.

2.6.3 Robustness checks

'Fake Hebei' placebo test

To ensure that our methodology is capturing the real impact of the inspection policy on air pollution across cities in Hebei, we conduct a series of placebo tests. Our first placebo test is to generate what we call a 'fake Hebei'. To do this we randomly choose a different eleven cities from all of the other provinces (none of which have been inspected at this point) and assume that during the same month of January 2016, that the Central Government Environmental Protection team were located in the cities that make up our 'fake Hebei'. We then interact these 'fake Hebei' dummies with same time period settings of pre, during and after. We use exactly same specifications for AQI as we use for Column (1) in Table 2.10 Since the initial one month inspection only targeted cities in Hebei, we expect there to be no impact of the 'pre (fake) Hebei, during (fake) Hebei and pre (fake) Hebei' on the AQI of the eleven cities that constitute 'fake Hebei' (Chen, Li and Lu, 2018). We generate our 'fake Hebei' 1,000 times and conduct the placebo test 1,000 times which gives us 1,000 coefficients on the 'pre (fake) Hebei, during (fake) Hebei and post (fake) Hebei' dummies.

[Figure 2.3 about here]

[Figure 2.4 about here]

[Figure 2.5 about here]

Figures 2.3, 2.4 and 2.5 present plots of the distribution of the t statistics for the 1,000 coefficients for '(fake) prehebei', '(fake) during hebei' and '(fake) afterhebei', respectively. We can see that t statistics of these 'fake Hebei' placebo tests are narrowly located around zero (the mean of t statistics are 0.131, 0.160 and 0.154 from Figures 2.3, 2.4 and 2.5, respectively), and nearly all lie between -1.96 and +1.96. This compares to the t statistics from our main result for real prehebei, during Hebei and afterhebei of -6.88, -4.87 and -8.58, respectively, which all lie in the left tale of Figures 2.3, 2.4 and 2.5, respectively. These results give us considerable confidence that our findings are capturing an inspection effect and hence these changes in pollution are not by chance (Fu and Gu, 2017).

Alternative benchmark settings

From Figure 2.2 it is clear that there are strong seasonality effects driving pollution levels. The AQI for example tends to be high in the winter and much lower in the summer. This raises the concern that the true policy effect may be affected by the baseline benchmark period setting, e.g., if the average air pollution concentrations are relatively low during the benchmark period, our true policy effect would be underestimated. Hence, we conduct two alternative benchmark settings as a further robustness check. First, instead of setting the benchmark period from December 2013 to 30 June 2015, we drop the data between

2015.01.01 to 2015.06.30, and set the benchmark as December 2013 to December 2015, other things remaining the same. Since we have a shorter benchmark period and that period had relatively higher air pollution levels, we expect the policy effect to be larger than the effect found in the main results. Table 2.15 presents the results of the shorter benchmark setting. Compared with the results of Table 2.2, the policy effect (e.g. the coefficients of prehebi, duringhebei and afterhebei) remain statistically significant and are indeed larger for AQI, PM2.5, PM10, CO and NO_2. If anything, this is further support for the effectiveness of the inspection policy.

[Table 2.15 about here]

Second, since Hebei was inspected during January 2016, instead of setting the benchmark period from December 2013 to 30 June 2015, we only set it as January 2014 and January 2015 as the benchmarks, in an attempt to rule out strong seasonality effects of different pollutant concentrations for during the benchmark period. Similarly, other things remaining the same, we expect the policy effect to be larger under this benchmark of January only settings. Table 2.16 reports the results and compared to the results in Table 2.2, the coefficients on prehebei, duringhebei and afterhebei are all statistically significant and again larger for AQI, $PM_{2.5}$, PM10, SO_2 CO and NO_2.

[Table 2.16 about here]

To summarise, our results suggest that the inspection policy has a statistically significant and negative impact on air pollution in Hebei province. The larger coefficients when we change the benchmark period suggests that our main results from Table 2.10 could be a lower bound.

Exclusion of bordering cities

Another concern one might have is the possibility of spatial spillovers from an inspection policy to those cities that border Hebei province (Hering and Poncet, 2014 and Chen, Li and Lu, 2018). One could argue that air pollution levels in these neighbouring cities could be affected by what happens in the Hebei cities once Hebei was identified as the first province to be inspected. For example, bordering cities may respond by also taking action to reduce pollution to reduce the chances of future punishment or from increased awareness of environmental issues. A second mechanism by which emissions might be influenced by the inspection in Hebie is through the impact on transboundary pollution so that a reduction in pollution in Hebei cities may reduce pollution in neighbouring cities. In contrast, one could also imagine a scenario where factories shut down in Hebei relocate to neighbouring cities and hence increasing emissions in these cities. To address these concerns we exclude those non-Hebei cities that border Hebei. Table 2.17 presents the results. The results in Columns (1) to (7) show that the coefficients for prehebi, duringhebei and afterhebei remain significant and negative and are similar to the results in Table 2.2.

[Table 2.17 about here]

2.7 Conclusions

After years of rapid economic development, China is increasingly turning its attention to environmental protection. Between 2016 and 2017, the Chinese central government implemented a Central Government Environmental Protection Inspection policy, dispatching inspection teams to different cities with the goal of improving the local air quality. To understand whether or not the policy was effective in reducing local air pollution we investigate the impact of the first inspection to take place as part of the CEIP that occurred in January

2016 in Hebei province.

Our main finding is that the inspection process had a significant impact on a range of different pollutants. Compared to cities outside of Hebei, and then compared to a benchmark period before the policy was announced, the CEIP inspection policy is estimated to have reduced AQI levels by 40.9%, PM2.5 levels by 47.6% and PM10 levels by 48.4% during the one month inspection period. Our results show that the policy remained effective up to 5 months after the inspection team left Hebei, with a further 47.3% fall in the AQI level, a 59.6% drop in PM2.5, a 55.3% drop in PM10, a 31.1% drop in SO_2, and a 31.7% drop of CO. Decomposing the before and after period into smaller time periods confirmed the general decline in pollution and showed that the largest falls tended to occur in the period immediately after the inspection team has left.

When we split Hebei's eleven cities into those that were criticized in the final report and those that were not, we find that during the inspection period, air quality improved more in the uncriticised cities, whereas after the inspection left, it was the criticised cities that experienced the highest air quality improvements. This is consistent with the eight criticized cities trying to catch up and learning from the actions of those cities that were not criticized.

A series of placebo tests and other robustness checks suggest that the air quality improvement in Hebei is not a chance finding. Our results are also robust to using different initial period benchmarks and to the exclusion of bordering cities. Our findings are consistent with those of Li et al. (2017) and Wu and Hu (2019) that temporary air pollution regulation in China can reduce air pollution substantially but tends to be short term. This study contributes to the literature on the impact of inspection policies on reducing air pollution and shows that despite widespread scepticism inspection policies can be an effective tool in incentivising local officials to improve their environmental performance.

Although we find that air pollution declined significantly as a result of the Hebei inspection, it is still important to identify and understand the channels by which pollution is reduced. For example, air pollution reductions might be a result of the shutting down of factories, cutting off electricity, employing and installing pollution-reduction technologies, or even moving factories outside the city boundaries. These actions can have very different impacts on social welfare. It is preferable that local officials and firms take positive action such as upgrading technologies, importing better pollution control equipment or adopting market mechanisms (e.g., cap-and-trade system) to deal with pollution issues rather than shutting down factories. An analysis of what has driving the reduction in pollution is a topic for future research.

Chapter 2 Figures and Tables

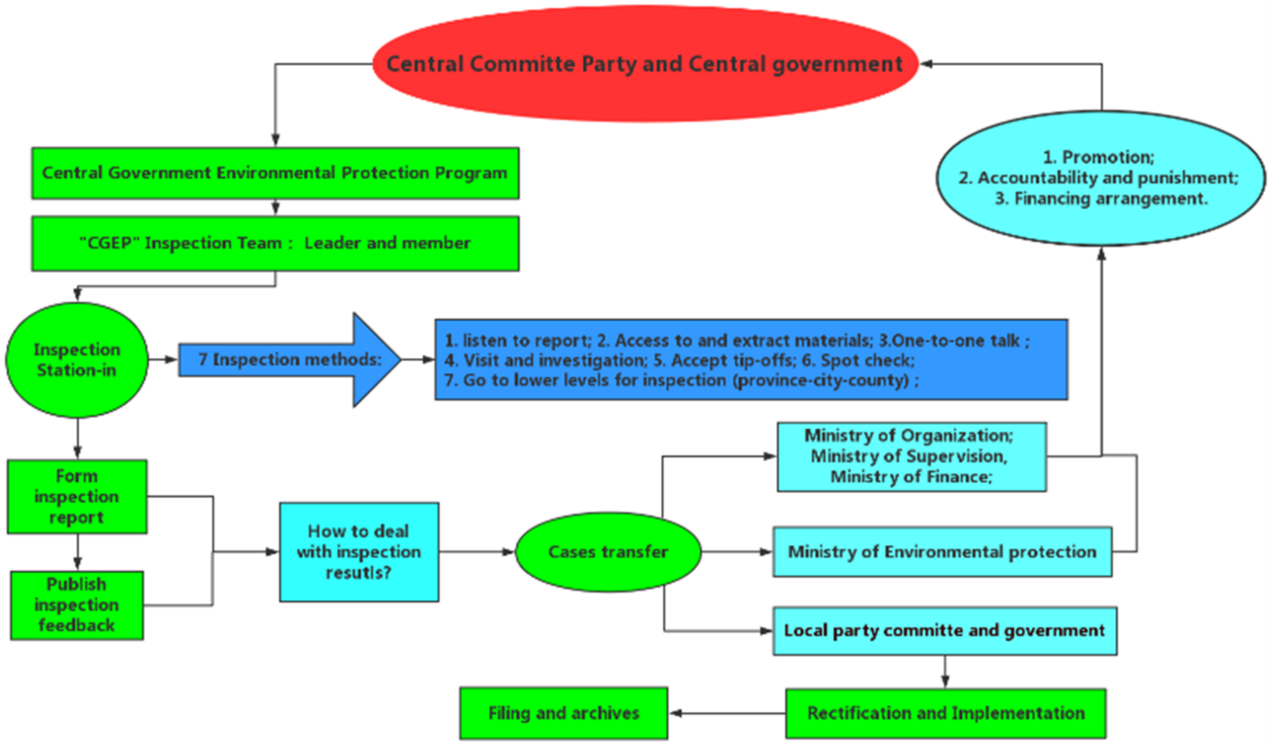


Figure 2.1: “CEIP” inspection process

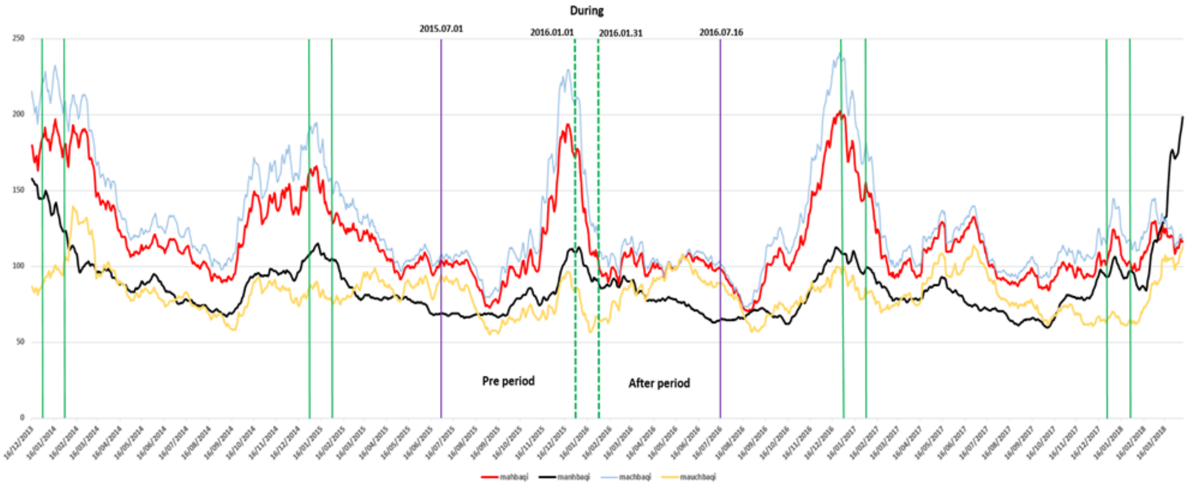


Figure 2.2: Time series of AQI by city group and time periods.

Note: 31-day moving average of a specific date t, mahbaqi: 31-day moving average level of AQI at a specific date in Hebei cities; manhbaqi:31-day moving average level of AQI at specific date in non-Hebei cities;machbaqi: 31-day moving average level of AQI at specific date in criticized Hebei cities; mauchbaqi: 31-day moving average level of AQI at specific date in non-criticized Hebei cities.

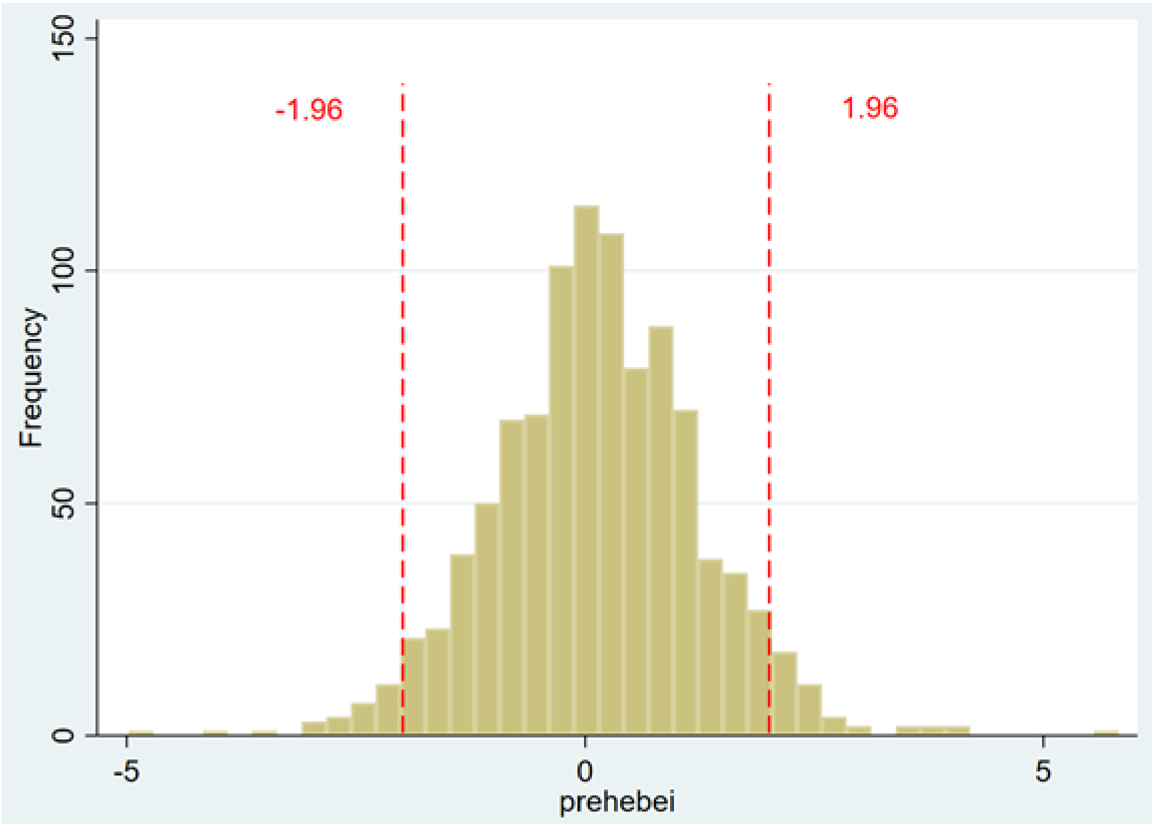


Figure 2.3: T statistics distribution of 1,000 times '(fake) prehebei'

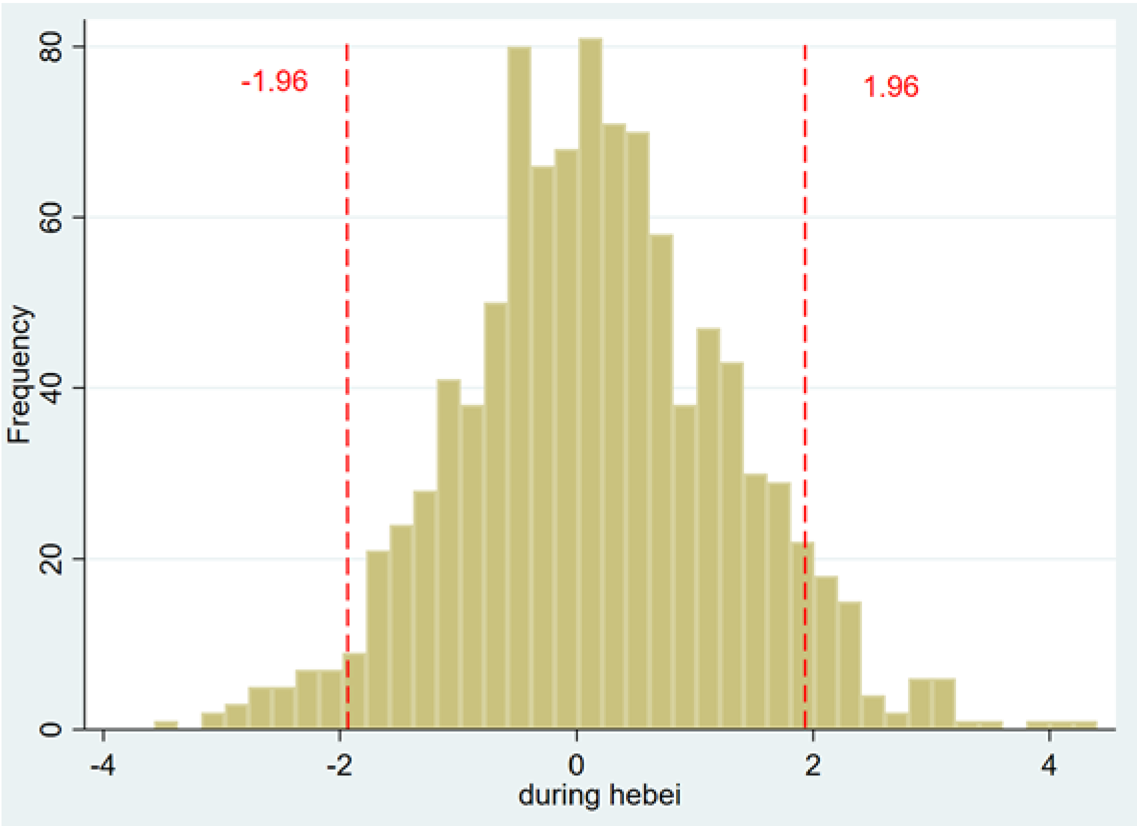


Figure 2.4: T statistics distribution of 1,000 times '(fake) during hebei'.

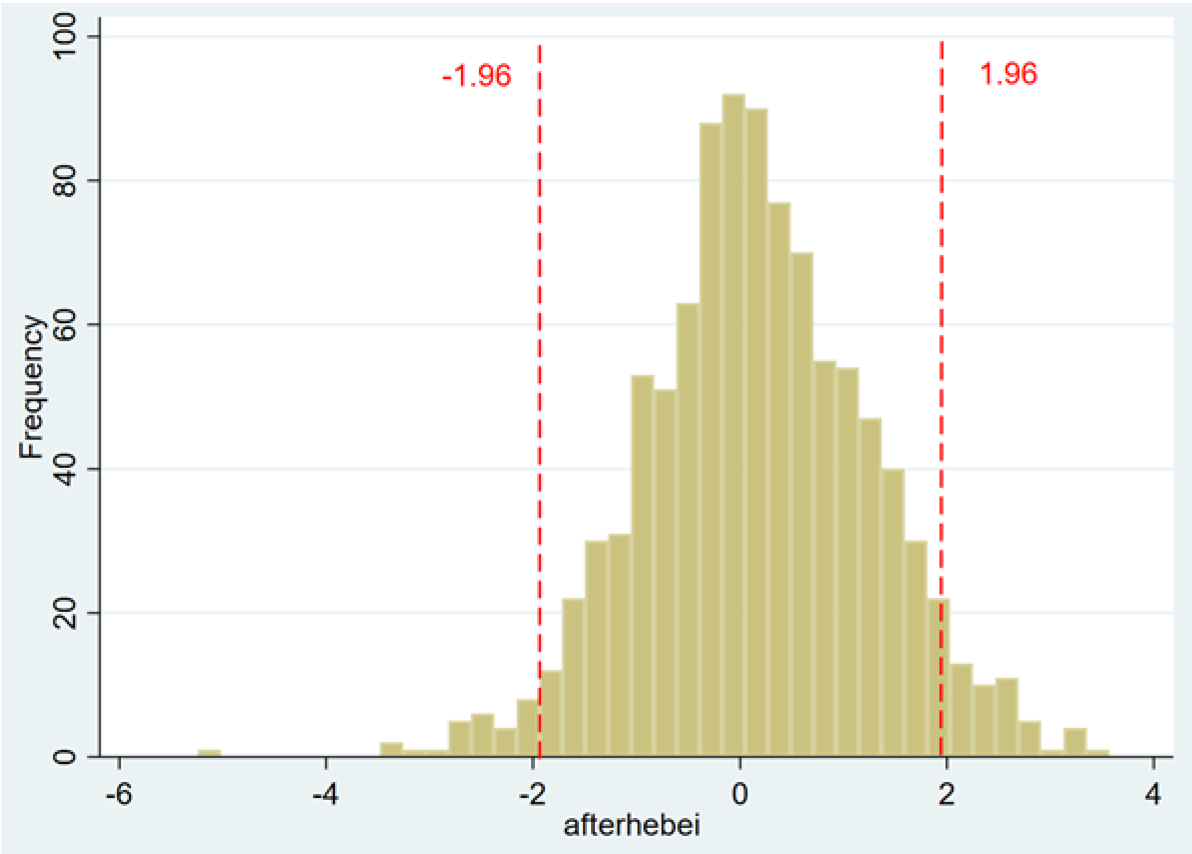


Figure 2.5: T statistics distribution of 1,000 times '(fake) afterhebei'.

Table 2.1: The health categories and meanings for AQI

AQI below	SO_2	NO_2	CO	O_3_8h	$PM_{2.5}$	PM10	Health risk category	Meaning
50	50	40	2	100	35	150	good	Satisfactory AQ, no risk
100	150	80	4	160	75	150	Moderate	Acceptable AQ, may be a moderate risk for a very small number of people who are usually sensitive to certain pollutants
150	475	180	14	215	115	250	Unhealthy for sensitive groups	Children, older adults, and people with lung and heart disease are at a greater risk
200	800	280	24	265	150	350	Unhealthy	Everyone begins to experience some adverse health effects, and sensitive groups experience more serious effects
300	1600	565	36	800	250	420	Very unhealthy	Everyone experience more serious health effects
400	2100	750	48	1000	350	500	Hazardous	Entire population is affected, stay indoors and avoid outdoor activities
500	2620	940	60	1200	500	600	Severe	Entire population is affected, stay indoors and avoid outdoor activities

Note: The upper limit AQI values and corresponding pollutants concentrations based on the Technical Regulation on Ambient Air Quality Index. Source: MEP, 2012 (Hu et al., 2015)

Table 2.2: Summary of daily average AQI by periods and city groups

AQI	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	130.71	147.46	86.04	90.65
Pre (2015.07.01 to 2015.12.31)	111.76	125.63	74.78	77.63
pre 3 (2015.07.01 to 2015.08.31)	96.86	102.62	81.49	68.91
pre 2 (2015.09.01 to 2015.10.31)	86.46	95.74	61.71	72.94
pre 1 (2015.11.01 to 2015.12.31)	152.22	178.92	81.02	91.19
During (2016.01.01 to 2016.01.31)	125.38	148.39	64.02	96.60
After (2016.02.01 to 2016.07.15)	101.22	106.14	88.10	78.85
after 1 (2016.02.01 to 2016.03.31)	101.81	110.82	77.77	89.17
after 2 (2016.04.01 to 2016.05.31)	99.12	101.02	94.05	77.38
after 3 (2016.06.01 to 2016.07.15)	103.28	106.84	93.80	67.08

Table 2.3: Summary of percentage change of daily average AQI by periods and city groups

AQI	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	130.71	147.46	86.04	90.65
Pre (2015.07.01 to 2015.12.31)	-14%	-15%	-13%	-14%
pre 3 (2015.07.01 to 2015.08.31)	-26%	-30%	-5%	-24%
pre 2 (2015.09.01 to 2015.10.31)	-34%	-35%	-28%	-24%
pre 1 (2015.11.01 to 2015.12.31)	16%	21%	-6%	1%
During (2016.01.01 to 2016.01.31)	-4%	1%	-26%	1%
After (2016.02.01 to 2016.07.15)	-23%	-28%	2%	13%
after 1 (2016.02.01 to 2016.03.31)	-22%	-25%	-10%	-2%
after 2 (2016.04.01 to 2016.05.31)	-24%	-31%	9%	-15%
after 3 (2016.06.01 to 2016.07.15)	-21%	-28%	9%	-26%

Table 2.4: Summary of daily average $PM_{2.5}$ by periods and city groups

$PM_{2.5}$	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	90.18	106.10	47.71	58.57
Pre (2015.07.01 to 2015.12.31)	75.47	89.87	37.06	46.92
pre 3 (2015.07.01 to 2015.08.31)	57.28	66.55	32.56	34.51
pre 2 (2015.09.01 to 2015.10.31)	51.59	60.81	27.02	42.34
pre 1 (2015.11.01 to 2015.12.31)	117.83	142.63	51.68	64.11
During (2016.01.01 to 2016.01.31)	92.35	112.92	37.51	68.38
After (2016.02.01 to 2016.07.15)	56.31	63.57	36.94	43.76
after 1 (2016.02.01 to 2016.03.31)	67.00	75.71	43.78	58.12
after 2 (2016.04.01 to 2016.05.31)	50.43	57.11	32.61	40.36
after 3 (2016.06.01 to 2016.07.15)	50.02	56.14	33.68	29.24

Table 2.5: Summary of daily average PM10 by periods and city groups

PM10	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	160.20	182.65	100.35	98.77
Pre (2015.07.01 to 2015.12.31)	126.23	144.87	76.55	78.98
pre 3 (2015.07.01 to 2015.08.31)	100.28	112.43	67.89	63.60
pre 2 (2015.09.01 to 2015.10.31)	103.00	116.26	67.63	75.79
pre 1 (2015.11.01 to 2015.12.31)	175.84	206.43	94.27	97.80
During (2016.01.01 to 2016.01.31)	141.03	166.86	72.14	104.06
After (2016.02.01 to 2016.07.15)	110.49	119.45	86.59	83.95
after 1 (2016.02.01 to 2016.03.31)	123.96	135.88	92.18	106.18
after 2 (2016.04.01 to 2016.05.31)	114.28	121.80	94.24	83.25
after 3 (2016.06.01 to 2016.07.15)	87.37	94.35	68.76	55.26

Table 2.6: Summary of daily average SO_2 by periods and city groups

SO_2	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	55.48	59.05	45.97	31.41
Pre (2015.07.01 to 2015.12.31)	32.98	36.62	23.27	22.46
pre 3 (2015.07.01 to 2015.08.31)	20.00	21.34	16.41	15.59
pre 2 (2015.09.01 to 2015.10.31)	25.24	28.41	16.78	19.69
pre 1 (2015.11.01 to 2015.12.31)	53.92	60.37	36.73	32.23
During (2016.01.01 to 2016.01.31)	62.82	73.17	35.23	35.47
After (2016.02.01 to 2016.07.15)	32.36	36.55	21.20	20.38
after 1 (2016.02.01 to 2016.03.31)	46.49	52.49	30.51	27.98
after 2 (2016.04.01 to 2016.05.31)	27.19	31.20	16.51	17.55
after 3 (2016.06.01 to 2016.07.15)	20.53	22.54	15.17	14.09

Table 2.7: Summary of daily average CO by periods and city groups

CO	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	1.56	1.69	1.19	1.18
Pre (2015.07.01 to 2015.12.31)	1.50	1.66	1.06	1.03
pre 3 (2015.07.01 to 2015.08.31)	0.96	1.04	0.76	0.82
pre 2 (2015.09.01 to 2015.10.31)	1.12	1.23	0.85	0.93
pre 1 (2015.11.01 to 2015.12.31)	2.41	2.73	1.57	1.34
During (2016.01.01 to 2016.01.31)	2.22	2.60	1.22	1.40
After (2016.02.01 to 2016.07.15)	1.16	1.28	0.84	0.97
after 1 (2016.02.01 to 2016.03.31)	1.52	1.69	1.07	1.15
after 2 (2016.04.01 to 2016.05.31)	0.96	1.05	0.73	0.90
after 3 (2016.06.01 to 2016.07.15)	0.96	1.07	0.67	0.81

Table 2.8: Summary of daily average NO_2 by periods and city groups

NO_2	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	47.16	51.33	36.02	33.43
Pre (2015.07.01 to 2015.12.31)	47.92	52.15	36.64	29.65
pre 3 (2015.07.01 to 2015.08.31)	32.52	34.28	27.82	21.90
pre 2 (2015.09.01 to 2015.10.31)	46.83	51.72	33.78	29.44
pre 1 (2015.11.01 to 2015.12.31)	64.66	70.73	48.46	37.74
During (2016.01.01 to 2016.01.31)	61.77	71.09	36.89	38.39
After (2016.02.01 to 2016.07.15)	43.63	47.38	33.64	27.79
after 1 (2016.02.01 to 2016.03.31)	51.10	55.90	38.31	32.71
after 2 (2016.04.01 to 2016.05.31)	42.31	45.82	32.96	27.43
after 3 (2016.06.01 to 2016.07.15)	35.46	38.14	28.33	21.73

Table 2.9: Summary of daily average O_3_8h by periods and city groups

O_3_8h	Hebei Cities	Criticized Hebei Cities	Un-criticized Hebei Cities	Non-Hebei Cities
Benchmark period (2013.12.02 to 2015.06.30)	89.26	90.71	85.41	85.52
Pre (2015.07.01 to 2015.12.31)	83.90	84.65	81.91	83.23
pre 3 (2015.07.01 to 2015.08.31)	127.04	128.25	123.79	105.01
pre 2 (2015.09.01 to 2015.10.31)	89.11	93.34	77.81	93.59
pre 1 (2015.11.01 to 2015.12.31)	34.86	31.64	43.45	50.72
During (2016.01.01 to 2016.01.31)	43.27	38.49	56.01	52.69
After (2016.02.01 to 2016.07.15)	119.45	118.75	121.33	97.88
after 1 (2016.02.01 to 2016.03.31)	80.75	80.21	82.18	82.14
after 2 (2016.04.01 to 2016.05.31)	133.60	131.39	139.51	106.91
after 3 (2016.06.01 to 2016.07.15)	151.88	153.01	148.87	106.62

Table 2.10: Main results of inspection on air quality in Hebei cities (time setting 1)

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$\ln SO_2$	lnCO	$\ln NO_2$	$\ln O_3$_8h
prehebei	-0.193***	-0.230***	-0.238***	-0.266***	-0.061	-0.009	-0.150***
	(0.028)	(0.066)	(0.046)	(0.063)	(0.049)	(0.047)	(0.041)
duringhebei	-0.409***	-0.476***	-0.484***	-0.009	0.032	-0.087	-0.210***
	(0.084)	(0.110)	(0.083)	(0.099)	(0.094)	(0.062)	(0.046)
afterhebei	-0.473***	-0.596***	-0.553***	-0.311***	-0.317***	-0.123*	0.305**
	(0.055)	(0.090)	(0.049)	(0.068)	(0.081)	(0.071)	(0.131)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	260,448	260,448	260,448	260,448	260,448	260,448	260,448
R-squared	0.289	0.376	0.339	0.403	0.361	0.410	0.455
No. of citynumber	334	334	334	334	334	334	334

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.11: Main results of inspection on air quality in Hebei cities (time setting 2)

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.075**	-0.069	-0.142***	-0.465***	-0.190***	-0.017	0.150**
	(0.034)	(0.061)	(0.045)	(0.091)	(0.065)	(0.066)	(0.075)
duringhebei	-0.201***	-0.195**	-0.314***	-0.358***	-0.192***	-0.100	0.316***
	(0.060)	(0.09616)	(0.08384)	(0.084)	(0.073)	(0.064)	(0.090)
post1hebei	-0.301***	-0.345***	-0.433***	-0.569***	-0.445***	-0.121	0.705***
	(0.068)	(0.083)	(0.065)	(0.112)	(0.100)	(0.101	(0.185)
post2hebei	-0.159	-0.208**	-0.256**	-0.899***	-0.771***	-0.171	1.170***
	(0.116)	(0.091)	(0.100)	(0.194)	(0.14901)	(0.145)	(0.239)
post3hebei	0.033	0.115	-0.169	-1.118***	-0.784***	-0.137	1.537***
	(0.133)	(0.086)	(0.125)	(0.266)	(0.192)	(0.182)	(0.298)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	260,448	260,448	260,448	260,448	260,448	260,448	260,448
R-squared	0.290	0.377	0.339	0.403	0.362	0.410	0.457
No. of citynumber	334	334	334	334	334	334	334

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.12: Main results of inspection on air quality in Hebei cities (time setting 3)

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
pre3hebei	0.021	0.090*	-0.051	-0.383***	-0.120*	-0.019	0.127**
	(0.034)	(0.052)	(0.048)	(0.084)	(0.067)	(0.071)	(0.063)
pre2hebei	-0.204***	-0.331***	-0.306***	-0.515***	-0.126***	0.027	-0.182**
	(0.048)	(0.084)	(0.063)	(0.076)	(0.047)	(0.067)	(0.087)
pre1hebei	0.106	0.147	-0.042	-0.207***	0.221**	0.056	-0.534***
	(0.100)	(0.114)	(0.109)	(0.068)	(0.088)	(0.080)	(0.105)
duringhebei	-0.116	-0.127	-0.297**	-0.193***	0.119	-0.034	-0.277**
	(0.127)	(0.145)	(0.145)	(0.070)	(0.108)	(0.075)	(0.107)
post1hebei	-0.197*	-0.263**	-0.412***	-0.369***	-0.067	-0.040	-0.015
	(0.101)	(0.131)	(0.136)	(0.103)	(0.107)	(0.101)	(0.161)
post2hebei	-0.028	-0.104	-0.230	-0.646***	-0.295**	-0.069	0.266
	(0.124)	(0.156)	(0.184)	(0.155)	(0.134)	(0.135)	(0.207)
post3hebei	0.190	0.239	-0.137	-0.815***	-0.215	-0.016	0.454*
	(0.158)	(0.180)	(0.232)	(0.230)	(0.163)	(0.166)	(0.254)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	260,448	260,448	260,448	260,448	260,448	260,448	260,448
R-squared	0.290	0.378	0.339	0.403	0.362	0.410	0.458
No. of citynumber	334	334	334	334	334	334	334

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.13: The results of only criticised cities in Hebei

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.182***	-0.206***	-0.217***	-0.249***	-0.044	-0.047	-0.113**
	(0.034)	(0.066)	(0.056)	(0.062)	(0.059)	(0.052)	(0.044)
duringhebei	-0.314***	-0.398***	-0.393***	0.083	0.120	-0.042	-0.215***
	(0.090)	(0.134)	(0.091)	(0.084)	(0.102)	(0.063)	(0.053)
afterhebei	-0.531***	-0.646***	-0.558***	-0.319***	-0.354***	-0.194***	0.492***
	(0.050)	(0.110)	(0.049)	(0.059)	(0.105)	(0.060)	(0.111)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	257,580	257,580	257,580	257,580	257,580	257,580	257,580
R-squared	0.292	0.378	0.340	0.401	0.363	0.412	0.454
No. of citynumber	331	331	331	331	331	331	331

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.14: The results of only un-criticised cities in Hebei

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.228***	-0.299*	-0.304***	-0.312**	-0.118	0.090	-0.249***
	(0.034)	(0.159)	(0.059)	(0.152)	(0.075)	(0.073)	(0.062)
duringhebei	-0.663***	-0.687***	-0.736***	-0.259	-0.215*	-0.207*	-0.191**
	(0.068)	(0.110)	(0.055)	(0.219)	(0.119)	(0.121)	(0.077)
afterhebei	-0.313***	-0.463***	-0.542***	-0.299*	-0.229***	0.068	-0.183
	(0.095)	(0.104)	(0.114)	(0.178)	(0.068)	(0.156)	(0.164)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	252,800	252,800	252,800	252,800	252,800	252,800	252,800
R-squared	0.286	0.376	0.337	0.398	0.360	0.408	0.447
No. of citynumber	326	326	326	326	326	326	326

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.15: Results of inspection on air quality in Hebei using a 1 year benchmark

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.387***	-0.440***	-0.335***	-0.027	-0.264**	-0.071	0.214***
	(0.081)	(0.112)	(0.110)	(0.110)	(0.102)	(0.058)	(0.070)
duringhebei	-0.650***	-0.746***	-0.620***	0.240	-0.246*	-0.172**	0.330***
	(0.098)	(0.131)	(0.121)	(0.153)	(0.128)	(0.080)	(0.080)
afterhebei	-0.764***	-0.906***	-0.705***	-0.029	-0.685***	-0.229**	0.958***
	(0.143)	(0.159)	(0.137)	(0.135)	(0.159)	(0.090)	(0.178)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	200,019	200,019	200,019	200,019	200,019	200,019	200,019
R-squared	0.304	0.391	0.357	0.442	0.377	0.429	0.483
No. of citynumber	334	334	334	334	334	334	334

Robust standard errors in parentheses. All pollutants are logged. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.16: Results of inspection on air quality in Hebei using a January benchmark

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.218**	-0.236**	-0.230***	-0.549***	-0.480***	-0.050	0.505***
	(0.086)	(0.093)	(0.078)	(0.114)	(0.125)	(0.079)	(0.113)
duringhebei	-0.497***	-0.553***	-0.516***	-0.255**	-0.419***	-0.144**	0.479***
	(0.064)	(0.095)	(0.071)	(0.127)	(0.113)	(0.071)	(0.095)
afterhebei	-0.635***	-0.754***	-0.629***	-0.494***	-0.798***	-0.195*	1.014***
	(0.133)	(0.136)	(0.105)	(0.132)	(0.168)	(0.100)	(0.199)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	143,436	143,436	143,436	143,436	143,436	143,436	143,436
R-squared	0.319	0.416	0.371	0.428	0.407	0.436	0.496
No. of citynumber	334	334	334	334	334	334	334

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Table 2.17: Results of inspection on air quality in Hebei after dropping bordering cities

VARIABLES	lnAQI	$PM_{2.5}$	PM10	$lnSO_2$	lnCO	$lnNO_2$	lnO_3_8h
prehebei	-0.204***	-0.237***	-0.249***	-0.275***	-0.062	-0.014	-0.156***
	(0.028)	(0.066)	(0.046)	(0.063)	(0.049)	(0.047)	(0.041)
duringhebei	-0.428***	-0.499***	-0.504***	-0.002	0.035	-0.097	-0.217***
	(0.084)	(0.111)	(0.083)	(0.100)	(0.094)	(0.063)	(0.046)
afterhebei	-0.492***	-0.620***	-0.570***	-0.328***	-0.321***	-0.134*	0.314**
	(0.055)	(0.091)	(0.050)	(0.069)	(0.081)	(0.071)	(0.131)
Weather	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y
CS linear trend	Y	Y	Y	Y	Y	Y	Y
CS quadratic trend	Y	Y	Y	Y	Y	Y	Y
Observations	246,755	246,755	246,755	246,755	246,755	246,755	246,755
R-squared	0.291	0.381	0.342	0.396	0.361	0.410	0.445
No. of citynumber	318	318	318	318	318	318	318

Robust standard errors in parentheses. All pollutants are logged. *** p<0.01, ** p<0.05, * p<0.1

Chapter Three

The Impact of the Wuhan Covid-19 Lockdown on Air Pollution and Health: A Machine Learning and Augmented Synthetic Control Approach

3.1 Introduction

At the time of writing, much of the world remains in the grip of the Covid-19 pandemic. The full social and economic consequences of the pandemic and its restrictions on our day-to-day activities will be far-reaching and will take time to fully identify and quantify. The primary method of slowing the infection rate has been to impose strict social distancing regulations known as 'lockdowns' where people are restricted to their own homes and all but essential economic activity ceases. Interestingly, one consequence of these lockdowns that became apparent from an early stage was a perceived improvement in air quality. As many shops and businesses closed, industrial activity and vehicle use in cities fell dramatically and reports emerged of pollution levels being considerably below those experienced in normal conditions. These reports first emerged in China but have since appeared in many other countries.[1] Such improvements in air quality and the likely associated health benefits have raised the prospect of an unlikely silver lining to the otherwise overwhelmingly negative impacts of the pandemic.

These apparent improvements in air quality have, however, raised a number of questions. First, which pollutants have actually fallen? Most media reports refer only to a reduction in Nitrogen Dioxide (NO_2) with no discussion of changes in other pollutants. Second, what benchmark is being used to measure any reduction? If pollution levels are changing year-on-year, comparisons with previous years may be misleading. Similarly, concentrations of many pollutants are significantly influenced by local weather conditions again making it difficult to compare emission levels with previous years or contemporaneously with other cities. Finally, what are the likely health benefits of any reductions in pollution? Using the example of the first Covid-19 lockdown, that happened in Wuhan, China, this paper addresses each of these questions in turn.

There are a number of reasons why it is important to understand how the response to the Covid-19 pandemic has affected pollution and pollution-related mortality. This pandemic and society's response to it has been unprecedented in modern times and the social and economic impacts will be diverse and long-lasting. In order to hasten the recovery and to learn lessons for future pandemics it is vital that we understand every aspect of the economic, social and health impacts of Covid-19 and the policies used to tackle it. More specifically,

[1]New York Times (2020) Watch the footprint of coronavirus spread across countries, March 17th. https://www.nytimes.com/interactive/2020/climate/coronavirus-pollution.html; Guardian (2020) Coronavirus pandemic leading to huge drop in air pollution, March 23rd. https://www.theguardian.com/environment/2020/mar/23/coronavirus-pandemic-leading-to-huge-drop-in-air-pollution.

it is important to understand how pollution (and health) responds to changes in social and economic activity. A city-level lockdown is an extreme case but, non-linearities aside, the pollution response to it informs us how different pollutants may respond to milder forms of restrictions on human activities such as congestion charging, pedestrianised zones and urban planning more generally. Furthermore, calculating the pollution and health benefits of the lockdown provides a unique opportunity to identify the costs incurred by society in going about its day-to-day business during “normal" times.

It is also important to understand how improved air quality as a result of the lockdown has lessened the strain on health services within cities such as Wuhan by reducing morbidity and mortality. Air pollution in China regularly exceeds World Health Organisation (WHO) guidelines and, in the absence of these pollution reductions, hospital admissions would almost certainly have been even higher given the clear links previously found between pollution, hospital admissions and mortality (eg. Maddison, 2005; Shang et al., 2013; Lagravinese et al., 2014; Cheung, He and Pan, 2020; Deryugina et al., 2019).[2] Relatedly, there have been reports that exposure to pollution may increase Covid-19 mortality raising the possibility that Covid-19 death rates in cities such as Wuhan may have been even higher if the lockdown had not improved air quality (see for example Wu, Nethery et al., 2020). Summing these latter two points, the cleaner air resulting from the lockdown may have increased the ability of hospitals to accommodate Covid-19 patients and directly reduced the number of Covid-19 deaths.

While it would therefore be useful if we could identify the impact of the Wuhan lockdown on pollution and health, isolating the effect of a policy intervention on pollutant concentrations is challenging since such concentrations are jointly influenced by meteoro-

[2] WHO health guidelines are that annual concentrations of pollution should be below 40 μg/m^3 for nitrogen dioxide and 20 μg/m^3 for coarse particulate matter (PM10), while the 24 hour mean of sulphur dioxide should be below 20 μg/m^3. The WHO does not provide guidelines for ambient concentrations of carbon monoxide.

logical conditions and emission levels. The influence of wind speed, wind direction, and temperature on pollution concentrations will often be greater than the effect of any policy intervention thereby significantly complicating attempts to isolate policy effects (Anh, Duc and Azzi, 1997). Traditional attempts to address the impacts of weather on pollution trends have been econometric in nature and tend to struggle with the fact that the effects of weather on observed pollution trends tend to be non-linear, subject to interaction effects and not independent of each other. Attention has therefore turned to the use of predictive machine learning methods as a means of more effectively capturing the influence of meteorological variables on pollution. Grange, Carslaw et al. (2018) for instance develop a weather normalisation technique based on the random forest machine learning model and remove the effect of weather conditions from Swiss PM10 concentrations data. They argue that this technique performs better than more traditional techniques and benefits from the fact that it does not need to conform to strict parametric assumptions. Grange and Carslaw (2019) use the same technique to examine NO_2 and NOx concentrations in London to isolate the effect of the Central London congestion charge. They identify clear features in the pollution data that were not detectable prior to weather normalisation. Finally, Vu et al. (2019) utilise a similar random forest machine learning approach to weather normalise six key pollutants in Beijing from 2013 to 2017. Their analysis reveals the extent to which meteorological variables influence observed pollution data and allows them to identify the effect of the 2013 Beijing Clean Air Action plan.

The purpose of this paper therefore is to quantify the causal impact of the Wuhan lockdown on local air pollution levels. To do this we apply a combination of state-of-the-art machine learning and synthetic control methods to a number of different air pollutants in Wuhan, China. Wuhan is a city of approximately 11.1 million people and was the largest of the 17 cities in Hubei province to be locked down at 10:00AM on 23rd January 2020.[3] Other

[3]An easing of the strict lockdown started on Wednesday 8th April 2020.

large cities in China did not lockdown for at least another two weeks, providing us with a unique natural experiment to investigate how air pollution levels respond to a sudden and abrupt decrease in economic activity. Our contribution is three-fold. First, we apply the latest weather normalisation techniques developed by Grange, Carslaw et al. (2018) and Vu et al. (2019) to remove the effect of local weather conditions on pollution concentrations. To do so we utilise hourly city-level concentrations of four pollutants: sulphur dioxide (SO_2); nitrogen dioxide (NO_2); carbon monoxide (CO); and particulate matter (PM10) between January 2013 and February 2020, for thirty Chinese cities. Second, taking the weather normalised concentrations data, we apply the recently developed "ridge augmented synthetic control approach" (ridge ASCM) developed by Ben-Michael, Feller and Rothstein (2021) to investigate the causal impact of the Wuhan lockdown on local air pollution levels. Ben-Michael, Feller and Rothstein (2021) improve upon the standard synthetic control method by removing the bias that can result from imbalance in pre-treatment outcomes. Third, using a selection of mortality estimates from the existing literature we calculate the potential deaths prevented in Wuhan city, Hubei province and China as a whole, due to improved air quality.

To briefly summarise our results, we find that Wuhan experienced a significant reduction in concentrations of NO_2 and PM10 as a result of the Covid-19 lockdown. Concentrations of NO_2 fell by as much as 24 μg/m^3 during our analysis period in January/February 2020 (a reduction of 63% from the pre-lockdown level of 38 μg/m^3), while PM10 concentrations fell by approximately 22 μg/m^3, albeit for a shorter period (a reduction of 35% from the pre-lockdown level of 62 μg/m^3). It is notable that these reductions brought NO_2 concentrations from a level very close to the WHO safe limit (40μg/m^3) to well within the limit, while PM10 fell from a level way beyond the safe limit (20 μg/m^3) to a level that was still in excess of the safe limit. Perhaps surprisingly, we find no significant reductions in concentrations of SO_2 or CO. Our analysis of the mortality effects associated with the reduced NO_2 concentrations suggests that the lockdowns may have prevented up to 496 deaths in Wuhan, 3,368 in Hubei province and 10,822 in China as a whole.

The remainder of the paper proceeds as follows. Section 3.2 describes our data and methodology. Section 3.3 provides our results and Section 3.4 examines the health implications of our findings. Section 3.5 concludes.

3.2 Data and Methodology

City-level hourly concentrations of four pollutants (SO_2, NO_2, CO, PM10) for thirty Chinese cities were collected from 'Qingyue Open Environmental Data Center' between 18th January 2013 to 29th February 2020.[4] City-level hourly pollution concentrations were calculated by averaging across all of the monitoring stations for each city. Similar data from the same source has been used in a number of other studies including He, Fan and Zhou (2016) and Qin and Zhu (2018). The meteorological data is from the "worldmet" R package (NOAA, 2016) developed by Grange, Carslaw et al. (2018) and includes information on temperature, relative humidity, wind direction, wind speed and air pressure. We then match the city-level hourly pollution data with the city-level hourly meteorological data to generate our data sample.[5] Table 3.1 provides the sources and health impacts of each pollutant and shows they are produced by differing combinations of electricity generation, industrial processes and road traffic. Available evidence suggests that road traffic is likely to be the largest source of CO and NO_2 concentrations in Chinese cities while electricity generation and coal burning will be the largest sources of SO_2. Sources of PM10 are difficult to quantify but will

[4]Thanks to Qingyue Open Environmental Data Center (https://data.epmap.org) for support on Environmental data processing. Thirty Chinese cities include: Wuhan, Shijiazhuang, Zhengzhou, Kunming, Beijing, Shanghai, Guangzhou, Chongqing, Tianjin, Shenyang, Hefei, Changsha, Jinan, Changchun, Guiyang, Xian, Fuzhou, Hangzhou, Taiyuan, Harbin, Huhehaote, Nanning, Nanjing, Chengde, Tangshan, Cangzhou, Xingtai, Baoding, Qinhuangdao, and Zhangjiakou.

[5]Appendix Table B.1 presents detailed information by meteorological monitoring station.

include all of those previously mentioned (Zhao, Wang et al., 2013) (US EPA 2020).

[Table 3.1 about here]

This paper uses two steps to identify the causal impact of the Wuhan lockdown on local air pollution levels. The first step is to conduct a random forest-based weather normalisation on four pollutants separately for thirty Chinese cities to obtain both hourly-observed and weather normalised pollution concentrations. The second step is to aggregate the hourly weather normalised pollution numbers into daily values, and then to take these daily observations for the thirty cities and use a (ridge) augmented synthetic control method on this data to estimate how concentrations levels in Wuhan have changed relative to the synthetic control.

The weather normalisation procedure is conducted on the observed hourly pollution data before running the synthetic control method for two main reasons: First, for policy evaluation analysis within environmental economics, it is difficult to evaluate the efficacy of policy on air pollutant concentration levels since the change of pollutant concentration levels are co-influenced by both meteorological conditions and emission levels. This means that it is difficult to clearly identify whether an improvement in air quality is due to a true fall in emissions or is simply a result of weather conditions that give the appearance of lower concentrations at the measurement stations (Grange, Carslaw et al., 2018; Grange and Carslaw, 2019). As these studies have shown, the best way of accounting for local meteorological conditions is to remove their impact from the pollution concentration observations. Stripping out the local weather effects allows policy makers and social planners to make better informed decisions on the efficacy of previous air pollution interventions which, in turn, will help guide future policy decisions (Wise and Comrie, 2005).

The second advantage of weather normalising is that, according to Abadie (2019), a key principle behind the synthetic control method is the comparative case study, where the impact of a policy intervention can be estimated by comparing the movement of the outcome variable of interest between a single treatment unit and a number of control units. Ideally, the control units should be as similar as possible to the single treatment unit but not exposed to the policy intervention. Abadie (2019) emphasises that, if the outcome variable is highly volatile, researchers will not be able to detect the effect of the policy intervention, no matter what the size of the 'real' intervention effect might be. In our case, daily air pollution concentration levels in Wuhan are extremely volatile which would lead to potential problems of overfitting. If there exists substantial volatility in the outcome variable, Abadie (2019) advises that it is removed in both the treatment and control units prior to applying the synthetic control method. Therefore, weather noise is removed from the observed air pollution concentrations for all thirty cities using the machine learning algorithm. The result is a more reliable estimation of the Wuhan lockdown effect on local air pollution levels, i.e. the pure human activity induced pollution with the natural variability due to weather conditions removed.

3.2.1 Machine learning

In recent years the use of machine learning (ML) techniques has grown rapidly due, in part, to the availability of 'big data' and improved computational power. Supervised ML focuses on prediction problems, given a set of data that contains the outcome variable (the variable of interest) and predictors (a set of independent variables that are used to predict the outcome variables). The whole dataset is split into a training set (used to build up the prediction model) and a test dataset (used to test the prediction accuracy /performance of the model). It is referred to as supervised ML because the outcome variable is available to guide and oversee the process of building the prediction model.

Machine learning is a powerful tool as it provides a way to analyse both linear and nonlinear relationships within the data (Varian, 2014). Increasingly, economists use ML methods in combination with other data analytical approaches. For example, Mullainathan and Spiess (2017) demonstrate the importance of using supervised ML methods in regression analysis while Athey and Imbens (2019) discuss the differences between econometrics and machine learning in terms of goals, methods and settings, and demonstrate the gains from interacting ML and econometrics.

Decision Trees and Random Forests

The meteorological normalisation technique applied in this paper is based on the random forest algorithm. A regression tree is obtained by binary recursively partitioning a single predictor each time over a threshold until a purity of the node is reached (i.e., the node cannot be further split) (Breiman et al., 1984). A decision tree model is easy to train and is highly interpretable. However, decision trees can be prone to overfitting i.e. decision tree predictions can be inaccurate (Hastie, Tibshirani and Friedman, 2009). Hence, predictions obtained from decision trees alone are optimal for a given (training) dataset, but could result in low prediction accuracy for a new dataset (Athey and Imbens, 2019).

To overcome the inherent disadvantage with decision trees, Breiman (2001) introduced the random forest algorithm. The performance of an algorithm is improved if it can be used on a larger number of datasets. One solution, when there is just one dataset, is to add randomness to the data by use of the bootstrap and bagging method (Varian, 2014). Bootstrapping refers to randomly sampling (with replacement) observations from the original dataset, and bagging refers to the process of obtaining an estimation by averaging results from a large number of bootstrapped samples. The random forest algorithm essentially consists of a large number of individual decision trees (grown from different bootstrap samples), and is obtained by averaging the estimations from the whole forest. Compared with a single decision tree, the random forest approach can greatly increase the performance of the prediction.

The random forest approach is relatively simple and fast to train and performs well even when using high dimensionality data (i.e. a large number of predictors / features/ independents). Random forests also allow for a more flexible relationship than that allowed by a simple linear model, as it relaxes the critical assumptions on data that are always required by conventional regression methods (e.g., sample normality, homoscedasticity independence, etc.). In addition, interactions and correlations between the predictors are not restricted. More importantly, a random forest approach provides a measure of the importance of different variables and predictor selections (Varian, 2014; Ziegler and König, 2014).

Weather Normalisation

Grange, Carslaw et al. (2018) were the first to introduce machine learning techniques to weather normalise trends in air pollution data (i.e. the time and meteorological variables). Their approach was to apply a random forest algorithm to predict concentrations of different pollutants at a specific time using a 're-sampled predictor data set'. Take March 15, 2013 for an example. The time and meteorological variables on any given day in the original predictor dataset are randomly selected. The random forest predictive algorithm is repeated 1,000 times and then the different predictors are fed into the random forest model which in turn predicts the concentrations of the different pollutants on March 15, 2013. The final weather-normalised concentrations on this day are obtained by averaging these 1,000 predictions for each pollutant. Note that Grange, Carslaw et al. (2018) not only de-weathered from observed concentration levels, but also removed time trends from the data. The disadvantage is that this approach of de-weathering and removing the time trends can lead to an inability to detect the seasonal variation in the weather normalised concentrations data. This also makes it harder to compare the same time period in different years (which we utilize later in our sensitivity checks).

The solution to the seasonal/time trend problem discussed above is to extend the weather normalizing procedure by de-weathering using only the pollution concentration observations (Vu et al., 2019). The Vu et al. (2019) algorithm includes a new predictor data set that is generated by randomly selecting only the weather variables from the original dataset. For example, for 09:00, 15 March 2013, only the weather variables were randomly selected from the original data set within a 4-week range to construct the new predictors data (i.e. at 09:00 on any date between 1 to 29 March on any year between 2013 to 2018). This process is conducted 1,000 times, and the results fed into the random forest model to give 1,000 predicted concentrations for that specific hour of 09:00, 15 March 2013 using 1,000 columns of randomly sampled weather predictors. The final weather normalised concentration level at 09:00, 15 March 2013 is calculated by averaging the 1,000 predicted concentrations. This means it is still possible to detect seasonal variation within the weather normalized concentrations (Vu et al., 2019).

In this paper we apply the weather normalised procedure of Vu et al. (2019) using the 'rmweather' R packages developed by (Grange, Carslaw et al., 2018). A decision-tree-based random forest model is grown for each of our four air pollutant concentrations for each of our thirty cities, as dependent (output) variables, and the time and meteorological variables as predictors (input variables). Each variable is shown in Table 3.2. For an illustration of the process for building a random forest model and how the weather normalisation process is conducted see (Vu et al., 2019). The whole observed data was randomly sampled into a training set (80%) and a test set (20%). The training set was used to train the random forest model and the test set to test model performance.

[Table 3.2 about here]

Following Vu et al. (2019), a forest of 300 ($n\_trees = 300$) is used and the number of times we sample the whole data and then predict is 300 ($n\_samples = 300$). The number of variables that may split at each node is three ($mtry = 3$) and the minimum size of terminal nodes for the model is three ($min\_node\_size = 3$). For the weather normalisation procedure, the meteorological variables are randomly selected 300 times (within a four-week range) from the observed meteorological dataset (between January 2013 to February 2020). The selection was repeated 300 times and then fed into the random forest model to predict the concentration levels. The final weather normalised concentrations are found by averaging the 300 predicted values from each hour.[6]

3.2.2 The augmented synthetic control method

The synthetic control method (SCM) was first developed by Abadie and Gardeazabal (2003) and has since been used to investigate a number of different questions particularly in labour, development and health economics (see e.g. Cavallo et al., 2013; Kreif et al., 2016; Dustmann, Schönberg and Stuhler, 2017; Mohan, 2017; Xu, 2017). Athey and Imbens (2017) argue that the synthetic control method is "arguably the most important innovation in the policy evaluation literature in the last 15 years".

The design of the SCM is similar to that of the traditional difference-in-difference setting where the goal is to find an appropriate control unit that is comparable to the treatment unit (the city or country that is exposed to an intervention). In this paper, as we are interested in testing the effect of the Wuhan lockdown on local air pollution levels, the ideal solution would be to find a city in China that did not experience a lockdown but is very similar to Wuhan across a range of different characteristics (e.g., the level of economic development, industrial structure, population, current pollution levels, etc.). However, in

[6] The code from Vu et al. (2019) is available from: https://*github.com/tuanvvu/Air_Quality_Trend_Analysis*.

reality no one city is likely to match Wuhan that closely. By taking a SCM approach we employ a data-driven procedure that uses a weighted average of a group of control cities to simulate or construct an artificial or 'synthetic' Wuhan. The goal of the synthetic Wuhan is to reproduce the trajectory of the air pollution levels in real Wuhan before the lockdown. Then, after the lockdown, the difference in the trajectories between the synthetic and real Wuhan can be summarised as the causal impact of the lockdown. In a sense, the synthetic Wuhan is the counterfactual air pollution evolution that Wuhan would have experienced had it not been locked down (Abadie, Diamond and Hainmueller, 2015).

There are a number of advantages with taking a SCM approach. For example, no extrapolation is needed and the synthetic weights are calculated and chosen without using the post intervention data that rules out the risk of specification cherry picking or p-hacking. Moreover, the contribution of each control unit to the overall synthetic unit is explicitly presented so the transparency of the counterfactual allows one to validate the weights using expert knowledge (Abadie, 2019). However, Abadie, Diamond and Hainmueller (2015) caution that the SCM may not provide meaningful estimations if the outcome trajectory of the synthetic unit does not closely match the outcome trajectory of the treatment unit before the intervention.

One solution to concerns about outcome trajectories is proposed by Ben-Michael, Feller and Rothstein (2021) who propose an augmented synthetic control method (ASCM). The ASCM extends the SCM to those cases where a good pre-intervention match between treatment and synthetic unit is not achievable. ASCM uses an outcome model to estimate the bias due to the poor pre-intervention match and then corrects for the bias in the original SCM estimate. The Ben-Michael, Feller and Rothstein (2021) approach is to use a ridge-regularized linear regression model that relaxes the non-negative weights restriction of the original SCM and allows for negative weights within the Ridge ASCM.

In this paper we follow the conventional panel data setting used by Ben-Michael, Feller and Rothstein (2021) given by:

$$Y_{it} = \begin{cases} Y_{it}(0) \quad if \quad W_i = 0 \quad or \quad t \leq T_0 \\ Y_{it}(1) \quad if \quad W_i = 1 \quad and \quad t > T_0 \end{cases} \tag{3.1}$$

Where Y_{it} is the outcome variable of interest, the weather normalised air pollutant concentration levels for four different pollutants, for city i and date t (where i=1,...,N and t =1,..., T), W_i refers to the indicator that city i received the order to lockdown at time T_0 $\leq T$, where $W_i = 0$ is that there never was a lockdown intervention. T_0 refers to the date of lockdown. $Y_{it}(0)$ and $Y_{it}(1)$ refer to the outcome variable of city i in date t within the control group and treatment group (Wuhan in our case), separately.

The estimated treatment effect of interest, the effect of the Wuhan lockdown on local air pollution levels, is given by: $Y_1(1)$ - $Y_1(0)$ = Y_1 - $Y_1(0)$. The SCM imputes the $Y_1(0)$ as a weighted average of the outcome variable within the control group, $Y_0'\gamma$. Ben-Michael, Feller and Rothstein (2021) explain that the way to choose the weights is the solution to a constrained optimization problem. In the special case where the working outcome model is a ridge-regularized linear model, the bias corrector estimator for $Y_1(0)$ can be written as:

$$\hat{Y}_1^{aug}(0) = \sum_{W_i=0} \hat{\gamma}^{scm} Y_i + (X_1 - \sum_{W_1=0} \hat{\gamma}^{scm} X_i)\hat{\eta}^r \tag{3.2}$$

The ridge ASCM can enhance the pre-intervention fit between the synthetic and treatment units compared to the SCM alone by allowing for negative weights. It can also directly penalize the potential extrapolation. Within the ridge ASCM, the hyper-parameter γ plays a significant role in identifying the trade-off between a better pre-intervention match and a larger approximation error.

Our target city, Wuhan, was given the order to lockdown on January 23rd, 2020. The other 29 cities in our sample did not lockdown on this date.[7] However, although they did not lockdown immediately, the majority of the cities in the control group entered a lockdown period between the 3rd and 5th of February 2020. Therefore, in the analysis we examine data up to the 3rd February. This means our analysis is limited to a twelve-day post lockdown period.[8] We use a one-month (thirty days to be precise) pre-period to construct our synthetic Wuhan. Ordinarily, we would set the Wuhan lockdown date as January 23rd 2020 to match the official government announcement that Wuhan would be locked down at 10:00 am on that day. However, following Abadie (2019), if there is an anticipation effect, the researcher should backdate the intervention date to allow for the full extent of the policy intervention to be fully estimated. We therefore test a number of different starting dates and reassuringly our results are not sensitive to the choice of date.

Nevertheless, we set January 21st 2020 as the intervention starting date, since human activity that might affect local air pollution levels may already have been adjusted before the official announcement. More importantly, it is reasonable to believe that some lockdown measures and regulations were being adopted by local government officials prior to the official announcement as it is likely that local officials would have known some time in advance despite things moving so fast during this difficult period. Our choice was also influenced by the trend in NO_2 concentrations that showed a clear reaction on that date.

Finally, before we show the results it is worth putting the Wuhan lockdown in context for those less familiar with the economy of Wuhan and how it relates to our group of control

[7] Appendix Table B.2 provides the list of cities within our control group (and the different city groupings that we use in our sensitivity analysis).

[8] While the 29 cities in our control group were not placed in lockdown during the twelve-day period of our analysis, we cannot rule out that economic and social activity levels in these cities may have been reduced if individuals and businesses responded to what was happening in Wuhan. Although likely to be small in magnitude, if true, this suggests our estimates of pollution reductions are conservative estimates.

cities. Table B.3 in the appendix includes summary statistics for Wuhan and the averages for the other 29 cities in the control group while Figure B.1 (see Appendix two) presents a map of China showing Wuhan and the control cities. The other cities are fairly evenly distributed. Table B.3 indicates that Wuhan is only slightly larger than the control cities in terms of population but has a higher population growth rate. However, it is geographically smaller on average (less than half the size) so has a higher population density. On average it is richer than the average of the control cities and is ranked around fifth in terms of per capita gross regional product. Wuhan is a city of 11 million people and a major industrial hub. The dominant industries include automobiles, manufacturing of electronic and optical communication equipment, pharmaceutical and chemical manufacturing, and iron and steel manufacturing. The automobile sector is a particularly important and includes, for example, the $9.4 billion Chinese automotive company Dongfeng Motor Corp, that has joint venture partners with Nissan and Honda.

In terms of air pollution Wuhan is not particularly out of line with other cities of comparable size. Figure 3.1 plots the annual average observed concentrations for the four pollutants that we use in the paper across the 30 cities in our data sample for the period 2013 to 2019. Wuhan is denoted by the red line. As can be seen, Wuhan is towards the lower end in terms of SO_2, but has relatively higher levels of NO_2. In terms of CO and PM10, Wuhan is around the average. This figure shows that Wuhan is fairly typical.

[Figure 3.1 about here]

In terms of the lockdown itself, all transport in and out of Wuhan was shut down, including the closure of public transit, trains, airports, and major highways. In addition, in a now familiar story across the world, all shops were closed except those selling essentials, all private vehicles were banned (except for those with a special permit), all public transport was banned (except for a small number of taxis), public gatherings were prohibited, and there was a policy of enclosed community management. However, key producers of steel, chemicals and semiconductors remained in operation as well as electric utilities.

In addition, it is possible to get some idea of the reduced movement of people within Wuhan. If one looks at Baidu Migration data (provided by Baidu which is the dominant search engine in China), based on the Location Based Service platform of Baidu Maps, we can observe real-time population movements including a "daily out-flow migration index of a city", a "daily in-flow migration index of a city" and a "daily within-city migration index of a city" (Fang, Wang and Yang, 2020). For this paper we looked at the "within city migration index" to give an indication of the intensity of the within city traffic movement before the Wuhan lockdown (22 Jan 2020) and after the lockdown. What the results show is that movement levels fell to very low levels in Wuhan compared to the other 29 cities. This indicates how effective the lockdown was in terms of restricting movement and how such restrictions were certainly not in place in the other 29 cities for this period. The reduced movement of people also helps to explain the reduction of NO_2 which is a result of the "traffic lockdown", i.e. the restriction of traffic mobility/ or the reduction in traffic-relate emissions.

3.3 Results

3.3.1 Machine learning results

Figure 3.2 presents a plot of daily pollution concentrations to show the overall trends in the observed data (grey line) and the weather normalised (red line) data for SO_2 ($\mu g/m^3$), NO_2 $\mu g/m^3$), CO (mg/m^3) and PM10 ($\mu g/m^3$) respectively between January 2013 and February

2020.[9] It can be seen that both observed and weather normalised concentration levels have generally fallen over time, particularly so for SO_2. This reduction has been driven in part by strict government regulations and a desire to reduce local air pollutants. More importantly, Figure 3.2 illustrates the impact of our weather normalization process with clear differences being seen between the observed concentrations and weather normalised concentrations with the latter being a much smoother data series.

[Figure 3.2 about here]

Concentrating on the more recent period, Figure 3.3 presents the daily plots of observed and weather normalised trends for SO_2, NO_2, CO and PM10 in Wuhan between 21st December 2019 to the 3rd of February 2020. Again, it is clear that the trends in the weather normalised pollutants are less volatile and noisy compared to the observed values and shows the extent to which weather conditions influence recorded pollution levels from stations. Figure 3.3 also illustrates how difficult and potentially misleading it would be to identify the effect of the lockdown (the dotted vertical line) on pollution concentrations using observed values of pollution only.

[Figure 3.3 about here]

3.3.2 The Impact of the Wuhan lockdown on local air pollution using ridge ASCM

To present the results we consider each of our four pollutants in turn. The plots in figures 3.4 to 3.9 are all plotted using plus or minus one standard error. Figure 3.4 (left) plots the difference in the weather normalised NO_2 (NO_2wn) levels between synthetic Wuhan and Wuhan. Figure 3.4 (right) plots the trend in the weather normalised NO_2 level of both

[9] Where "wn" refers to a weather normalised pollutant e.g. SO_2wn is weather normalised SO_2.

synthetic and real Wuhan. The vertical line again refers to the Wuhan lockdown date. For NO_2 we chose January 21st 2020 as the intervention start date as we found a significant anticipation effect for NO_2wn, i.e., the NO_2wn in Wuhan began to fall significantly and substantially below that of synthetic Wuhan from January 21st, 2020. As can be seen, the synthetic Wuhan does a good job in simulating the NO_2wn trend in Wuhan before the lockdown. Both trends were around 45 to 52 μg/m^3 between December 21st and the 27th (notably above the WHO safe limit of 40 μg/m^3) before they began to fall in January 2020 to around 35 to 40 μg/m^3. The fall coincides with the Spring break in China where economic activity usually drops considerably. After the 21st January 2020, a large and significant gap opens up between NO_2wn emissions in Wuhan and synthetic Wuhan with a peak difference of around 24 μg/m^3, equivalent to a reduction of 63% of the level of NO_2 concentrations (38 μg/m^3) immediately prior to the lockdown. As time goes on the gap between the series closes a little but is still more than 15 μg/m^3 at the end of the twelve-day period. Notably, NO_2 has fallen to a limit that is now comfortably below the WHO safe limit. The right figure plots the trend between synthetic and real Wuhan between December 21st 2019 and 3rd February 2020. The blue vertical line again refers to the intervention date (21st January 2020). The synthetic Wuhan weather normalised NO_2 levels were consistently between 33 to 40 μg/m^3, whereas the level in Wuhan dropped substantially to around 20 μg/m^3 three to four days after lockdown, and remained below 20 μg/m^3 until the end of study period. The results show that the lockdown led to a large reduction in weather normalised NO_2 level in Wuhan.

[Figure 3.4 about here]

We now consider SO_2. Figure 3.5 (left) plots the difference in the weather normalised SO_2 (SO_2wn) level between synthetic Wuhan and real Wuhan while Figure 3.5 (right) plots the evolution of trends in weather normalised SO_2 levels for both synthetic and real Wuhan. The vertical line is drawn on the Wuhan lockdown date of January 22, 2020. As shown, differences in SO_2wn between the synthetic and real Wuhan are negligible suggesting that the other 29 cities did a good job in simulating the trajectory of pollution concentrations in Wuhan. After the lockdown, the SO_2wn level in Wuhan was around 3 μg/m^3 lower than if Wuhan had not been locked down. However, the reduction disappears three to four days after lockdown and returns to the same trend that the other 29 cities were following. It is worth noting that even at the three to four-day mark, which is equivalent to a 3 μg/m^3 reduction in SO_2 in Wuhan, the reduction is only marginally significant.

[Figure 3.5 about here]

Moving on to CO, Figure 3.6 plots the results for weather normalised CO levels. In this case, synthetic Wuhan is not a good match with the pre-policy real Wuhan. This means we cannot confidently draw conclusions on the impact of the Wuhan lockdown on local CO levels.

[Figure 3.6 about here]

Finally, Figure 3.7 (left) plots the difference in the weather normalised PM10 (PM10wn) level between synthetic Wuhan and real Wuhan. Figure 3.7 (right) plots the trend of weather normalised PM10 level of both synthetic and real Wuhan. The vertical line coincides with a Wuhan lockdown date of January 22nd, 2020. The trajectories of synthetic Wuhan and real Wuhan were closely matched prior to the lockdown. After the lockdown the trends begin to diverge with the difference increasing to 22 μg/m^3 four to five days after lockdown (a reduction of 35% from the pre-lockdown level of 62 μg/m^3). The fall in Wuhan became significant on the third or fourth day. Notice that after seven to eight days the difference in the trends became insignificant. Thus, the lockdown of Wuhan led to a significant but short-lived reduction in PM10 levels and did so from levels that were way above the WHO safe limit of 20 μg/m^3 to levels that were still beyond this limit.

[Figure 3.7 about here]

To summarise the baseline results, they demonstrate that, relative to the control, the Wuhan lockdown led to a large and significant reduction in NO_2 concentrations, a smaller and more short-term reduction in PM10, but no significant fall in SO_2 levels. For CO the pre-policy fit was not considered strong enough for us to draw any firm conclusions.

3.3.3 Placebo tests

To validate our baseline results we follow Abadie, Diamond and Hainmueller (2015) and conduct a series of placebo tests. We begin with an in-time placebo test and then estimate an in-place placebo test and finally we rerun our estimations using alternative control groups for NO_2 and PM10. The placebo test results give us confidence that our main findings are not through chance.

In-time placebo test

For the in-time placebo test we assume that the Wuhan shutdown happened on the same date but one or two years earlier, in either 2018 or 2019. Figure 3.8 shows the results of in-time placebo tests for NO_2wn. Apart from the date of the lockdown we use the exact setting and run the exact same code for the placebo test. On the left-hand side of Figure 3.8 we focus on the data period between 21st December 2018 to 3rd February 2019 and then set the fake lockdown to be 21st January 2019. The right-hand side of Figure 3.8 presents the results for the period between 21st December 2017 and 3rd February 2018 with a fake lockdown set at 21st January 2018. For both in-time placebo tests we did not find any significant reductions in NO_2 for these two fake lockdown dates. Figure 3.9 presents the results for PM10wn. Again, for both 2018 and 2019 there was no significant difference between synthetic Wuhan and real Wuhan.

[Figure 3.8 about here]

[Figure 3.9 about here]

In-place placebo test

Our second placebo test is an in-place test. We randomly assign the lockdown policy to one of the other 29 control cities. Given there was no lockdown in any other city on that date we would not expect to find any sizable reduction effect. Our approach is to assign each of the other cities to be a 'synthetic Wuhan' and again use the exact same setting and code to run the ridge ASCM model. Figure 3.10 plots the difference between a synthetic trend of 29 different lines using 29 different control cities, plus our main findings on NO_2 for the real Wuhan lockdown (the red line). The real Wuhan stands out from the other 29 lines, none of which showed a similar reduction (over 20 μg/m^3).

[Figure 3.10 about here]

The results for PM10wn are a little different from the NO_2wn results in that we found similar size effects for four of our synthetic Wuhan lines. However, the results for these four cities are not significant. If we drop these four lines (representing Shijiazhuang, Jinan, Hangzhou and Huhehaote) we have the right-hand figure, where the red line stands out in the early period of the lockdown. The results are consistent with our baseline findings for Wuhan weather normalised PM10 levels where we only found a significant reduction two to seven days after lockdown which is where the red line on the right figure shows the largest reduction compared to the remaining grey lines.

[Figure 3.11 about here]

Alternative control groups

Our final sensitivity check is to use a range of different control groups to run the ridge ASCM model to check whether the results are sensitive to the initial choice of our 29 large cities. In addition to the full 29 city control group, we also re-estimate the results using four alternative control groups that we call synthetic control groups 1, 2, 3 and 4 (Syn_CG1, Syn_CG2, Syn_CG3 and Syn_CG4). The detailed list of each control group is provided in Appendix Table B.2. The alternative control groups use Province capitals only; Northern cities only (that may be more similar to Wuhan); a smaller group of cities that did not experience a lock down before March 2020; and a final group that did lock down after 3rd February.

Figure 3.12 shows the results from creating a synthetic Wuhan from four alternative control groups on weather normalised NO_2 and PM10, respectively. Figure 3.12 shows that all five control groups closely match the pre-lockdown trends for NO_2wn and PM10wn. The five different controls also show similar post intervention trajectories suggesting that our findings of the causal impact of Wuhan lockdown on local NO_2 and PM10 level are not sensitive to the choice of control group.

[Figure 3.12 about here]

3.4 The Health Implications of China's Falling Pollution

Having established the impact of Wuhan's lockdown on pollution concentrations we undertake a simple back of the envelope exercise to calculate the potential lives saved as a result of the improved air quality. For simplicity we focus only on the reduction in NO_2

concentrations.

Our results in Figure 3.4 indicate that the reduction in concentrations of NO_2 varied between 15 μg/m^3 and 24 μg/m^3 during the period between the start of the lockdown and the end of our estimation period in early February. Since we do not have a usable control group beyond early February, we are unable to estimate how long the reductions in pollution continued for but, for the purposes of this exercise, we model lives saved if concentrations fell by 20 μg/m^3, and a more conservative estimate of 10 μg/m^3, over the full 2.5 months of the lockdown.

To begin, we draw upon a number of studies that have estimated the mortality effects of NO_2 concentrations. Next, from the National Bureau of Statistics we calculate the monthly mortality rate in Wuhan (0.045917%) which we apply to the population of Wuhan which was 11.081m in 2019.[10] We then calculate how much lower mortality would have been over 2.5 months as a result of our estimated reduction in pollution.

Table 3.3 summarises the various studies that have estimated the mortality effects of NO_2 concentrations and presents the range of estimated effects. Table 3.4 then utilises each of these effects, in the manner described above, to produce our estimates of lives saved. As can be seen, the estimated lives saved in Wuhan city as a result of the full 2.5 month lockdown range from 183 to 496 for a 20 μg/m^3 reduction in NO_2 and between 92 and 248 for a 10 μg/m^3 reduction.

[Table 3.3 about here]

When Wuhan went into lockdown on 23rd January it did so along with 16 other cities within Hubei province, affecting a total population of 59.02 million. While our analysis of the reduction in NO_2 concentrations is within Wuhan city, it does not seem unreasonable to

[10] http://cjrb.cjn.cn/images/2019-03/25/6/25R06-07C_Print.pdf.

assume all cities in Hubei province experienced a similar reduction in pollution given they were subject to an equally stringent lockdown for the same length of time. Table 3.4 therefore also reports lives saved as a result of a 20 $\mu g/m^3$ reduction in NO_2 concentrations across the whole of Hubei province. These range from 1,228 to 3,368 for a 20 $\mu g/m^3$ reduction in NO_2 and between 614 and 1,684 for a 10 $\mu g/m^3$ reduction.

For completeness, we extend our analysis to all regions subject to lockdown within China. By early February 2020, a total population of over 233 million were subject to formal lockdown (including Hubei's 59 million).[11] While it is difficult to be clear of the strength and duration of all the lockdowns outside of Hubei we here assume they resulted in the same reduction of 20 $\mu g/m^3$ NO_2 concentrations and did so over a slightly shortened lockdown period of 2 months. Table 3.4 provides the results and indicates that lives saved range from 3,940 to 10,822 for a 20 $\mu g/m^3$ reduction in NO_2 and between 1,970 and 5,411 for a 10 $\mu g/m^3$ reduction

[Table 3.4 about here]

It is important to stress that these are little more than back of the envelope calculations and rely on a number of assumptions, in addition to those already pointed out regarding the stringency and duration of the lockdowns. First, we are modelling lives saved as a result of a reduction in concentrations of a single pollutant, NO_2. A similar exercise could be undertaken for our estimated reduction in PM10 concentrations. However, there remains some uncertainty as to whether health impacts of different pollutants, particularly NO_2 and particulate matter, are truly independent of each other given how highly correlated they tend to be. Nevertheless, some evidence of independence has been found by Faustini, Rapp and Forastiere (2014) suggesting that our focus on NO_2 may provide an underestimate

[11] See https://en.wikipedia.org/wiki/2020_Hubei_lockdowns. The figure of 233 million is likely to be a conservative estimate since many other parts of China had some restrictions on activities and/or were experiencing informal lockdowns as individuals chose to stay at home where possible.

of the true mortality benefits of the reduced concentrations of these two pollutants. Second, we are assuming that the mortality response is proportionate to the reduction in pollution i.e. a 20 $\mu g/m^3$ reduction in concentrations has double the mortality effect of a 10 $\mu g/m^3$ reduction. Similarly, we are assuming that a two-month reduction in pollution has double the mortality benefits of a one-month reduction. Third, in predicting the possible lives saved due to a lockdown-induced reduction in pollution we are ignoring any other potential mortality effects caused by the lockdown such as increased exposure to indoor pollution, mental health effects, reduced road traffic accidents and so on. Finally, there is a possibility that those most susceptible to pollution exposure, i.e. those with underlying respiratory or other health conditions, are also those most susceptible to Covid-19. As such, if these individuals are dying from Covid-19 then we may be over-estimating the lives saved due to cleaner air.

Nevertheless, our results suggest that the lockdowns in China resulted in significant reductions in mortality as a result of improvements in air quality alone.

3.5 Discussion and Conclusions

Faced with a pandemic that is unprecedented in modern times, governments around the world have introduced strict lockdowns to try to control the spread of Covid-19. Inevitably, such a stringent, far-reaching policy will have wide-ranging impacts in addition to that of disease control. Using the example of Wuhan's Covid-19 lockdown, this paper has examined one such impact, the perceived reduction in air pollution due to reductions in traffic volumes and economic activity more generally.

We adopted a two-stage approach. First, to isolate the impact of the lockdown on pollution concentrations we removed the confounding effects of weather conditions using a random forests machine learning approach (Grange, Carslaw et al., 2018; Vu et al., 2019).

This approach overcomes the difficulties of econometrically controlling for non-independent, non-linear weather conditions. Our analysis reveals the importance of removing weather conditions from pollution patterns. Analysing observed (non-weather normalised) pollution levels, or pollution levels where weather conditions have not been fully controlled for, could provide misleading conclusions as to the impact of the lockdown. Second, we adopt a new Augmented Synthetic Control Method Ben-Michael, Feller and Rothstein (2021) to examine how weather normalised concentrations of four pollutants responded to the lockdown using a control of 29 other Chinese cities that were not in lockdown.

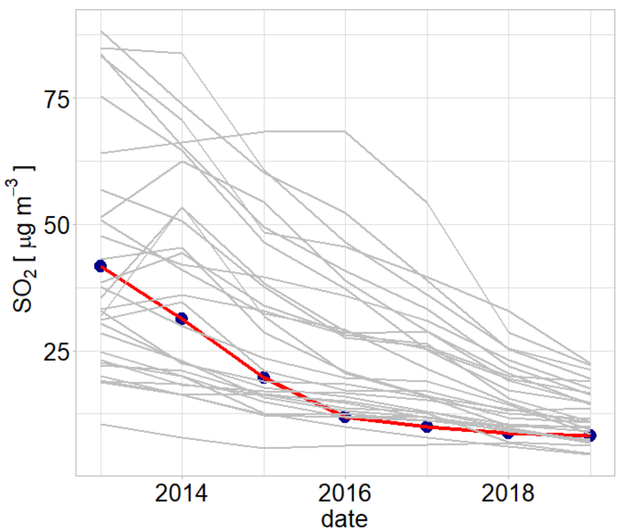
Our results indicate that the impact of the lockdown varied by pollutant, a nuance that newspaper reports of cleaner post-lockdown air have generally failed to acknowledge. We find that concentrations of NO_2, a pollutant closely tied to traffic volumes and fossil fuel use, fell by as much as 24 $\mu g/m^3$ following the lockdown (a 63% fall) although this reduction declined to 16 $\mu g/m^3$ by the end of our twelve-day window of analysis. Prior to the lockdown NO_2 concentrations were very close to the WHO health limit and so this reduction brought those concentrations to within safe limits. Concentrations of PM10 also fell by over 20 $\mu g/m^3$ although this reduction was short term and not statistically significant for the duration of our twelve-day window. Interestingly, concentrations of SO_2 and CO did not fall in a statistically significant manner following the lockdown. In the case of SO_2 this is likely to reflect the country's reliance on coal-fired power plants and the fact that temperatures were relatively low in Wuhan through much of this period, resulting in a need for domestic heating. It is less clear why CO, a pollutant largely emitted by transport, did not fall following the lockdown.

Finally, we employ a selection of estimates of the mortality effects associated with NO_2 concentrations to calculate the potential lives saved as a result of the cleaner air. We find that reduced NO_2 concentrations following lockdown may have prevented as many as 496 deaths in Wuhan city, 3,368 deaths in Hubei province and 10,822 deaths in China as a whole. While these potential deaths prevented may outweigh the official Chinese death toll from Covid-19 itself, our findings should not in any way be interpreted as implying that the pandemic has yielded net benefits to China. As we have pointed out, our estimates of deaths prevented are little more than back of the envelope calculations and should be treated with a degree of caution.

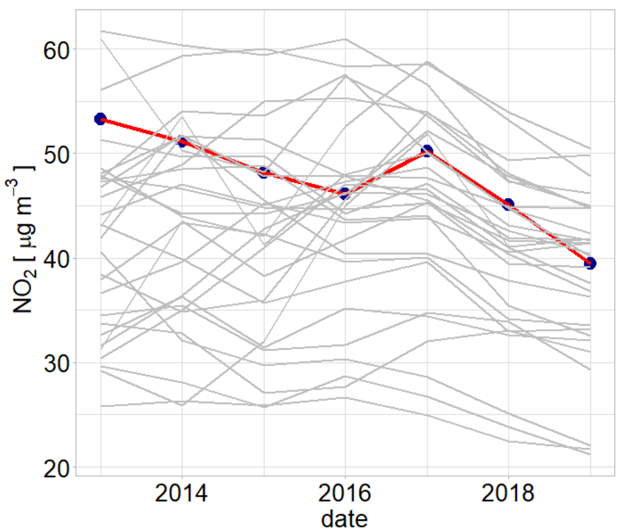
While a city-level lockdown may provide some clues as to how milder forms of restrictions on human activities might impact human health such as: congestion charging; pedestrianised zones; and urban planning more generally; because these all happened at the same time during a lockdown, estimating the individual impacts would be a challenge. However, the large NO_2 effect does suggest that policies to reduce emissions from vehicles, such as a push for the electrification of cars and buses, would have considerable health benefits. How one would measure the costs incurred by society following a return to business as usual is also a challenge. One approach is to estimate the health costs incurred at the city level using published hospital statistics data and then using micro-simulation for modelling the long term impacts (Public Health England, 2020). A second approach is to elicit a value of statistical life (VSL) in an air pollution context. For example, the OECD (2012) developed a new method for calculating country-specific VSL and estimated the cost of deaths from outdoor pollution for OECD countries to be almost $ 1.5 trillion in 2013.

Finally, despite the inherent difficulties in estimating the cost savings from any new emission reductions, the purpose of our analysis is to show that a policy as stringent as a lockdown has far-reaching implications which extend well beyond the primary purpose of disease control. Indeed, since air pollution, Covid-19, and the health of the population more generally are inextricably linked, then policy makers need to be aware of these interactions when formulating policy in the ongoing fight against Covid-19 and future pandemics.

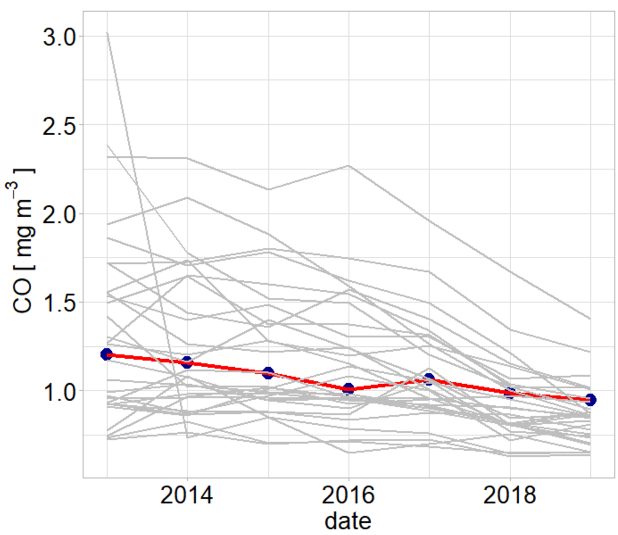
Chapter 3 Figures and Tables



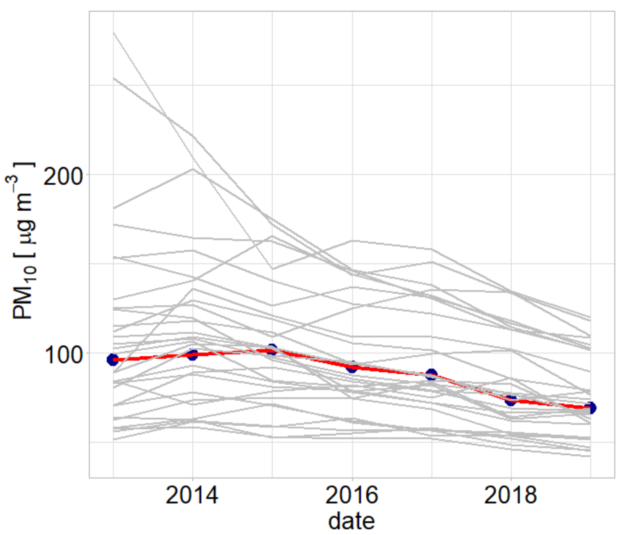
(a) SO_2



(b) NO_2



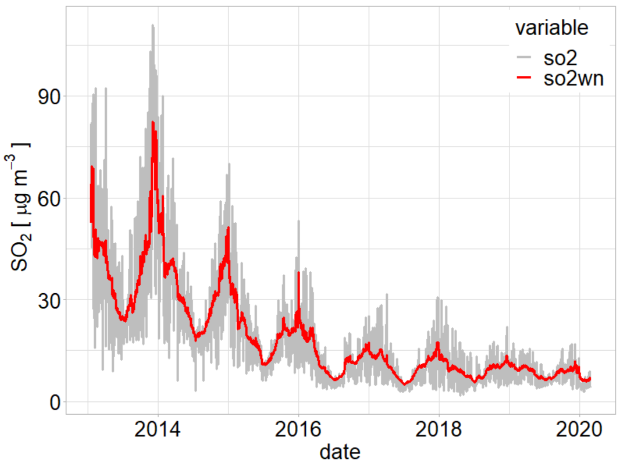
(c) CO



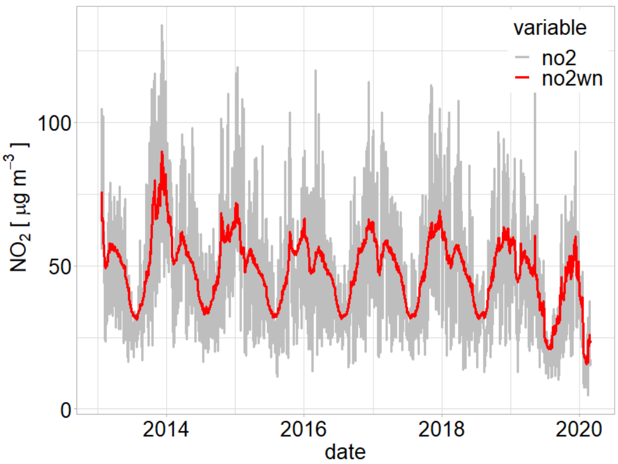
(d) PM10

Note: Wuhan is denoted by the red line.

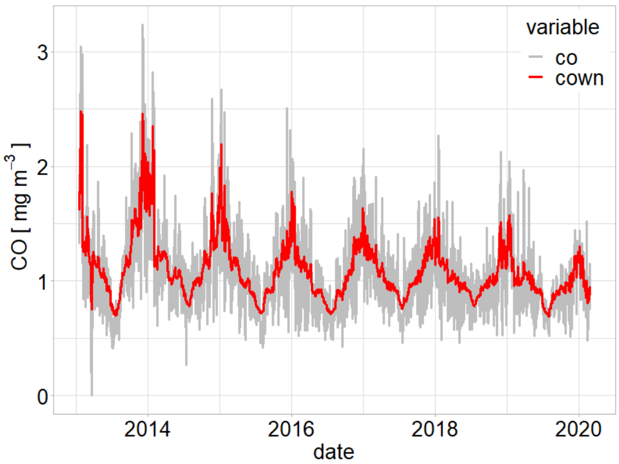
Figure 3.1: The annual average observed concentrations of SO_2, NO_2, CO and PM10 in Wuhan and 29 control cities between 2013 and 2019.



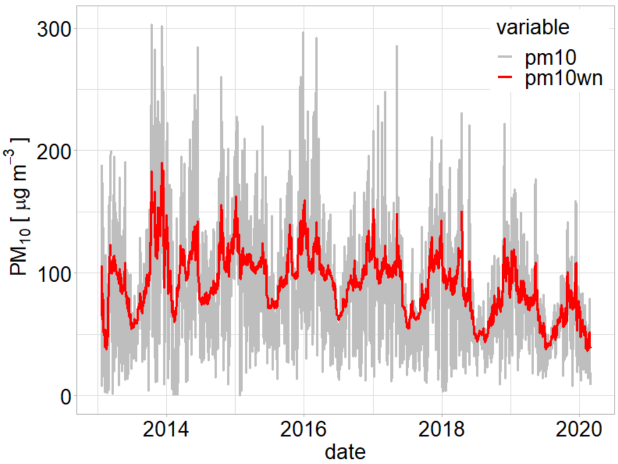
(a) SO_2



(b) NO_2

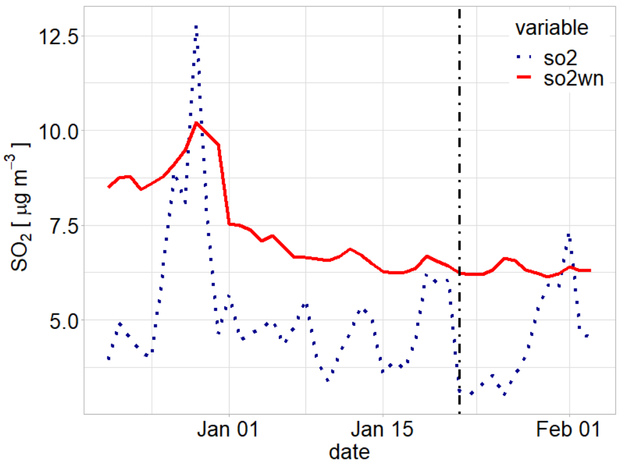


(c) CO

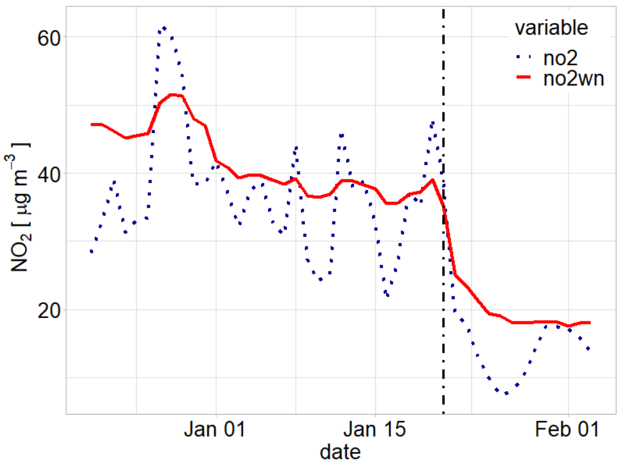


(d) PM10

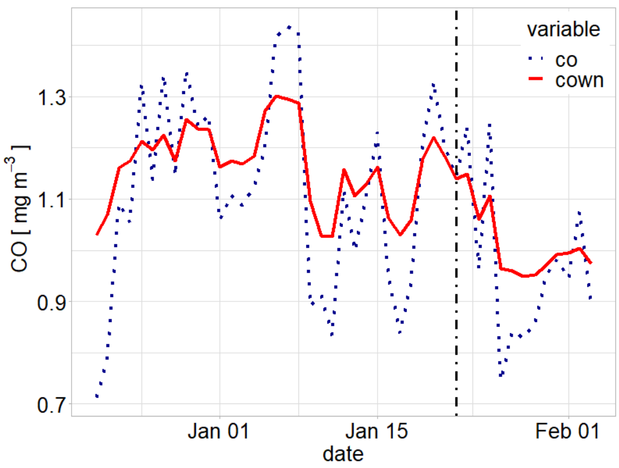
Figure 3.2: Daily averages of observed and weather normalised concentrations of SO_2, NO_2, CO and PM10 in Wuhan between January 2013 and February 2020.



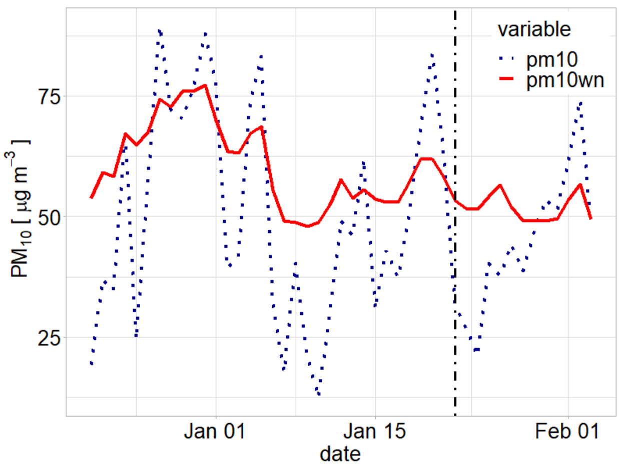
(a) SO_2



(b) NO_2

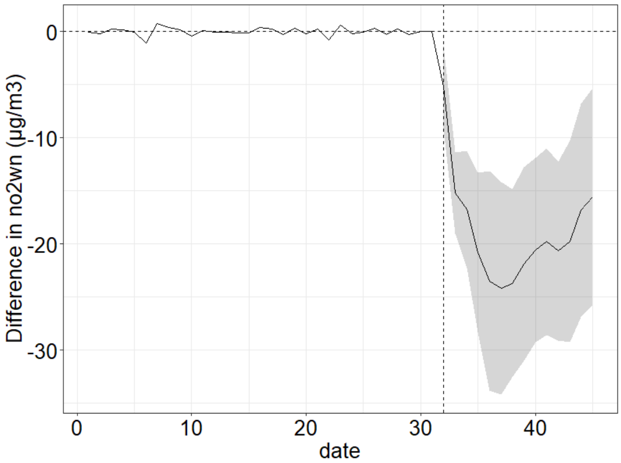


(c) CO

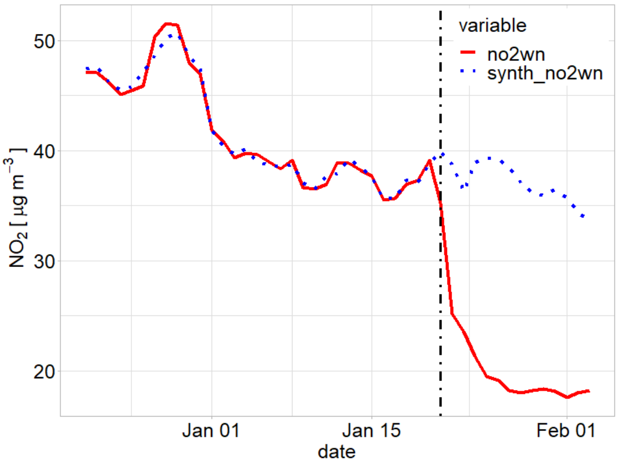


(d) PM10

Figure 3.3: The comparison of daily observed and weather normalised concentrations of SO_2, NO_2, CO and PM10 in Wuhan between 21st December 2019 and 3rd February 2020.



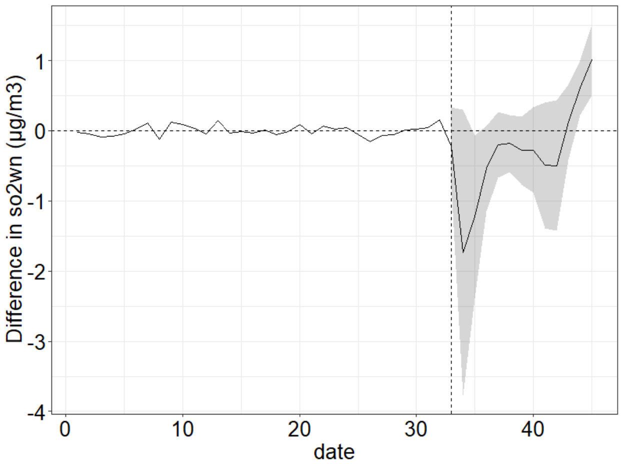
(a) NO_2



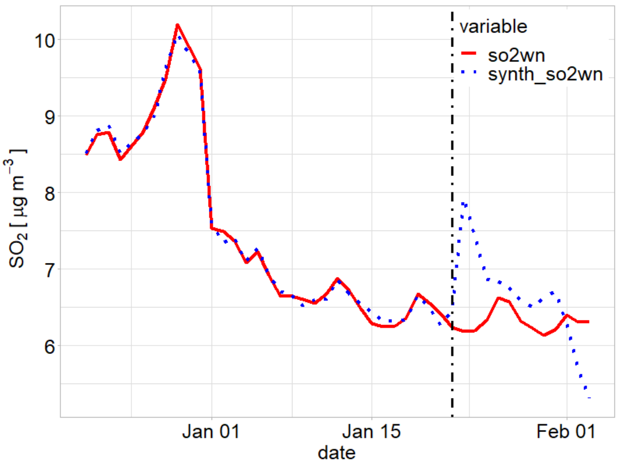
(b) NO_2

Note: Left hand figure shows point estimate ± one standard error of the ATT.

Figure 3.4: Ridge ASCM results on weather normalised NO_2 concentrations in Wuhan.



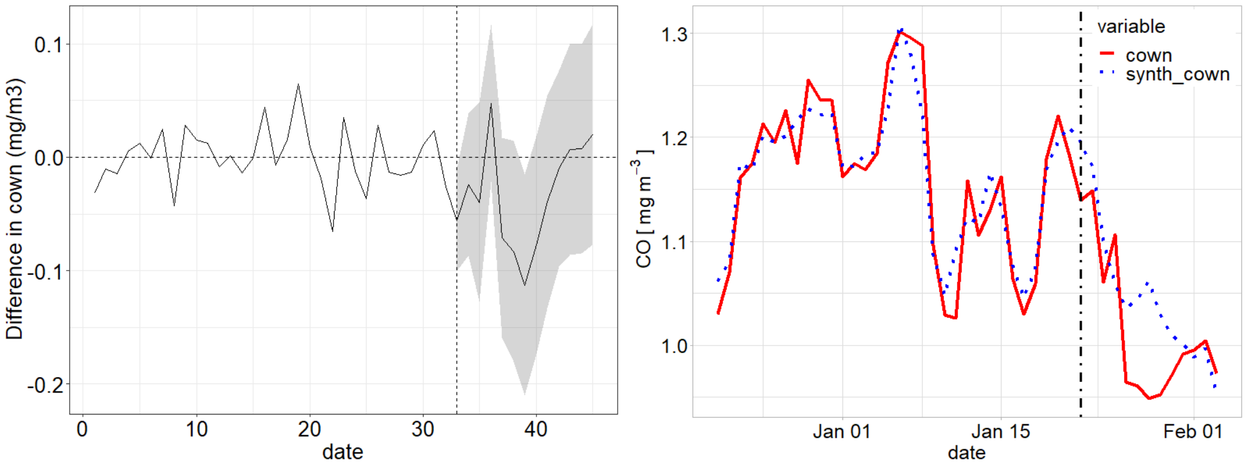
(a) SO_2



(b) SO_2

Note: Left hand figure shows point estimate ± one standard error of the ATT.

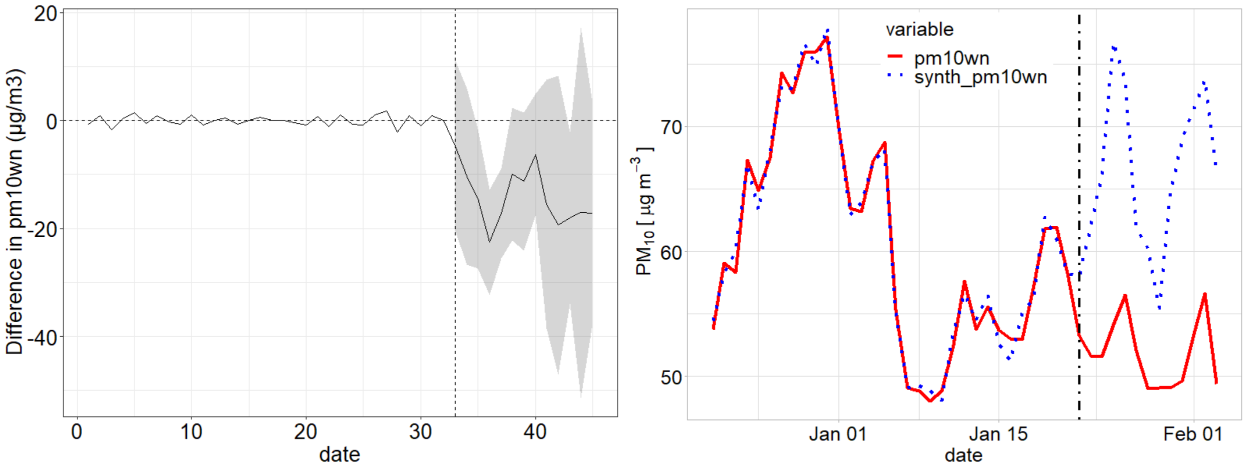
Figure 3.5: Ridge ASCM results on weather normalised SO_2 concentrations in Wuhan.



(a) CO (b) CO

Note: Left hand figure shows point estimate ± one standard error of the ATT.

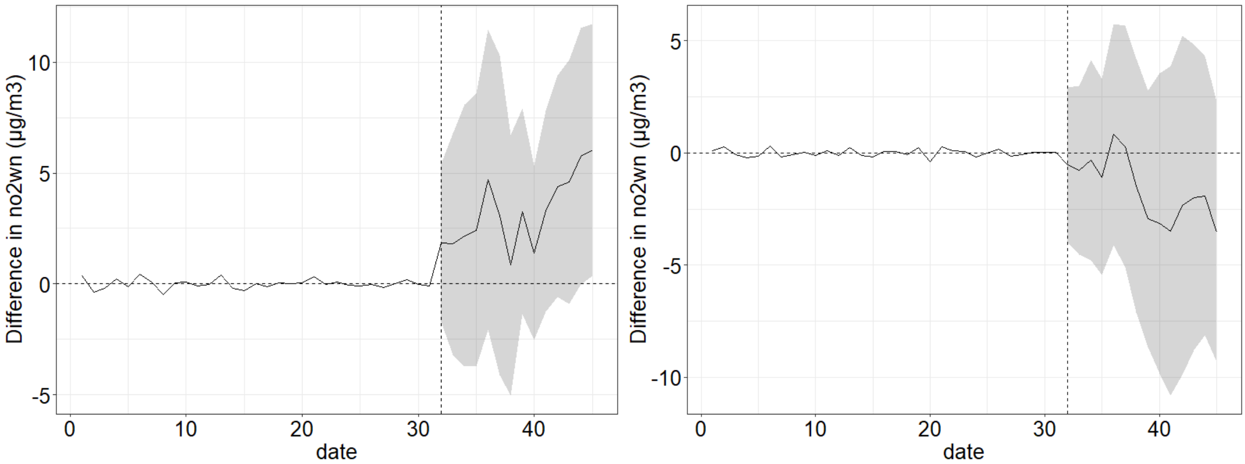
Figure 3.6: Ridge ASCM results on weather normalised CO concentrations in Wuhan.



(a) PM_{10} (b) PM_{10}

Note: Left hand figure shows point estimate ± one standard error of the ATT.

Figure 3.7: Ridge ASCM results on weather normalised PM10 concentrations in Wuhan.

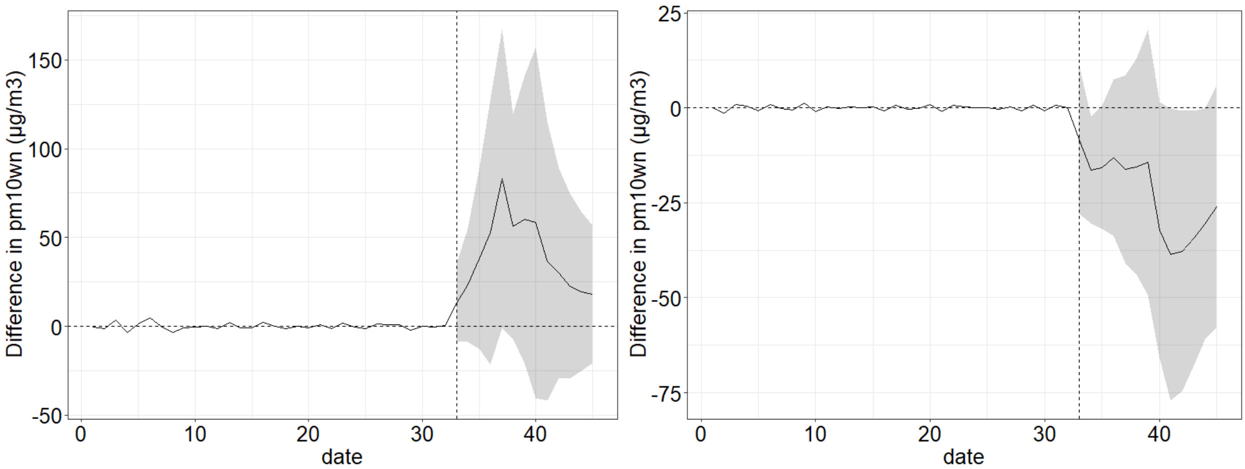


(a) NO_2wn 2019

(b) NO_2wn 2018

Note: Both figures show point estimate ± one standard error of the ATT.

Figure 3.8: The in-time placebo test results of NO_2wn using 21st January 2019 (left) and 21st January 2018 (right) as Wuhan lockdown date.

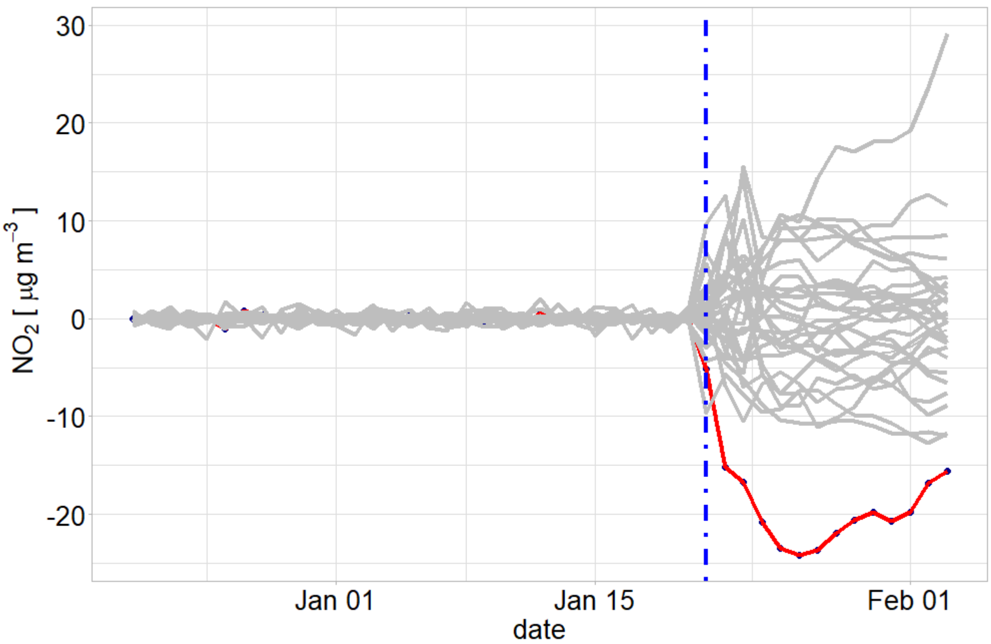


(a) PM10wn 2019

(b) PM10wn 2018

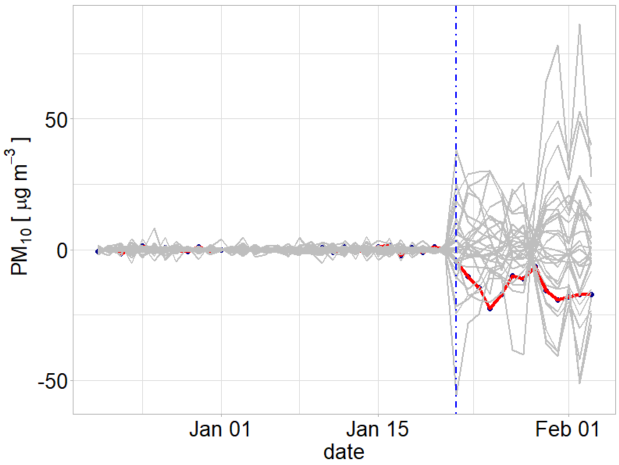
Note: Both figures show point estimate ± one standard error of the ATT.

Figure 3.9: The in-time placebo test results of PM10wn using 22nd January 2019 (left) and 22nd January 2018 (right) as Wuhan lockdown date.

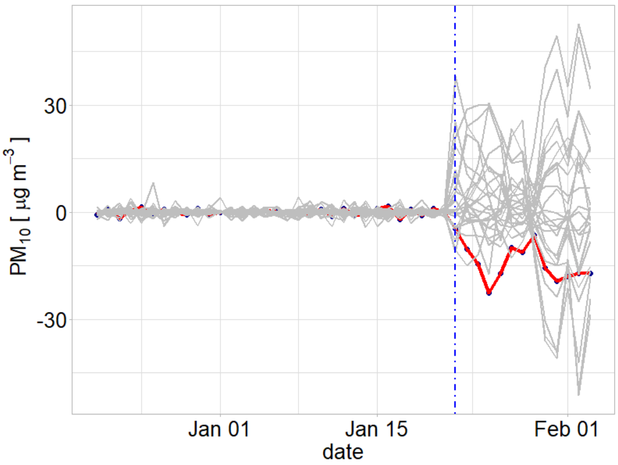


Note: We randomly assign the lockdown policy to one of the other 29 control cities and compare with Wuhan (in red).

Figure 3.10: The results of in-place placebo test on NO_2wn



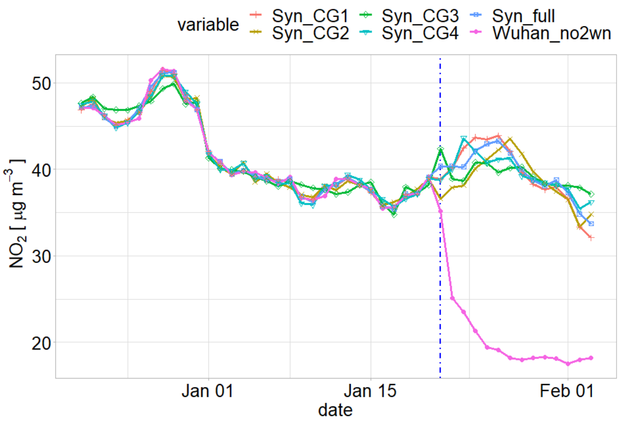
(a) PM10wn



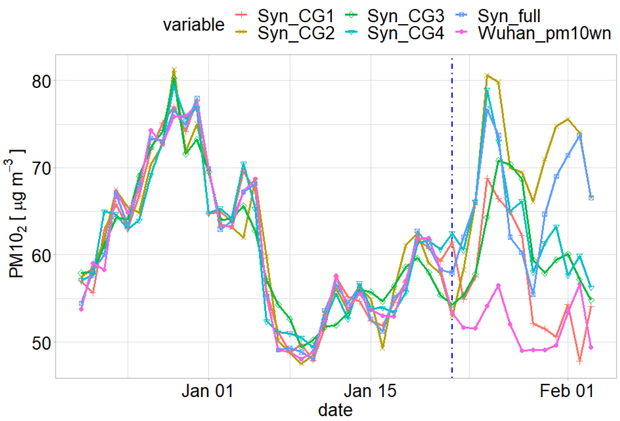
(b) PM10wn

Note: The left figure plots the results using all 30 cities, the right figure plots the results after dropping Shijiazhuang, Jinan, Hangzhou, Huhehaote.

Figure 3.11: The results of in-place placebo test on PM10wn.



(a) NO_2wn



(b) PM10wn

Note: Appendix Table A2 defines each control group.

Figure 3.12: The results of alternative control group tests on NO_2wn (left) and PM10wn (right).

Table 3.1: Sources and health effects of our four pollutants

Pollutant	Sources	Health Effects
Nitrogen Dioxide (NO_2)	Combustion processes, mainly for power generation, heating and motor vehicles	Respiratory difficulties, reduced lung function
Sulphur Dioxide (SO_2)	The burning of sulphur-containing fossil fuels, mainly from power generation, domestic heating and transport	Respiratory difficulties, irritation of the eyes
Coarse Particulate Matter (PM10)	Road transport and the burning of fuels for industrial, commercial and domestic uses	Respiratory difficulties and cardiovascular disease
Carbon Monoxide (CO)	Incomplete burning of fossil fuels in transport and industrial processes	Cardiovascular disease

Source: WHO (2018).

Table 3.2: A list of input and output variables used in this study

Input variables (predictors, independent variables)	
Time variables	
date_unix	Number of seconds since 1970-01-01, represents the trend in pollutant emissions
day_Julian	Day of the year, represents the seasonal variation
weekday	Day of the week, represents the weekly variation
hour	Hour of the day, represents the hourly variation
Meteorological variables	
temp	Temperature (degrees Celsius)
wd	Wind direction (m/s)
ws	Wind speed (in degrees, 90 is from the east)
RH	Relative humidity (%)
pressure	Atmospheric pressure (millibars)
Output variables (dependent variables): Air pollutant concentrations:	
SO_2 (μg/m^3), NO_2 (μg/m^3), CO (mg/m^3), PM10 (μg/m^3)	

Table 3.3: Previous literature on the NO_2 mortality association

Author	Study region	Period	Key finding
Tao et al. (2012)	Pearl River Delta of southern China	2006-2008	"10 μg/m^3 increases in average NO_2 concentrations over the previous 2 days were associated with 1.95% increase in total mortality"
Faustini et al. (2014)	Meta-analysis of 23 studies	2004-2013	"An increase of 10 μg/m^3 in the annual NO_2 concentration was associated with a 1.04% rise in mortality"
Mills et al. (2015)	Quantitative systematic review of 204 global studies	Before 2011	"A 10 μg/m^3 increase in 24 h NO_2 was associated with increases in all-cause mortality of 0.71%"
Chen et al. (2018)	272 Chinese cities	2013-2015	"A 10 μg/m^3 increase of NO_2 concentrations ... resulted in increments of 0.9% for total mortality"
Atkinson et al. (2018)	48 studies of 28 cohorts (global)	Before 2014	"Each 10μg/m^3 increment in NO_2 is associated with 1.02 increase in all-cause mortality"

Table 3.4: Previous literature on the NO_2 mortality association

Region:	Wuhan	Wuhan	Hubei	Hubei	China	China
Mortality effects:	20 $\mu g/m^3$	10 $\mu g/m^3$	20 $\mu g/m^3$	10 $\mu g/m^3$	20 $\mu g/m^3$	10 $\mu g/m^3$
Tao et al. (2012): 1.95%	496	248	3,368	1,684	10,822	5,411
Faustini et al. (2014): 1.04%	265	133	1,795	898	5,772	2,886
Mills et al. (2015): 0.71%	183	92	1,228	614	3,940	1,970
Chen et al. (2018): 0.90%	230	115	1,555	778	4,994	2,497
Atkinson et al. (2018): 1.02%	260	130	1,763	882	5,660	2,830

Note: Since the mortality effects in the previous literature are estimated for a 10 $\mu g/m^3$ change in NO_2 concentrations we here double them to capture a 20 $\mu g/m^3$ change. The period of lockdown is assumed to be 2.5 months in Wuhan and Hubei province and 2 months for China. Monthly mortality rates are 0.045917% in Wuhan, 0.058333% in Hubei and 0.05941667% for China as a whole. The locked down populations sizes are 11.08m in Wuhan, 59.17m in Hubei and 233.5m in China as a whole. All mortality rates and population levels are from the National Bureau of Statistics of China. A worked example: lives saved in Wuhan from a 20 $\mu g/m^3$ reduction in NO_2 using Tao et al.'s (2012) mortality estimates are calculated as 2.5((11,081,000 * 0.00045917)*(0.0195*2) = 496.

Chapter Four

Can city level environmental inspections reduce local air pollution? A machine learning and ASCM approach

4.1 Introduction

China's economic growth over the last three decades has been dramatic. However, one of the consequences of such rapid growth has been a deterioration in a range of environmental indicators including those associated with public health. For example, Rohde and Muller (2015) estimate that in 2014 air pollution was the cause of an additional 1.6 million deaths in China. Air pollution is also also linked to an reduced life expectancy as certain diseases become more prevalent (Yin et al., 2020).

To address the problems associated with air pollution the Chinese government has passed a range of different policies aimed at reducing the environmental impact of rapid economic growth. However, many of the earlier policies failed to meet their targets. A contributing factor to the relatively poor performance is the way in which central government policies were implemented and enforced at the local government level with local government officials willing to sacrifice the environment in pursuit of economic growth. Although it took some time, the Chinese central government began to change the incentive structure so that local government officials would give higher priority to environmental issues (Zheng and Kahn, 2017). In the past China has also implemented various policies aimed at reducing air pollution. For example, Chen, Jin et al. (2013) and He, Fan and Zhou (2016) estimated the air quality improvements from policies designed to reduce air pollution during 2008 Beijing Olympic Games and Liu et al. (2016) examined a similar "APEC blue" policy implemented ahead of the APEC 2014 meeting.

The purpose of this paper is to investigate the effectiveness of the Central Environmental Inspection Policy (CEIP) which, after being introduced in 2016, is now one of the central pillars of China's environmental policy regime. The objective of the CEIP is to encourage local government officials to exert greater effort to protect the environment and involved centrally-appointed inspection teams being dispatched to cities across China for about one month at a time to inspect both the local government officials themselves and the firms in the cities. The CEIP is an ambitious policy (and is still ongoing in 2021) and demonstrates the determination of the Chinese central government to reduce harmful local emissions. However, despite the apparently far reaching powers accorded to the CEIP there are still unanswered questions regarding the efficacy of the policy both in terms of reducing pollution and also the negative impact on the local economy.[1]

The contribution of this paper is to evaluate the effectiveness of the CEIP in reducing the emissions of a number of local air pollutants. Our identification strategy is to focus on Hebei province which was the first to be visited, without prior notice, by the inspection

[1]For details of the CEIP see http://www.scmp.com/news/china/policies-politics/article/1987159/china-sends-environmental-inspection-teams-8-more. [ADD in official link to a document that describes the policy - can be in Chinese:http://www.mee.gov.cn/gkml/sthjbgw/qt/201601/t20160104$_3$20975.htm]

team in January 2016. More specifically, we combine weather normalisation techniques (a random forest-based machine learning model) from atmospheric sciences and the Augmented synthetic control method (ASCM) to provide a causal estimate of the impact of a province having a visit from the CEIP inspection team on local air quality. First, we use machine learning techniques to remove the influence of weather conditions on observed air pollution concentrations, and second, we run a Ridge ASCM model on the weather normalised pollution concentrations taking advantage of the quasi-experimental design of the Hebei CEIP visit schedule.

There are a number of examples in the recent literature that have attempted to test the efficacy of different environmental policies aimed at reducing local air pollution. For example, for the US, Greenstone (2003), Greenstone (2004) and Auffhammer, Bento and Lowe (2009) show how the U.S. Clean Air Act improved air quality levels. However, in developing countries, issues related to compliance and enforcement of environmental policies associated with Weak institutions means that policies are often judged to be ineffective (Duflo et al., 2013; Greenstone and Hanna, 2014). In the case of China the results are mixed. Some studies find that environmental policies do improve environmental quality due, in part, to the administrative power that exists within the Chinese political system (Zheng and Kahn, 2013) while other studies, such as (Chen, Jin et al., 2013) and (He, Fan and Zhou, 2016) on the 2008 Beijing Olympic Games and Liu et al., (2016) for the "APEC blue" policy, while finding that the policies were effective, these effects were only relatively short term.

There are also a small number of papers that have examined various aspects of the CEIP. For example, using firm level data and regression discontinuity approach, Wang, Fan and Chen (2021) investigate the effectiveness of the CEIP on firm production and emission levels in a single Chinese province and find that it resulted in some polluting firms being shut down with a related drop in chemical oxygen demand (COD) emissions. Xiang and Gevelt (2020) took a different approach and interviewed different stakeholders in Hebei to try to understand how effective the CEIP was for the enforcement of environmental regulation and concluded that although it improved regulation enforcement, it was rather resource intensive and inefficient.

The research that is most closely related to our own are those that look at air pollution. For example, Wang, Sun and Zhang (2021) take a regression discontinuity approach to investigate the impacts of CEIP (looking at the first round and when provinces are inspected a second time) on Air Quality Index (AQI) and six pollutants. The results show an immediate improvement of AQI, $PM_{2.5}$ and PM10. Similarly, Jia and Chen (2019) take a DID approach to identify the effectiveness of the CEIP (1st and 2nd rounds) on local air quality levels and find improvements in both the AQI and $PM_{2.5}$ levels.

Compared to Wang, Sun and Zhang (2021) and Jia and Chen (2019), the design of our study means that we are able to address a number of possible concerns. First, our weather normalisation technique means we are able to carefully account for prevailing weather conditions where the failure to do so may lead to misleading results given the complex nature of atmospheric processes (e.g., changes in meteorological conditions might mask the real changes in emission strength). Second, we are able to examine the pre-parallel trend assumption (not formally tested in Jia and Chen (2019)) which which might affect the validity of DID estimations. In our case we were unable to validate the assumption. Third, previous studies have tended to use multiple provinces/cities as a treatment group which makes it difficult to identify a city-specific impact of the CEIP which is that the ASCM model allows us to do for Hebei as a whole, and for individual cities within Hebei.

To briefly summarise our results, we find that the Hebei inspection led to a substantial, but short term, reduction in the emissions of $PM_{2.5}$, PM10 and SO_2 in Hebei as a whole. The weather normalised value of $PM_{2.5}$ and PM10 fell by as much as 25.89 $\mu g/m^3$ (20.81%) and 47.80 $\mu g/m^3$ (22.61%) immediately after the inspection, although pollution levels gradually returned to previous levels within three months. Results from a standard DID model (where we are able to run tests for pre-parallel trends), show that the standard DID model overestimates the policy effect, partially due to the lack of a parallel trends prior to the policy being implemented, and the inability to properly control for the underlying weather conditions. In-time and in-place placebo tests confirm our results (and the use of alternative control groups). Simple back-of-the-envelope calculations suggest that, the 3-month reduction of PM10 that can be attributed to the CEIP reduced the number of deaths in Hebei province by between 565 and 1,048.

The remainder of the paper is organized as follow. Section 4.2 describes the CEIP inspection policy in more detail while Sections 4.3 and 4.4 discuss data and empirical strategy, respectively. Section 4.5 presents the main results and Section 4.6 discusses the potential benefit of the falling emissions attributable to the policy. Section 4.7 concludes.

4.2 Background: the Central Environmental Inspection Policy (CEIP)

In 2016 the Chinese government introduced the Central Environmental Inspection Policy (CEIP) which, after the first inspection in Hebei province, was rolled out systematically across the country on a province by province basis. The CEIP implementation involves sending centrally appointed inspection teams to a number of cities within a given province to investigate how successful local authorities have been in reducing emissions of local air pollutants through the enforcement of existing regulations. After the initial inspection for Hebei province in January 2016, the vast majority of the remaining provinces were divided into four groups which were inspected during November 2016, April 2017 and August 2017, respectively.[2]

Each inspection round consists of an inspection team that contains a number of central governmental officials, ministerial staff (mainly from the State Council, the Ministry of Ecological Environment and the Ministry of Natural Resources), and relevant experts. A team was then stationed in each of the targeted provinces for around one month. The central governmental inspectors are authorized to schedule meetings or one-to-one talks with top local government officials, to investigate firms without prior warning, and to hold any government officials to account for poor environmental performance. According to the official government reports, local governments were not able to anticipate an upcoming inspection (and hence respond in advance), since the inspection targets were kept secret. Even after a team arrives in a city they are authorized to meet any officials and to investigate any individual firm without prior warning. The general public are also able to participate by reporting local environmental pollution issues to the inspection teams who will process and record any reports that come in. For details on the operations and measures of the one-month inspection see Appendix C.1.

There are several key features that distinguish the CEIP policy from previous environmental policies. Most pertinent is that the CEIP has support from the highest level of the Chinese political system which means that the inspection teams are given a level of authority that means that first, the inspection teams themselves consist of high level state and ministerial level officials and second, that they are able to interview high level province and city level officials and also investigate any institution or firm within a given province or

[2]For completeness, after the initial Hebei inspection, Inner Mongolia, Heilongjiang, Jiangsu, Jiangxi, Henan, Guangxi, Yunnan and Ningxia were inspected between 15 July and 15 August 2016, Beijing, Shanghai, Hubei, Guangdong, Chongqing, Shaanxi and Gansu were inspected between 25 November to 25 December 2016, Tianjin, Shanxi, Liaoning, Anhui, Fujian, Hunan and Guizhou were inspected between 25 April to 25 May 2017, and finally, Jilin, Zhejiang, Shandong, Hainan, Sichuan, Tibet, Qinghai and Xinjiang were inspected from 15 August to 15 September 2017.

city.[3] Previous environmental protection policies have tended to rely heavily on "command-and-control" where pollution mitigation targets are set by the central government, and local (provincial and city level) government officials and firms are required to comply with. However, it has been argued that local governments often 'turn a blind eye' to the environmental violations by local firms as long as they contribute to economic growth, since the economic performance remain the major criteria for the promotion prospects of local officials (Wang, 2013).[4] In theory, the CEIP should have broken this link between local officials and local firms by assigning higher level officials to inspect lower level officials with the former given the power to influence the promotion prospects of thousands of local officials depending on the results of the inspection. The prospect of being inspected at any time, should therefore, provide a strong incentive for local officials to enforce environmental regulations and reduce local air pollution.

For the purpose of this study, we focus on the inspection on Hebei province in January 2016 (Figure C.1 in the appendix illustrates the location of Hebei cities and control cities). We focus on the Hebei inspection for several reasons. First, the Hebei inspection provides a "quasi-experimental" design in the sense that during January 2016, only certain cities in Hebei were inspected (the treatment group), while no other city in China was inspected which provides us with a viable control group. More specifically, our treatment group consists of eight individual cities in Hebei province and our control group consists of 27 cities from other Chinese provinces. This identification strategy allows us to employ difference-in-difference and synthetic control methods for policy evaluation (Abadie and Cattaneo, 2018).

The second reason to focus on Hebei province is that historically, Hebei experienced

[3] China has a five-layer vertical political governance system at the nation, province, city, county and township levels. Promotion of lower level government officials is decided by the level above. In this case the political level of inspection teams is usually higher than that of inspected targets (Zheng and Kahn, 2013).

[4] For more details see http://www.scmp.com/news/china/policies-politics/article/1987159/china-sends-environmental-inspection-teams-8-more.

relatively high levels of air pollution. A study by the Ministry of Ecological Environment (MEE) in 2013 suggests that seven of the ten most polluted cities in China were from Hebei province due it part to the province's reliance on heavy industry such as energy generation, steel, machinery, and chemical industries. As a result, Hebei is one of the highest energy consumers in China and the largest producer of steel and because of the provinces dependency on heavy industry there has been reluctance from local officials and firms to the implementation and enforcement of existing environmental protection policies (Wang, 2013). A third reason is that, Hebei was the first province to be inspected which means we do not have to address anticipation effects where local officials will have learnt from the experience of those provinces that went before them. Although historically Hebei has been one of the most polluted provinces in China there was no reason for officials in Hebei to think that they would be the first province to be inspected or when they would be inspected.

To get some idea on how seriously the policy was implemented, a government report provides information on the strength of the inspection. For example, according to the feedback published by the Hebei inspection team in May 2016, (by the 8th April 2016), 2,856 environmental issues were reported to the team by the general public and were subsequently addressed, 200 firms were found to have issues with illegal pollution and were banned and shut down, 123 individuals were detained, 65 government officials were questioned about environmental pollution issues, 60 were criticised and 366 were held accountable for poor environmental performance.[5]

4.3 Data

Our data are obtained from 'Qingyue Open Environmental Data Center' and consists of two years of hourly air quality data (between 1st June 2014 and 30th May 2016) for four

[5] For details see http://www.gov.cn/xinwen/2016-05/03/content_5070077.htm.

pollutants ($PM_{2.5}$, PM10, NO_2 and SO_2) from 35 Chinese cities. City level pollutant concentrations are calculated by averaging across all the state control air quality monitoring stations within each city. Eight of the 35 cities are located in Hebei province (the treatment group) with the remaining 27 cities being the control group.[6] Our air pollution data starts in 2014 as this reflects the time when there was a significance increase in the quality and reliability of air quality data from surface monitoring stations across China thanks to the central government's self declared “war" on pollution (Barwick et al., 2019; Greenstone, He et al., 2021).

After initial data processing, we follow Grange, Carslaw et al. (2018) and Shi et al. (2021) and match the hourly air quality data with hourly meteorological data. Table 4.1 describes the input variables (predictors), including four time variables and ten meteorological variables. The meteorological data are obtained from three different sources. The surface observed meteorological data include temperature (temp), relative humidity (rh), wind direction (wd) and wind speed (ws) are collected from the Integrated Surface Database (ISD) of National Oceanic and Atmospheric Administration (NOAA) using the ‘worldmet’ R package (https://CRAN.R-project.org/package=worldmet).

[Table 4.1 about here]

Data on the boundary layer height (blh), total cloud cover (tcc), surface net solar radiation (ssr), surface pressure (sp) and total precipitation (tp) are collected from the

[6] Thanks to Qingyue Open Environmental Data Center (https://data.epmap.org) for support with processing the data. The 27 control cities are: Changchun, Changsha, Chengdu, Chongqing, Fuzhou, Guangzhou, Guiyang, Hangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Lhasa, Nanchang, Nanjing, Nanning, Shanghai, Shenyang, Taiyuan, Urumqi, Wuhan, Xian, Xining, Yinchuan, Zhengzhou. The treated cities in Hebei province are: Shijiazhuang, Cangzhou, Xingtai, Baoding, Tangshan, Zhangjiakou, Qinhuangdao, Chengde. The cities of Handan, Langfang and Hengshui are also located in Hebei province but were not included in the analysis due data availability issues.

ERA5 reanalysis data set (ERA5 hourly data on single levels from 1979 to present) from the ECMWF (European Centre for Medium-Range Weather Forecasts). The reanalysis data set was regridded into 0.25 degrees (regular lar-lon grids), and the hourly data are extracted from the NetCDF format based on the corresponding latitude and longitude of each of the 35 cities. (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview).

Finally, to take into account the transport of air mass clusters, we calculated the 72-hour back trajectories using the HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) model from NOAA for each of the 35 cities. We then clustered the 72-hour back trajectories into 12 clusters using (the Euclidian distance) 'openair' R package (https://CRAN.R-project.org/package=openair). Figure C.2 in the appendix shows the plots of 72-hour back-trajectory clusters in Beijing, Shanghai, Lanzhou and Changsha, respectively. The 72-hour back-trajectory cluster can be used to represent the air masses that the city were exposed to and thereafter account for the potential spatial spillover issues due to regional transport of pollutant concentrations (Zongbo to elaborate a bit on why it is important to include the "72-hour back-trajectory cluster"in the analysis).[7]

As a result, ten meteorological variables and four time variables are used as input variables to predict hourly air pollution concentrations and then to conduct the weather normalisation procedure to obtain the hourly deweathered (weather normalised) pollution concentrations. The hourly observed and weather normalisation data are then converted into weekly frequencies to enable us to run the DID and Ridge ASCM model estimations.

[7]See Shi et al. (2021) for detailed description.

4.4 Empirical strategy

The standard approach to policy evaluation is to use DID methods. In this study we adopt the standard DID approach to identify the policy effect, but also provide an alternative approach that combines machine learning and and ASCM to estimate the causal impact of the CEIP on air quality levels in cities in Hebei.

Our ML ASCM approach uses a weather normalisation machine learning technique to decouple the effects of meteorological conditions from observed air pollutant concentrations. We then use the deweathered (weather normalised) air pollutant concentrations to run a Ridge ASCM (Augmented Synthetic Control Method) model to investigate the 'real' ("more reliable") causal impact of the inspection policy, i.e. the anthropogenic effect (the effect that can only be attributed to emission reduction measures) and not the effect on the observed pollutant concentration levels that are interactively determined by emissions and weather conditions (Cole, Elliott and Liu, 2020).

We take an interdisciplinary approach to identifying a causal relationship between an inspection and pollution concentrations for a number of reasons. From a natural (atmospheric) science perspective, identifying the impact of air pollution control policies is challenging since the air pollutant concentrations are collectively affected by meteorological conditions and emission levels. This means environmental scientists tend to rely on chemical transport models to evaluate policy effects. However, the high uncertainty in emission inventories affect the robustness of their results (Vu et al., 2019). Statistical analysis is also widely used by environmental scientists to examine trends in air quality trends but are limited when it comes to taking into account issues related to multicollinearity and interaction effects between different explanatory variables (Grange, Carslaw et al., 2018).

The solution, proposed by Grange, Carslaw et al. (2018) and Grange and Carslaw

(2019), is to employ a machine learning based weather normalisation technique, which can address most of the challenges discussed above when analysing trends in air quality. This method has been adopted by numerous studies to investigate the response of air quality to a range of different policies implemented in different regions across the world (e.g., (Zhang, Vu et al., 2019; Cole, Elliott and Liu, 2020; Shi et al., 2021). For example, Vu et al. (2019) applied a machine learning based random forest model to evaluate the impact of Beijing Clean Air Action Plan implemented in 2013 and found that a series of primary control policies under Beijing Clean Air Action plan led to a significant reduction in emissions from 2013 to 2017 but also that meteorological variables had a significant impact on observed air pollutant concentrations. As Vu et al. (2019) point out, they were unable to attribute the reduction in air pollution to any one policy as a number of policies were implemented over the same period.

Some of the concerns raised by Vu et al. (2019) can be addressed using standard econometric tools (e.g., DID), which have the advantage of identifying the causal impact of certain policy interventions under the context of 'quasi-experimental or experimental designs' (Greenstone and Gayer, 2009). The key idea for the causal inference in policy evaluation is to identify the impact of an exogenous policy shock on the variable of interest, if both the states of receiving and not receiving policy treatment can be observed. Athey and Imbens (2017) provides a comprehensive overview of the identification strategies that have been widely used in economic studies for evaluating the causal impact of the policy intervention, including the Regression Discontinuity Design (RDD), the Difference-in-Difference (DID) and the Synthetic Control Method (SCM).

Here we combine weather normalisation techniques from environmental science with advanced econometric techniques which, when used together can help overcome certain weaknesses with each method separately. For example, the ASCM can help identifying the causal impact (under the natural experiment design of the CEIP) that weather normalisation techniques alone are not able to address, while weather normalisation helps to remove the impact of weather conditions on air pollutant concentrations that standard statistical analysis is not able to address.

As a result, our empirical strategy is three fold. First, we use the standard DID approach to identify the policy effect, conducting the appropriate pre-parallel trend tests, which are crucial to get a valid and unbiased estimation of the CEIP effect. Second, we compare the DID method results with the ML methods for dealing with the weather variables to show trends in air quality and show that, in our case, the ML approach has several advantages. Third, under the “quasi" natural experiment design of the CEIP, we adopt the Ridge ASCM approach on weather normalised air pollution concentration data, to investigate the causal impact of the policy on emissions (i.e., the pure changes in emission levels after removing the impact of local weather conditions). We argue that the ML technique offers an alternative way to provide valid and reliable policy evaluation when studying air pollution related policies.

4.4.1 DID and the pre-parallel trends

DID methods are the standard approach for policy evaluation in environmental economics. To identify the impact of the CEIP on air quality in Hebei province we follow He and Wang (2017) and He, Pan and Tanaka (2020) and estimate the relative change of air quality between Hebei (eight city average) and 27 control cities across China. To ensure we have a consistent sample all the data are converted from hourly to weekly intervals and January 2015 to December 2015 was set as the pre inspection period, and January to May 2016 as the post inspection period. We drop the data after June 2016 since the next round of inspections started after this date and the inclusion of cities in provinces then being inspected may bias the results. Our estimating equation is given by:

$$Y_{it} = \beta \times Inspection_{it} + \alpha \times X_{it} + \mu_i + \pi_t + \epsilon_{it} \quad (4.1)$$

where Y_{it} refers to the concentration level (or the logged form) of observed and weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in city i and week t. $Inspection_{it}$ represents the inspection status in city i and week t, which is equal to 1 if city i was inspected in week t and 0 otherwise. X_{it} refer to a vector of weather controls that includes temperature, relative humidity, wind direction, wind speed, boundary layer height, total cloud cover, surface net solar radiation, surface pressure and total precipitation. μ_i and π_t denote the city and week fixed effects, respectively. City fixed effects μ_i control for the city specific time-invariant factors, like the geographical conditions and short term economic and industrial variables, and the week fixed effects π_t control for the shocks that are common to all cities in a give week (He, Pan and Tanaka, 2020).

Our variable of interest is β, which estimates the difference in air pollution levels between Hebei (treated group) and the 27 cities (control group) before and after the inspection period, i.e. the potential impact of an inspection on air quality levels in inspected cities. To ensure that the results are comparable between DID and the ML+ASCM, we average the data for the eight Hebei cities to create one "city" to represent 'Hebei'.[8]

The underlying assumption of an unbiased estimation of β is that the trajectory of air pollution concentrations between treatment and control cities share parallel trends before Hebei was inspected in January 201. Therefore, following He and Wang (2017), we test the parallel trend assumption using the following event study approach:

[8]We also estimated our results for each city separately using both DID, and ML + ASCM but the general conclusions are unchanged but are available from the authors upon request.

$$Y_{it} = \sum_{m \geq -6, m \neq -1}^{m=2} \beta^m Inspection_{it}^m + \alpha \times X_{it} + \mu_i + \pi_t + \epsilon_{it} \tag{4.2}$$

where $Inspection_{it}^m$ are a series of dummy variables that denote the inspection status at different periods, m refers to the 2-month window period before and after the inspection, for example, $m = -6$ denotes January and February 2015, $m = -5$ denotes March and April 2015, $m = -4$ denotes May and June 2015, $m = -3$ denotes July and August 2015, $m = -2$ denotes September and October 2015, $m = 0$ denotes the inspection month (January 2016), $m = 1$ denote February and March 2016, $m = 2$ denotes April and May 2016. Note that $m \neq -1$ suggests that the period of November and December 2015 was omitted so that the inspection effects are relative to the period immediately before the inspection starts (November and December 2015 acts as the "reference" month). The coefficients β^m measure the difference in air quality between the treated and control cities in period m relative to the difference 2-month prior to the inspection. To ensure the pre-parallel trend assumption holds for the DID method, there should be no significant difference of the air quality between the treated and control cities prior to the policy being implemented, i.e. the estimation of β^{-6} to β^2 should be close to 0.[9]

4.4.2 Weather conditions and air quality: Comparing DID and Machine learning

There is a growing literature that combines machine learning and econometrics to solve the real world problems. Varian (2014) provides a comprehensive introduction to the main ML methods that can be used as 'new tricks' in economic studies (e.g., classification and regression trees, Random forest, etc.) while Mullainathan and Spiess (2017) discuss how

[9] We also show the results of using 1-month windows to estimate the pre-parallel trend test in Figure C.3 of the appendix.

supervised machine learning can be used in empirical analyses. Athey and Imbens (2019) summarise both supervised and unsupervised ML tools and argue that the interaction of ML and econometrics can provide better performance in certain cases.

Econometrics and ML are both about "learning from data". However, they differ in terms of their goals, methods and settings (Athey and Imbens, 2019). Econometrics tries to identify, detect and summarise relationships among data, for example using linear regressions to produce parameters estimations, while machine learning sees data as a pure prediction problem, i.e. predicting the outcome variable using a set of predictors. The main difference is that econometrics is “model driven" and assumes that the data structure is given by the existing economic theory, while ML has no restrictions on the data structure and thus is a “data driven" procedure. Econometrics has the advantages that it can provide causal inference and can use confidence intervals which can be difficult for ML. In short, while both methods are useful, for this study where we want to control for the impact of weather conditions on air pollutant concentrations, we argue that ML (Random forest model) is more appropriate.

Our argument is that rather than specifying a specific data structure in advance, ML (Random forest) has the advantage of being “data driven", which means that through the use of an algorithm it is possible to discover a complex data structure and allow for more flexible functional forms (Mullainathan and Spiess, 2017). ML also has the advantage of being nonparametric, and thus does not require assumptions such as sample normality, homoscedasticity and independence to hold. In the case of the analysis of air pollution, concentrations vary a lot depending on the different underlying weather conditions. For example, high wind speed is known to reduce pollutant concentrations by increasing atmospheric dispersion, but this relationship is generally non-linear. The impact of temperature and atmospheric pressure on pollutant concentrations is likely to also be non-linear, and can interact with other weather variables Grange and Carslaw (2019). Modelling the atmospheric process can be challenging and burdensome for traditional statistical analysis due to the nature of complexity, non-linearities and collinearity (of the atmospheric process). For example, OLS regressions need to “hand-curate" which functional form or interaction terms should be included in the regressions (Mullainathan and Spiess, 2017).

Developed by Breiman (2001), Random forest (RF) is a decision tree based ensemble learning algorithm that belongs to supervised ML. The key technique of the RF model is the bagging (bootstrap aggregation) technique, which refers to randomly sampling data observations (and predictors) from the original data set and then predicting the outcome variable using different sampling data and the sampling predictors. The final prediction is determined by multiple decision trees rather than the individual trees (thus why it is called the ensemble learning algorithm), which can help overcome the disadvantage of over fitting of the single decision tree model. As we discuss later, due to its non-black-box nature, RF models can be investigated with partial dependence plots to demonstrate the relationships between variables and a variable's importance as a predictor. Such results can be useful in understanding the atmospheric process. It is therefore, useful to use ML techniques as a complement to conventional statistical analysis.

4.4.3 Weather normalisation using Machine learning

The random forest based weather normalisation technique was originally developed by Grange, Carslaw et al. (2018), where a set of explanatory variables (time variables and meteorological variables) from individual stations or cities were used to predict daily PM10 concentrations using a random forest machine learning model (to train the RF model).[10] To achieve weather

[10]Meteorological variables include temperature, relative humidity, wind speed, wind direction and pressure, time variables include date_unix (number of seconds since 1970-01-01, represents the trend in emissions), day_Julian (day of the year, represents seasonal variation), weekday (day of the week, represents weekly variation) and hour (hour of the day, represents hourly variation).

normalisation, the trained RF model from individual stations or cities was used to repeatedly sample and predict concentrations (for example 1000 times). When predicting each concentration, each time, a randomly selected set of explanatory variables (including both time and weather variables, but excluding the trend term date_unix) were chosen from the whole data period as inputs to give 1000 predictions of the specific time point. The 1000 predictions at the specific time point are then averaged using an arithmetic mean and hereafter referred to as the de-weathered concentration at that specific time point, i.e. the concentration level under average meteorological conditions (Grange, Carslaw et al., 2018). Due to Grange, Carslaw et al. (2018)'s inability to detect seasonal trends, Vu et al. (2019) modified the algorithm to include only the weather variables when randomly selecting the explanatory variables to predict the concentration at a specific time point. This allows researchers to detect the seasonal and diurnal variations after normalising for weather.

More recently, Shi et al. (2021) applied the methods of Grange, Carslaw et al. (2018) and Vu et al. (2019), to detect the impact of the COVID-19 lockdown on air quality in eleven global cities. In Shi et al. (2021), only the meteorological variables were normalised (similar to Vu et al. (2019)), and instead of re-sampling meteorological variables from a selected period (i.e., the 2-weeks before and after the selected time point) as in Vu et al. (2019), Shi et al. (2021) re-sampled from the whole data period (which is similar to Grange, Carslaw et al. (2018)). This method is more appropriate when trying to detect short term emission changes (such as the case of the COVID-19 lockdown (Shi et al., 2021)).

In our study, because the CEIP may have a sudden impact on air quality in a short term (i.e., lead to the short term change in emissions), we follow the weather normalisation technique used by Shi et al. (2021). Using the 'rmweather' R package developed by Grange and Carslaw (2019), we conduct the weather normalisation on $PM_{2.5}$, PM10, NO_2 and SO_2 for 35 Chinese cities separately.[11]

[11]The code used by (Shi et al., 2021) is available at https://$github.com/songnku/COVID-19-AQ$.

To illustrate, consider PM10 levels in Shijiazhuang city. First, we build up (train) the random forest model to predict hourly concentrations of PM10 in Shijiazhuang using data between 1st June 2014 and 30th May 2016 (a 2-year period). Our explanatory variables include meteorological variables (surface observed data: temperature, relative humidity, wind speed, wind direction; ERA5 Reanalysis data: boundary layer height, total cloud cover, surface net solar radiation, surface pressure, total precipitation; as well as back-trajectory clusters data) and time variables (date_unix, day_Julian, weekday and hour). We used 70% of the original data set to train the model and the remaining 30% was used to test the model. The Random Forest model performance is presented in Table 4.2.[12] By default, the 'rmweather' R package sets the total number of trees at 300 ($n\_trees = 300$) and the number of times to sample the whole data and then predict is also set at 300 ($n\_samples = 300$). The model was trained only on the observations that have non-missing wind speed data.

[Table 4.2 about here]

Second, as for normalising weather, only the ten meteorological variables were randomly selected as the new explanatory variables to obtain the predicted pollutant levels (using the previously trained RF model for that pollutant in the city). This process was then repeated 300 times to generate 300 predicted values for the pollutant. The final weather normalised value is obtained by calculating the mean of the 300 predictions. To be consistent, we run the same weather normalisation code on $PM_{2.5}$, PM10, NO_2 and SO_2 for all 35 cities (eight Hebei cities + 27 control cities) separately. This is in line with Abadie (2019) who suggests that the substantial volatility in the outcome variable should be removed in both the treatment and control units prior to applying the SCM. In the results section, we show the ASCM results using both observed and weather normalised concentrations of $PM_{2.5}$, PM10, NO_2 and SO_2.

[12]For simplicity, we only report the model performance for $PM_{2.5}$, PM10, NO_2 and SO_2 in Shijiazhuang city. The model performance for the other 34 cities is also available if requested.

4.4.4 The augmented synthetic control method (ASCM)

Although DID has often been used in policy evaluation studies there is a problem when the pre-parallel trend assumption can not be identified. To address such concerns, Abadie and Gardeazabal (2003), Abadie, Diamond and Hainmueller (2010) and Abadie, Diamond and Hainmueller (2015) developed the synthetic control method, which is regarded as "most important innovation in the policy evaluation literature in the last 15 years" (Athey and Imbens, 2017). The key idea behind the SCM is the "comparative case study", which uses a set of control units to construct a suitable comparison group. Instead of using a single control unit or using the simple average of the whole control group (like the DID does), the SCM uses the weighted average of the control group which provides a more suitable comparison group (therefore addresses the problem of non pre-parallel trends). By using a weighted average combination of a set of control units, the SCM is able to closely approximate the movement of outcome variables for the treatment unit before policy intervention (so called the "synthetic unit"), and then after the intervention the difference between the changes in the treatment unit and the synthetic unit can be regarded as the causal impact of the policy intervention (Abadie, Diamond and Hainmueller, 2010; Abadie, Diamond and Hainmueller, 2015).

To employ the standard synthetic control method (SCM), we need to ensure that the treatment and control groups are relative similar in terms of the factors that affect emission levels of $PM_{2.5}$, PM10, NO_2 and SO_2 (for example, the level of economic development, population, energy consumption, etc (Isaksen, 2020)). However, such city level data is only available at monthly or quarterly frequencies. Since Hebei was inspected in January 2016, we are interested in the direct and immediate effect. However, getting data at the weekly or even monthly data in some cases means that we are not able to use the standard SCM approach.

Extending the standard SCM, Ben-Michael, Feller and Rothstein (2021) propose the Augmented Synthetic Control Method (ASCM), which focuses on optimizing the pre-treatment fit and minimizing the imbalance of the outcomes prior to treatment between the treatment and weighted control units. This is feasible even when only the outcome variable is available for researchers. Moreover, Abadie, Diamond and Hainmueller (2010) suggest that researchers should not use SCM if the outcome variable of the synthetic unit does not closely match the outcome variable of the treatment prior to the intervention, which might lead to a biased estimation of the intervention effect. The ASCM extends the SCM to settings where pre-policy fit is not possible, and uses an outcome model to estimate the bias due to the imperfect pre-intervention fit, and then directly corrects the original SCM estimate using the outcome model. Thus, the ASCM estimator has the feature of being a "doubly robust estimation".

In this paper, to estimate the effectiveness of the Hebei inspection on local air quality, we use the canonical panel data setting in Ben-Michael, Feller and Rothstein (2021). Let $i = 1, \ldots, N$ refer to cities and $t = 1, \ldots, T$ refers to the time period by week (T_0 denotes the inspection time, i.e. the week start on 28 December 2015), W_i denotes the treatment status of city i at $T_0 < T$ where cities with W_i=0 never receive an inspection and W_1=1 refers to only one unit (Hebei) that does receive an inspection (assume Hebei is the first city), let $Y_{it}(0)$ and $Y_{it}(1)$ refer to the potential outcome (weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2) of city i in week t within the control group and treatment group (Hebei), separately.[13] We have the following observed outcomes:

[13]For simplicity, for the main results, we calculate the average weather normalised weekly $PM_{2.5}$, PM10, NO_2 and SO_2 data in eight Hebei cities and treat it as one city - "Hebei". We also show how the policy effected the eight cities separately.

$$Y_{it} = \begin{cases} Y_{it}(0) \quad if \quad W_i = 0 \quad or \quad t \leq T_0 \\ \\ Y_{it}(1) \quad if \quad W_i = 1 \quad and \quad t > T_0 \end{cases} \tag{4.3}$$

If we assume that the potential outcome of the control units are generated by a fixed component m_{it} and additive noise ϵ_{it}:

$$Y_{it}(0) = m_{it} + \epsilon_{it} \tag{4.4}$$

Then the estimated inspection effect on Hebei is given by:

$$\tau = \tau_{1t} = Y_{1t}(1) - Y_{1t}(0) \tag{4.5}$$

where $Y_{it}(1) = Y_{it}(0) + \tau_{it}$ is the inspected potential outcome and τ_{it} is the key estimated treatment effect. Our target is to get the estimation of the $Y_{1t}(0)$, following Ben-Michael, Feller and Rothstein (2021), where the fixed component is estimated via a ridge-regularized linear mode, and the Ridge ASCM estimator is given by:

$$\hat{Y}_{1T}^{aug}(0) = \sum_{W_i=0} \hat{\gamma}_i^{scm} Y_{iT} + (X_1 - \sum_{W_1=0} \hat{\gamma}_i^{scm} X_i)\hat{\eta}^{ridge} \tag{4.6}$$

Hebei was inspected in January 2016, to allow a relatively longer pre-treatment period for the Ridge ASCM to construct the synthetic Hebei, we use the 1-year pre inspection period (51-week) and 5-month post period (23-week). We set the inspection start week as the week starting on the date 28 December 2015. According to Ben-Michael, Feller and Rothstein (2021), the hyper-parameter λ^{ridge} in the ridge regression controls the bias-variance trade-off between a better pre-fit and lower estimation variance, and choosing λ is thus important in practise. In our paper, we use the default setting that selects λ^{ridge} through a cross-validation procedure for all pollutants in the Ridge ASCM analysis. We drop the data after June 2016 since the next inspection round for eight other provinces started around 15 July 2016.

4.5 Results

4.5.1 DID and parallel trend test results

First, we estimate the effects of the CEIP on air quality levels in Hebei using a DID model. The results are shown in Table 4.3. Columns 1, 2, 3 and 4 show the results for $PM_{2.5}$, PM10, NO_2 and SO_2, respectively. Rows (a) and (b) show the results without weather controls and rows (c) and (d) show the results that include weather controls. Our main focus of Table 4.3 is row (c) since it uses the observed pollutants and includes weather control variables, which is the same as equation 1 which is the equation that is usually applied in a standard DID estimation. Given both city, and week, fixed effects are included, the estimated coefficients in columns 1, 2 and 4 (row (c)) confirm the effectiveness of the CEIP, suggesting that, compared with the control group cities (that had no CEIP inspection, i.e., Jan to Dec 2015), the weekly $PM_{2.5}$, PM10 and SO_2 concentrations in Hebei were significantly reduced by 8.405 $\mu g/m^3$, 17.28 $\mu g/m^3$ and 2.765 $\mu g/m^3$ during Jan to May 2016, respectively. However, we also found that NO_2 levels increased by 2.286 $\mu g/m^3$. For comparison, row (b) shows the results of a DID estimation on weather normalised concentrations without weather controls. It can be seen that the estimated CEIP effects differs a lot depending on whether we use observed or weather normalised pollutant concentrations, suggesting that the impact of weather conditions on pollution concentration is complex and can lead to different estimation results.

[Table 4.3 about here]

The underlying assumption for an unbiased estimation of the inspection effects rely on the pre-parallel trend assumption, i.e. both control cities and treatment cities should have followed parallel trend in terms of the the trajectory of the air quality movement prior to the the policy implementation, this is to ensure that there is no systematic difference of the air pollution levels between treatment and control groups prior inspection, and the difference after policy implementation can be attributable to the policy itself only. Figure 4.1 plots the results of the pre-parallel trend test on observed $PM_{2.5}$, PM10, NO_2 and SO_2. More specifically, Figure 4.1 (a), (b), (c) and (d) plot the point estimate results of the estimated coefficients from equation (2), the vertical line refers to the policy start time.[14] Figure 4.1 implies the failure of the pre-parallel trend test for $PM_{2.5}$, PM10 and NO_2, suggesting there already exist significant and systematic difference of the air pollution levels ($PM_{2.5}$, PM10 and NO_2) between treatment and control groups prior inspection, and therefore the policy effect we got from Table 4.3 (row (c) and columns 1, 2 and 3) is likely to be biased. The ASCM we adopt in later part can potentially address such problem when the pre-parallel trend assumption is not feasible.[15]

[14]To ensure consistency with the later ASCM results, we set the same policy start time (the week starting from 28 December 2015) and used the same control group (full sample control group, 27 cities) in DID (and pre-trend test) as in the later ASCM analysis. Figure C.3 in the appendix plots the pre-trend test results of using every 1-month window, which also fail to support the pre-parallel trend assumption.

[15]We also show the results of pre-trend test using weather normalised value in appendix Figure C.4, the results show that even though taking the full control of weather impact, there still does not exist the parallel trend of weather normalised pollutants between the treated and control units prior the inspection. This

[Figure 4.1 about here]

4.5.2 Weather normalised results using machine learning

Before we get to the main results, we first illustrate the results of the random forest and weather normalisation process. Figure 4.2 plots the results of (machine-learning based) weather normalisation on $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei between January 2015 and May 2016. The blue dotted line refers to the weekly observed value and the red solid line denotes the weather normalised value, i.e. the pure trend in emissions without the confounding meteorological influences.[16] As shown in Figure 4.2, the observed weekly pollutants are highly volatile and there exists large differences between the observed and weather normalised air pollution concentrations. Figure 4.2 shows the impact of weather on pollution readings but also the benefit from weather normalising given week to week readings of actual data can be misleading making pollution levels appear better or worse than the underlying values.

[Figure 4.2 about here]

Figure 4.3 presents the broad picture in terms of the emissions of our four different pollutants by presenting the weekly concentrations of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei and the 27 control cities between January 2015 and May 2016. The weather normalised concentrations hereafter refer to the emission levels without the confounding weather conditions. After removing the meteorological impact, the concentrations of $PM_{2.5}$,

further demonstrate the importance of using alternative method like (A)SCM when pre-trend is not feasible in DID.

[16]Throughout the paper, "WN/wn" refers to the weather normalised pollution concentration value, for example, "$PM_{2.5}$wn", "PM10wn", "NO_2wn" and "SO_2wn" denote weather normalised concentration of $PM_{2.5}$, PM10, NO_2 and SO_2, respectively.

PM10, NO_2 and SO_2 in Hebei are relatively high compared to the cities in the control group. This raises the concern that simply using the (average of) the 27 cities as the control group may not provide a comparable control group to investigate the policy effect, since the differences of $PM_{2.5}$, PM10, NO_2 and SO_2 prior to the policy between two groups are, in some cases, or an order of magnitude lower (Abadie, Diamond and Hainmueller, 2010).

[Figure 4.3 about here]

Following Grange, Carslaw et al. (2018), the random forest model has the advantage of allowing us to observe the importance of each variable in predicting the outcome variable and the underlying relationship between variables. Figure 4.4 plots the result of variable importance for $PM_{2.5}$, PM10, NO_2 and SO_2 RF models in Shijiazhuang (using the same example as Table 4.2). As shown, the seasonal variation (day_julian) plays the most important role in predicting hourly $PM_{2.5}$, PM10, NO_2 and SO_2 concentrations in Shijiazhuang, while meteorological conditions such as temperature (temp), surface pressure (pressure) and relative humidity (RH) are also important explanatory predictors, while wind speed and wind direction have only limited relevance.

[Figure 4.4 about here]

Figures C.5, C.6, C.7 and C.8 in the appendix plot the partial dependences for $PM_{2.5}$, PM10, NO_2 and SO_2 RF model for Shijiazhuang, respectively. The vertical axes of each figure denotes the dependence of hourly pollution levels for each variable assuming all other explanatory variables are fixed at the average level (Grange, Carslaw et al., 2018). As shown, the concentration of the four pollutants depends on different meteorological variables. The plots of day_julian in $PM_{2.5}$, PM10, NO_2 and SO_2 suggest that the four pollutants have higher concentrations in the winter and lower in the summer, which is due to the higher emission levels and atmospheric stability during the winter. For example, PM10 is negatively related to surface pressure, and positively related to relative humidity although the majority of the weather variables have a non-linear relationship with $PM_{2.5}$, PM10, NO_2 and SO_2 concentrations. These non-linearities demonstrate one of the advantages of the RF ML model in that it can deal with weather conditions compared to the DID model approach, where the relationships are often assumed to be linear (in practise researchers had to hand-pick different functional forms for different weather variables). It is also important to note that the dependence of different air pollutants on different weather variables are different in different cities. Hence the RF ML model arguably provides a more flexible (flexible in the sense of allowing for non-linear relationship and regional diversity) way to account for the impact of weather conditions on different pollutants in different cities that can be missed in standard statistical analysis.

4.5.3 The Impact of the Hebei inspection on local air pollution using a Ridge ASCM approach

Our main results are presented in Figures 4.5 and 4.6 plot and show the results of a Ridge ASCM regression using weather normalised concentrations of $PM_{2.5}$ (Figure 4.5, row 1), PM10 (Figure 4.5, row 3) , NO_2 (Figure 4.6, row 1) and SO_2 (Figure 4.6 row 3), respectively. We include a 1-year pre-inspection period (51-weeks) and a 5-month post period (22-weeks). The vertical line is the date when the CEIP started in Hebei (the week starting the 28th December 2015).[17] For comparison, we use the same code, time period and cities to run the Ridge ASCM model on observed pollution data and the results are shown in Figure 4.5, row 2 (for $PM_{2.5}$), Figure 4.5, row 4 (for PM10), Figure 4.6, row 2 (for NO_2) and Figure 4.6, row 4 (SO_2).

[17]For the two weeks starting on the 13th April and the 20th April 2015, the observed data for the four pollutants in all 35 cities were missing. For these weeks we used cubic spline interpolation to fill in the missing data.

[Figure 4.5 about here]

[Figure 4.6 about here]

The interpretation of the results is that the first columns in Figures 4.5 and 4.6 is the difference in the weather normalised values of the four pollutants between Hebei and what we call in the paper Synthetic Hebei. For weather normalised pollutants, the differences were close to zero prior to the inspection (which occured at the dotted line) after which point the values began to diverge. Our results show that the CEIP lead to significant reductions (significant at the 95% CI using the Jackknife+ procedure to generate the CIs) of weather normalised $PM_{2.5}$, PM10 and SO_2 emission levels for the first 2 to 2.5 months after the inspection in Hebei as a whole. The significant effects gradually disappear after late March 2016 and became insignificant up until the end of our sample of the May 2016.

When we look at the observed values (not weather normalised) we still tend to see a fall in pollution values but the impact of the inspection on all four pollutants is insignificant (the CI are above zero at all points in time). It should be noted however that the pre-inspection fit is not as good for the observed data. The results demonstrate the importance of weather normalisation when looking at trends in air quality. In our case the real policy impact is masked by weather conditions at the time and the policy would be recorded as less effective than it actually was.

The second columns of Figures 4.5 and 4.6 plot the trends in weather normalised and observed levels for our four pollutants for Hebei (the red dotted line) and synthetic Hebei (the blue triangle line). For the weather normalised values, the synthetic Hebei simulates the trajectory of the four pollutants very closely prior to the inspection, but far less so afterwards (where the blue triangle line captures when predictions for pollution levels assuming there had not been an inspection).

To summarise our results, Figures 4.5 and 4.6 demonstrate that the CEIP led to significant and substantial reductions in weather normalised $PM_{2.5}$, PM10 and SO_2 in Hebei, however the effect only lasted for around 2 to 3 months. No significant reduction was found for weather normalised NO_2 which as we will discuss later is reassuring as NO_2 is principally associated with transport which would not have been impacted by the inspection teams (and in a sense can be thought of as a placebo test).

In Figures 4.5 and 4.6, the pollution levels were calculated using the arithmetic mean of the eight Hebei cities that were subject to inspection. As a robustness check we use three alternative methods to calculate the Hebei average and the results are shown in Table C.2 of the Appendix. Columns 1 to 4 in Table C.2 illustrate the impact of the CEIP on the four pollutants using the arithmetic, population weighted, GDP weighted and industrial output weighted means, respectively. The average ATT estimation of the CEIP impact using alternative weighting methods is fairly similar, suggesting that our main results (column 1, Table C.1) are not sensitive to how the Hebei average is calculated.[18]

The next stage of the analysis is to examine the differences across cities within Hebei. Figure 4.7 plots the Ridge ASCM results on weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 for each of the eight Hebei cities separately (the results with confidence intervals are shown in Figures C.9 and C.10 in the Appendix). The results suggest that the CEIP led to significant and substantial reductions in weather normalised $PM_{2.5}$ and PM10 concentrations in Shijiazhuang, Cangzhou, Xingtai, Baoding, Tangshan and Qinhuangdao. For Chengde a significant reduction was detected in PM10wn, but not for $PM_{2.5}$wn. The impact of the CEIP on weather normalised NO_2 and SO_2 differs across the eight cities. The inspection lead to a

[18]We also test for the pre-parallel trends for the four pollutants using alternative weighting methods and similar pre-parallel trends to those presented in Figure 4.1 are found with no pre-parallel trends detected. These results further demonstrate the importance of using ASCM when the pre-parallel trend assumption does not hold.

reduction in SO_2 in six out of eight cities, although there were significant increases in SO_2 in Xingtai and Tangshan. The city level results appear to suggest that the CEIP teams targeted the dirtier cities and were effective in reducing air pollution in those cities with the greatest air pollution problems, especially Baoding, Shijiazhuang, Xingtai and Tangshan (which were the relatively polluted cities in Hebei during the study period, according to the Figure C.11 in the appendix).[19]

[Figure 4.7 about here]

4.5.4 Placebo tests

Following Abadie, Diamond and Hainmueller (2015) and Cole, Elliott and Liu (2020), we conduct three different placebo tests. In addition to in-time and in-place tests we also re-estimate our results using a number of different control groups.

In-time placebo test

For the in-time placebo test we backdate the real inspection time six months and assume instead that Hebei was inspected in July 2015 instead of January 2016. We use the same code and city control group. Figure 4.8 plots the in-time placebo results of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2. As shown, there are no significant results for any of the four pollutants.[20]

[Figure 4.8 about here]

[19] Our city level results need to be treated with some caution since we do not have detailed information on the size of the teams that visited each city or the length of time that they stayed there.

[20] We also estimated in-time placebo tests backdating the real inspection time by 5/4/3/2/1 months, separately. In each case, there were no significant effects.

In-place placebo test

For the in-place placebo test, we assume that each of the city in our control groups experienced the inspection policy in January 2016, we run the exact same code on the 27 control cities one by one, and the differences between each treatment city and the synthetic treatment city are plotted in Figure 4.9 (weather normalised $PM_{2.5}$, PM10 and SO_2).

The left column in Figure 4.9 plots the results of using all cities in the control group for $PM_{2.5}$, PM10 and SO_2, respectively. The red line is our main Hebei result included for comparison. The left column indicates that the Hebei pollution levels are lower than the average (but less so for SO_2). However, once we exclude those cities that have a poor pre-inspection fit between their pollution levels and the synthetic group, we have the right column of Figure 4.9. The results shows the inspection effect of weather normalised $PM_{2.5}$, PM10 and SO_2 in Hebei was significantly different to the other 27 cities used for the in place test. It is reassuring that the time that Hebei stands out is the same period where we find significant results in our main results (i.e., between 1 January to around March 2016). Our in place tests are consistent with the notion that the impact of the inspection in Hebei on emissions was significant and not driven by chance (Abadie, Diamond and Hainmueller, 2015).

[Figure 4.9 about here]

Alternative control groups

To obtain unbiased estimations of the (Ridge) ASCM (as well as DID) we use the Stable Unit Treatment Variable Assumption (SUTVA), which assumes that all the cities in the control group (door pool) should not be affected by the inspection in Hebei (Imbens and Rubin, 2015; Ben-Michael, Feller and Rothstein, 2021). In this study, there are two possible violations of the SUTVA assumption. The first is that air pollution from one city can, under certain wind conditions, impact pollution in other cities, making it difficult to capture the true effect of a city level inspection. However, our study mitigates against this concern by the inclusion of back-trajectory clusters to account for the air pollution transportation as part of the weather normalisation machine learning procedure. The second concern is how cities respond to being inspected. For example, when cities in Hebei were inspected, local officials and firms in other heavily polluted cities in other provinces may respond to the nes of the Hebei inspection by increasing or decreasing their own emissions that would violate the SUTVA assumption, i.e., cities in the control group might have been affected by the Hebei inspection. Although we believe such actions to be unlikely we re-estimate our main results using a number of different control groups on the understand that if we find similar results using different cities as the control group, the violation of the SUTVA assumption does not undermine our main results. We re-estimate our results using four different control groups.

First, in addition to the “Beijing-Tianjin-Hebei (BTH)“ region, air pollution is also a pressing problem in the “Northeastern three provinces (Heilongjiang-Jilin-Liaoning” region. It is possible that cities in this region may have taken some emission reduction activities once they learned that Hebei was to be inspected inspected although they would not have known if the Hebei inspection would be rolled out across China or when. However, for our first alternative control group we exclude the cities in “Northeastern three provinces” from the original full control group that we call SCG1.

Second, we create a control group, SCG2, that excludes southern Chinese cities (these are cities located south of the “Qin mountain and Huai river line”) from the original full control group. This choice is motivated by the relatively high pollution levels in northern Chinese cities (due in part to the winter heating policy (known as the “Huai River” policy) (Ebenstein, Fan, Greenstone, He and Zhou, 2017). Therefore, we drop ten southern cities from the full control group to get a group of cities that are more “similar” to Hebei in terms of air pollution levels.[21] A third alternative control group, SCG3, excludes those cities that have relatively clean air quality (cities in Yunnan, Guizhou, Guangxi and Fujian) from the control group. [22]

Figure 4.10 presents the results from the three alternative control groups using weather normalised PM10, $PM_{2.5}$ and SO_2, respectively (SFull is a reminder of the main results using the full control group). All control groups simulate the movement of weather normalised $PM_{2.5}$ and SO_2 with Hebei before inspection, and show similar effects of the inspection suggesting that our main results of the causal impacts of Hebei inspection on weather normalised $PM_{2.5}$ and SO_2 are not sensitive to the choice of control group.

[Figure 4.10 about here]

4.6 A comparison of DID and ML+ASCM

We now compare the results of our main results using ML + Ridge ASCM with a more traditional DID approach. The results are show in Table 4.4. Rows a1 and a2 in Table 4.4 show the maximum change of air quality (in levels and percent) in Hebei due to the CEIP. Our main results (ML+Ridge ASCM) who that the CEIP led to a significant reduction in $PM_{2.5}$, PM10, NO_2 and SO_2 by as much as 20.81%, 22.61%, 4.32% and 8.59%, respectively. Row a3 in Table 4.4 shows the average ATT estimate from the Ridge ASCM estimation on weather normalised values, i.e. the average treatment effect of the inspection on air quality in Hebei after the inspection. On average, the CEIP led to significant reduction of $PM_{2.5}$wn

[21] Changsha, Wuhan, Hefei and Kunming are kept in SCG2 because they have partial access to the centralised winter heating programs.

[22] See Table C.3 in the Appendix for the full list of cities in each group. We use identical code and the same time period for the analysis.

and PM10wn by 4.69 $\mu g/m^3$ and 20.80 $\mu g/m^3$ during the 5-month period after inspection. It also led to the reduction of weather normalised NO_2 and SO_2 by 1.11 $\mu g/m^3$ and 1.26 $\mu g/m^3$, although the significance is limited. Row a4 in Table 4.4 represents the corresponding average change (in %, derived from row a3).

Rows b (b1 and b2) and c (c1 and c2) in Table 4.4 present the estimated coefficient of a DID model using observed values (b1, b2) and weather normalised values (c1, c2), respectively (all include weather controls). The results of DID using observed values suggest that, compared with control cities (that were not inspected during January 2016), the CEIP significantly reduced $PM_{2.5}$, PM10 and SO_2 in Hebei by 8.41 $\mu g/m^3$, 17.29 $\mu g/m^3$ and 2.77 $\mu g/m^3$ in Hebei, respectively. However, NO_2 increased significantly 2.29 $\mu g/m^3$. The DID estimation for $PM_{2.5}$, PM10 and NO_2 may be biased for two reasons. First, bias introduced by not having a good parallel trend prior the policy implementation between treated and control units (as shown in Figure 4.1 (a), (b) and (c)). Second, bias due to an inappropriate control for weather conditions (Equation 1).

To investigate further, Rows (c1 and c2) in Table 4.4 present the DID results using weather normalised values (plus weather controls). The pre-parallel trends still do not hold when using weather normalised values (as shown in Figure C.4 of the Appendix). This suggests that the DID estimates may still be biased even though the impact of weather conditions was properly controlled. This means any existing bias comes from the parallel trends assumption.

Our comparison suggests that the lack of parallel pre-trends and the failure to control appropriately for meteorological conditions, the traditional DID method (Equation 2) using observed air pollutant concentrations (plus weather controls and fixed effects) tends to over-estimate the CEIP effect for $PM_{2.5}$ and SO_2 (while underestimate policy effect for PM10), and shows an increase in NO_2 (5.1% increase vs. 2.48% reduction).

[Table 4.4 about here]

4.7 The benefits of reducing pollution

After investigating the causal impact of the CEIP on air quality, we conduct a simple back of the envelope exercise to illustrate the potential health and economic benefits from reducing pollution in Chinese cities. We focus on PM10 for illustrative purposes. Table 4.5 describes the health impacts from our 4 pollutants.

[Table 4.5 about here]

Our main results from Figures 4.5 and 4.6 and Table 4.4 suggest that the CEIP led to a significant reduction (around 20.81 μg/m^3) of weather normalised PM10 for around 3 months in Hebei. For the purpose of this study we conservatively assume that on average Hebei experienced 15 μg/m^3 reduction of PM10 emissions for around 3 months before returning to previous levels. We also assume similar effects of the CEIP inspection for each round that occurred throughout 2016 and 2017 for China as a whole (a 15 μg/m^3 PM10 reduction that lasted for 3-months.

We select four key studies that have investigated the PM10-mortality association in China presents in Table 4.6. We then collect the data of annual population and mortality in Hebei and China in 2015 and calculate the number of lives saved in Hebei and China due to the falling PM10 concentrations resulting from the CEIP. Table 4.7 gives an idea of the potential lives that can be saved from reducing PM10 and suggests that the reduction of PM10 due to the CEIP would save between 565 and 1,048 lives in Hebei and between 13,195 amd 23,824 lives in China as a whole.

[Table 4.6 about here]
Table 4.7 about here

It is important to emphasize that our estimation for the potential lives saved is highly speculative. For example, for simplicity we only focus on PM10 (we could do a similar exercise for $PM_{2.5}$) although we acknowledge that there is uncertainty regarding the exposure-response coefficients between a single pollutant and multi pollutant model. Second, we assume the PM10 mortality association for a change in 15 $\mu g/m^3$ is simply 1.5 times of that for 10 $\mu g/m^3$. Third, the inspection is estimated to cause air pollution reductions, but this might be at the cost of shutting down factories, laying off workers and reducing employment which may have a negative impact on health and mortality that we are not capturing.

4.8 Discussion and Conclusions

This paper provides an ex post empirical study to evaluate the effectiveness of the CEIP on air quality. More specifically, We employ a combination of weather normalisation techniques and an ASCM model to estimate the causal impact of the CEIP on air quality levels in Hebei province. We use hourly air pollution and meteorological data in 35 Chinese cities to decouple the effects of meteorology from observed pollutant concentrations, and then run a Ridge ASCM model using the weather normalised concentrations. Weather normalisation is important as it allows us to adjust the data for the fact that pollutant concentrations observed from surface monitoring stations are co-influenced by emission levels and meteorological conditions the latter making it a challenge to evaluate the efficacy of policy interventions. The weather normalisation method provides us with a reliable way to account for the effect of local weather conditions on observed concentrations so we can measure the “real” change in emission levels from by human activities (Grange and Carslaw, 2019; Vu et al., 2019).

In addition to the Ridge ASCM model results we also present results from a traditional

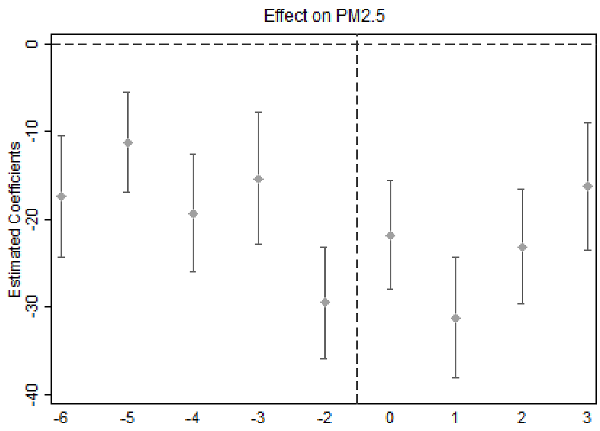
DID approach. The problem with this approach is that it failed the trend test suggest that the DID results may be biased as the trends in air quality between treatment and control group before the intervention were not close enough. A comparison between the DID and machine learning results demonstrates that features of ML (e.g., the flexible functional form, the non-black-box, the non-parametric and non-linearity of data structure) in air quality analysis provide a useful tool for use in other air pollution studies.

Our baseline results from the Ridge ASCM model using weather normalised pollution concentrations suggest that, the CEIP was effective in reducing air pollution and led to significant reductions of emissions (i.e., weather normalised) of $PM_{2.5}$, PM10 and SO_2 by 4.69 $\mu g/m^3$ (5.94%), 20.80 $\mu g/m^3$ (13.25%) and 1.26 $\mu g/m^3$ (2.62%) in Hebei as a whole. However, the three pollutants rebounded back to their normal levels in a relatively short time (around three months). When we looked at each city in Hebei separately, we found that the policy was more effective in reducing PM10 and $PM_{2.5}$ in heavily polluted cities such as Shijiazhuang and Baoding.

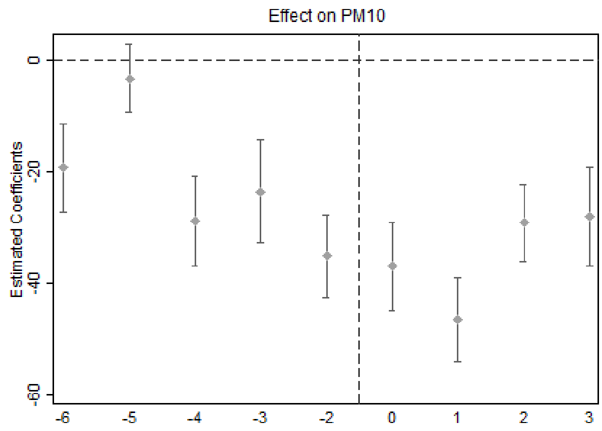
A comparison between DID and our ASCM approaches showed that the standard DID model tended to overestimate the policy effect on $PM_{2.5}$ and SO_2 (and underestimate the policy effect on PM10) with bias from both the failure of the pre-parallel trend assumption and failing to appropriately control for meteorological conditions (or a combined effect of the two). We argue that our method can provide an alternative approach for investigating the effectiveness of air pollution control policies, especially when there are no parallel trends.

One of the limitations of our study is that we only focus on the first round of CEIP that took place in Hebei province in January 2016, whereas the whole China was inspected throughout 2016 and 2017. An extension to this research would be to extend the period and examine further rounds of inspections. However, it then becoming more challenging to find appropriate control groups.

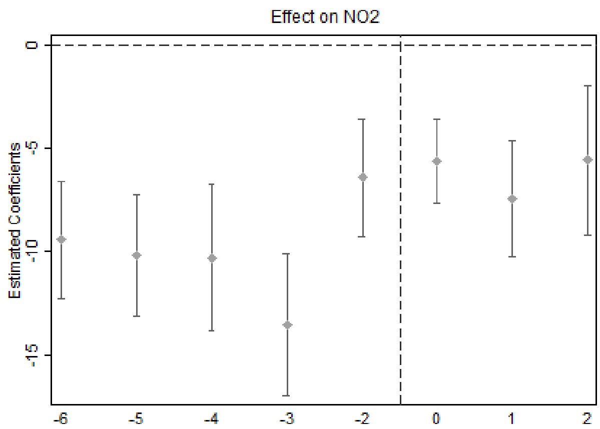
Chapter 4 Figures and Tables



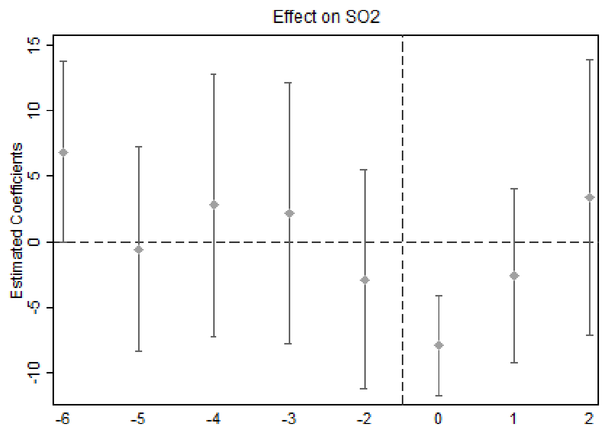
(a) $PM_{2.5}$ (every 2-month window)



(b) PM10 (every 2-month window)



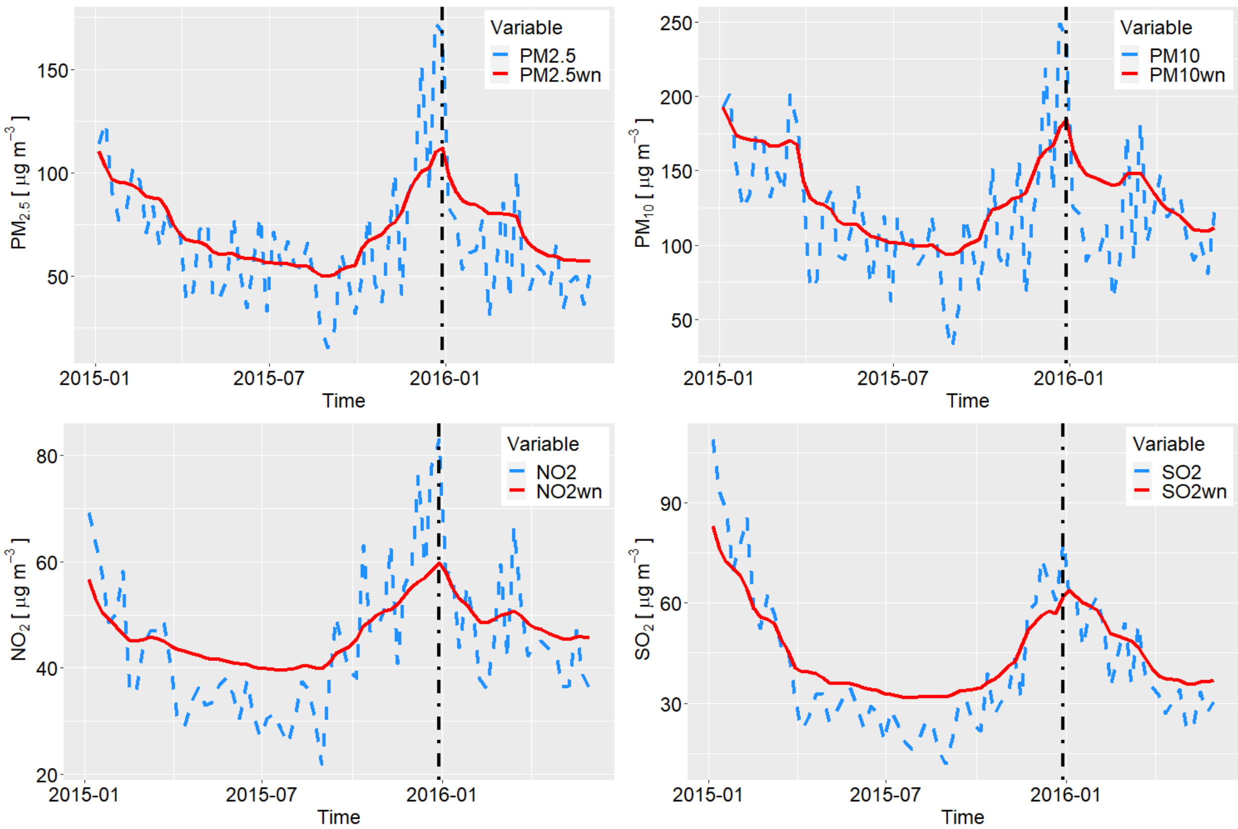
(c) NO_2 (every 2-month window)



(d) SO_2 (every 2-month window)

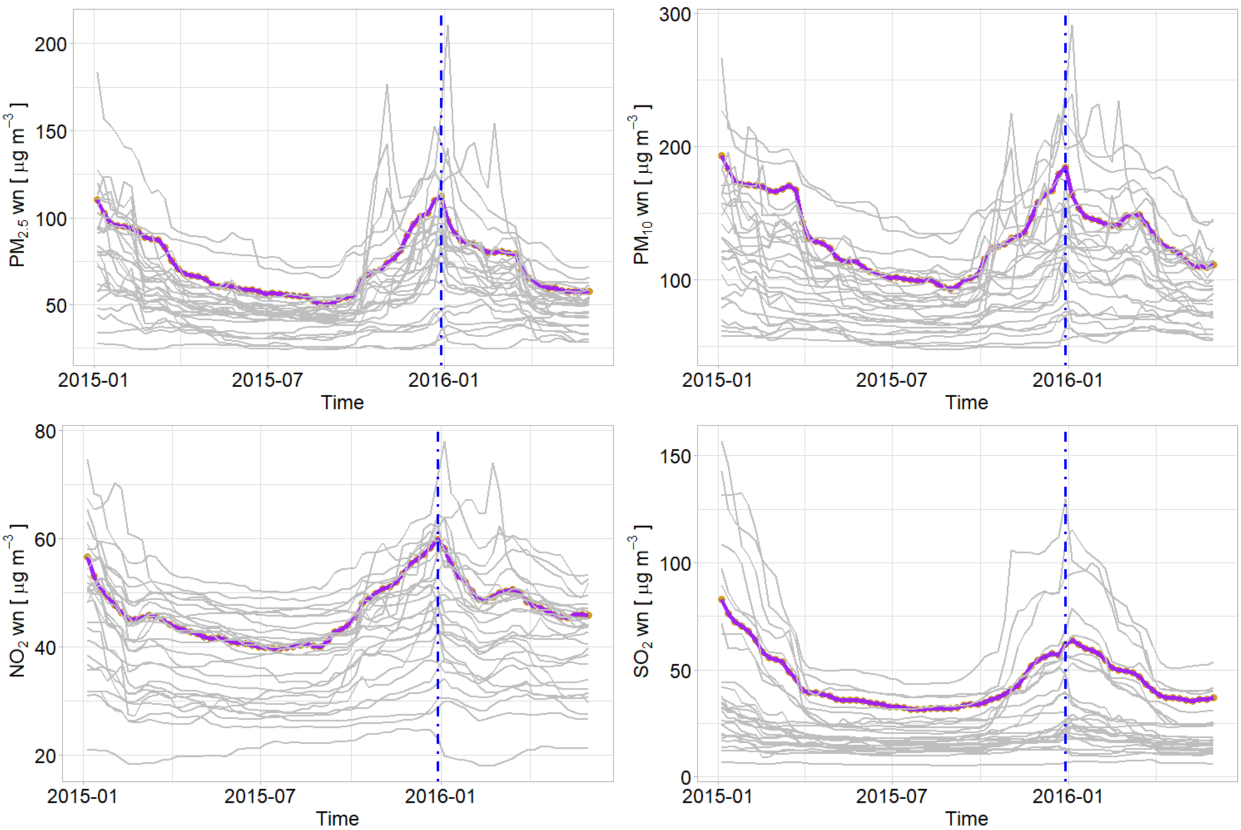
Note: The error bars refer to the 90% confidence intervals of the estimated coefficients.

Figure 4.1: The plot of pre-inspection parallel trend tests



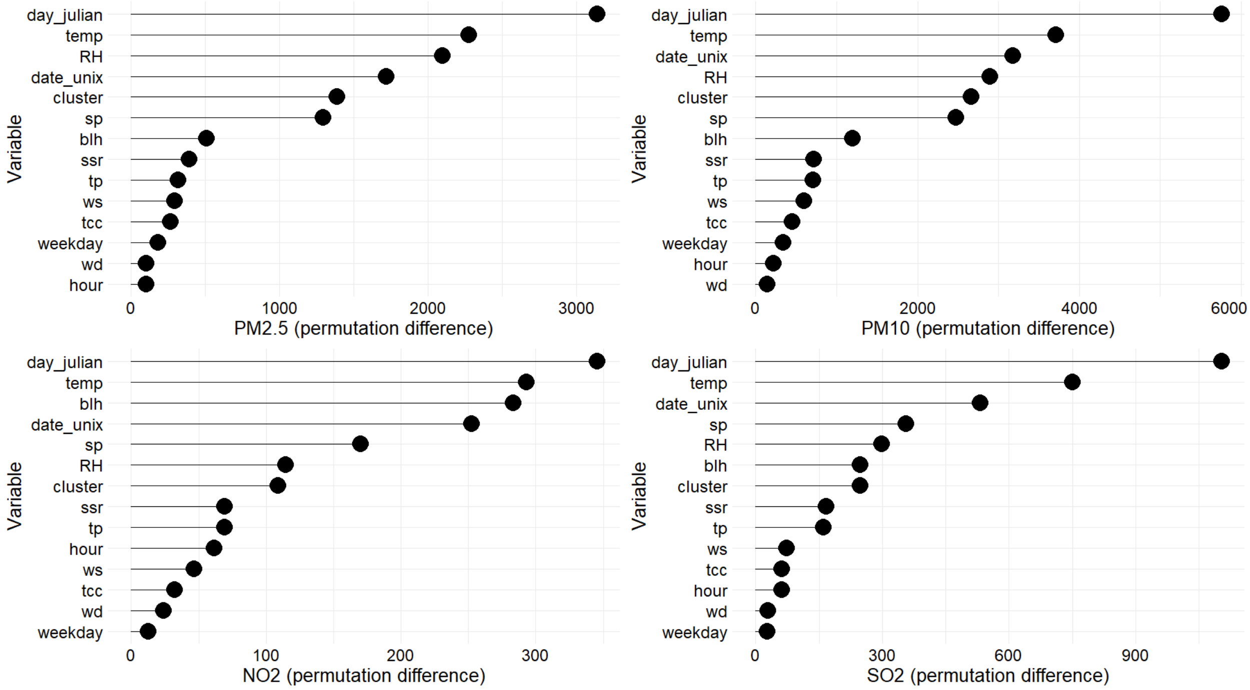
Note: The vertical line denotes 28 December 2015.

Figure 4.2: The comparison of weekly observed and weather normalised concentrations of $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei between January 2015 to June 2016.



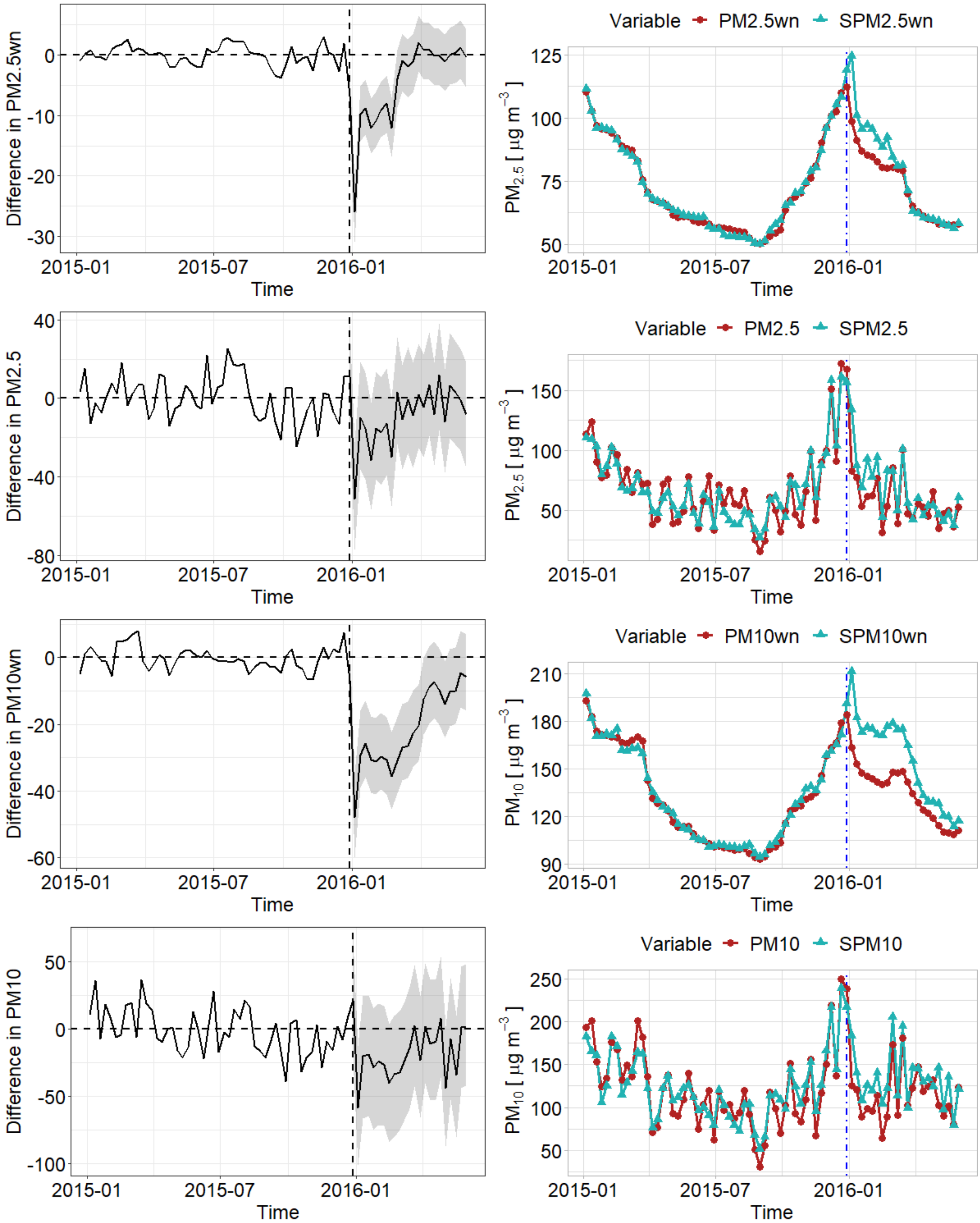
Note: Hebei is denoted by the purple line, the vertical line denotes 28 December 2015.

Figure 4.3: The weekly concentration of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei and 27 control cities between January 2015 to June 2016.



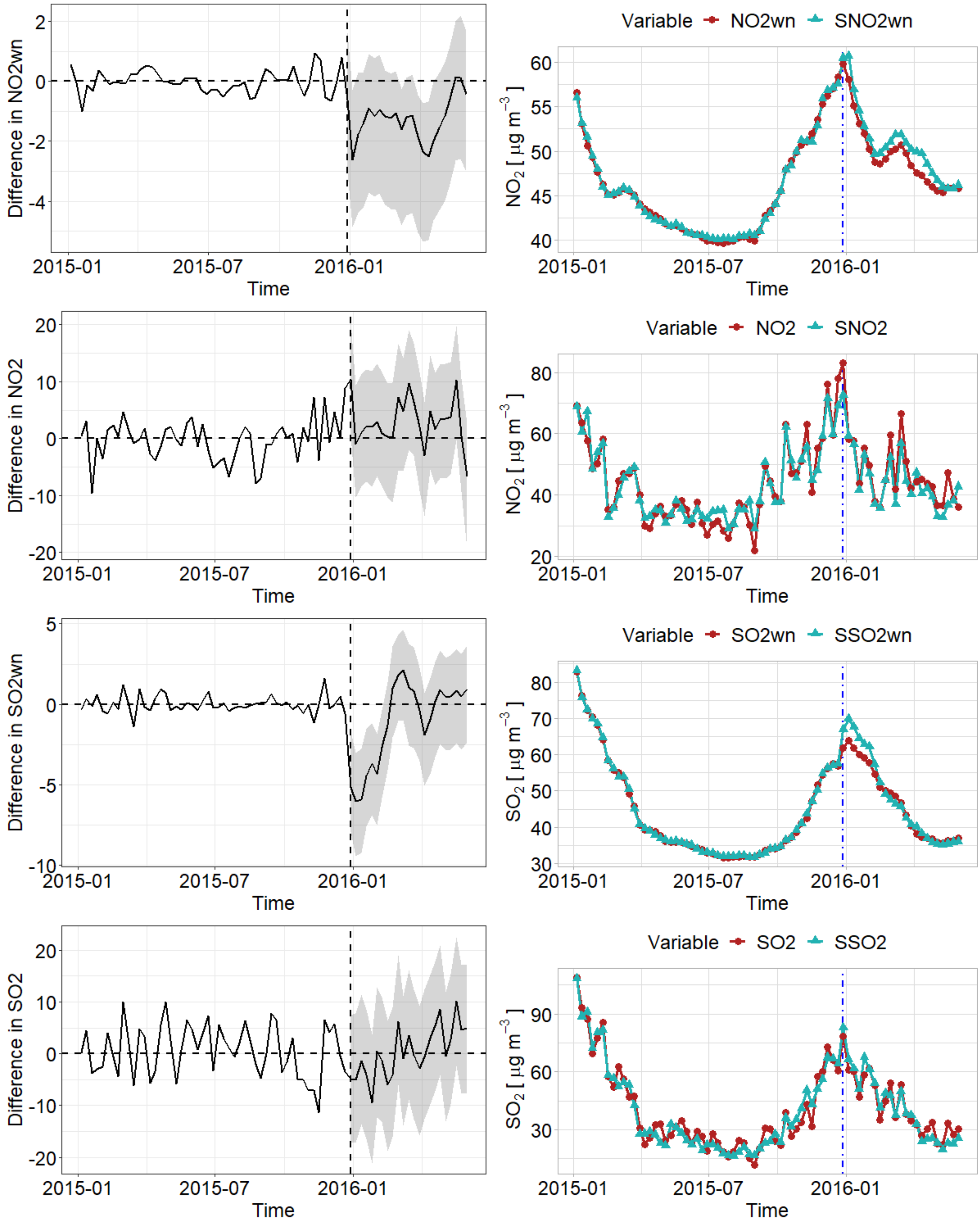
Note: "day_julian" refers to day of the year (represents the seasonal variation), "date_unix" refers to the Number of seconds since 1970-01-01 (represents the trend in pollutant emissions), "weekday" refers to Day of the week (represents the weekly variation), "hour" refers to Hour of the day (represents the hourly variation), "temp" refers to Temperature (degrees Celsius), "wd" refers to Wind Direction (m/s), "ws" refers to Wind Speed (in degrees, 90 is from the east), "RH" refers to Relative humidity (%), "pressure" refers to Atmospheric pressure (millibars).

Figure 4.4: The variable importance of $PM_{2.5}$, PM10, NO_2 and SO_2 random forest model in Shijiazhuang



Note: "PM2.5wn"refers to the weather normalised value of $PM_{2.5}$ whereas "PM2.5" refers to the observed value. The shaded area indicates the point-wise 95% confidence interval using Jackknife+ procedure

Figure 4.5: Ridge ASCM results on weather normalised and observed $PM_{2.5}$ and PM10 in Hebei.

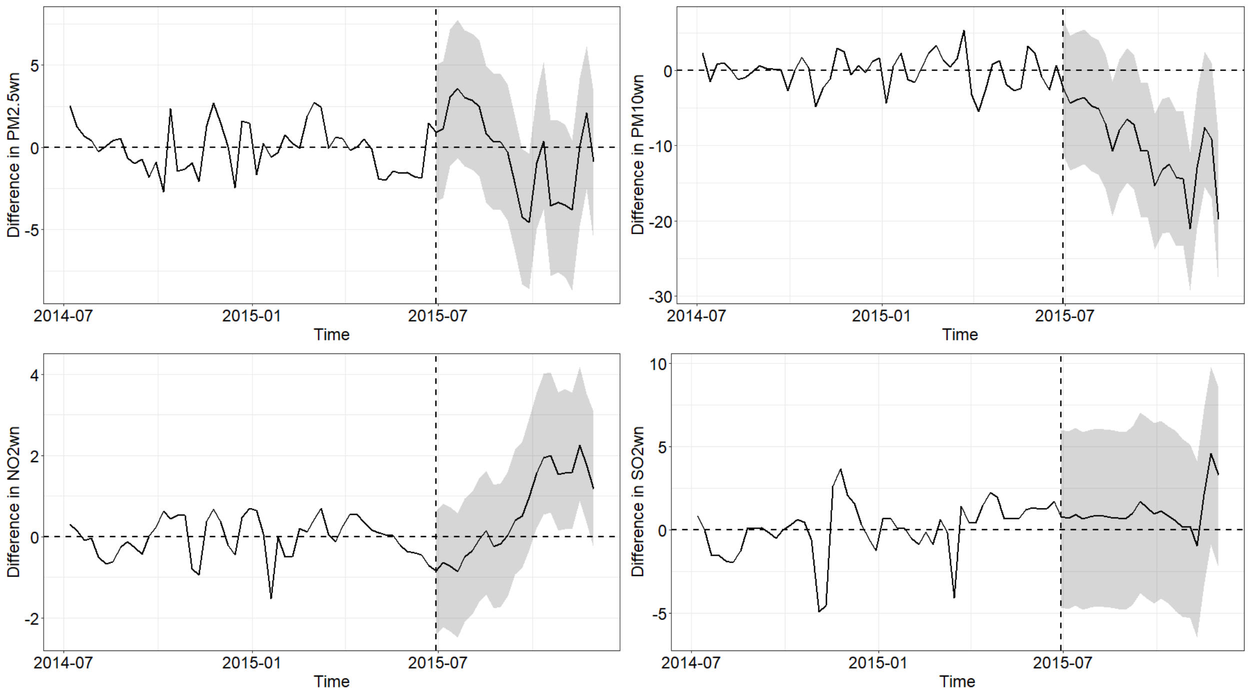


Note: The shaded area indicates the point-wise 95% confidence interval using Jackknife+ procedure

Figure 4.6: Ridge ASCM results on weather normalised and observed NO_2 and SO_2 in Hebei.

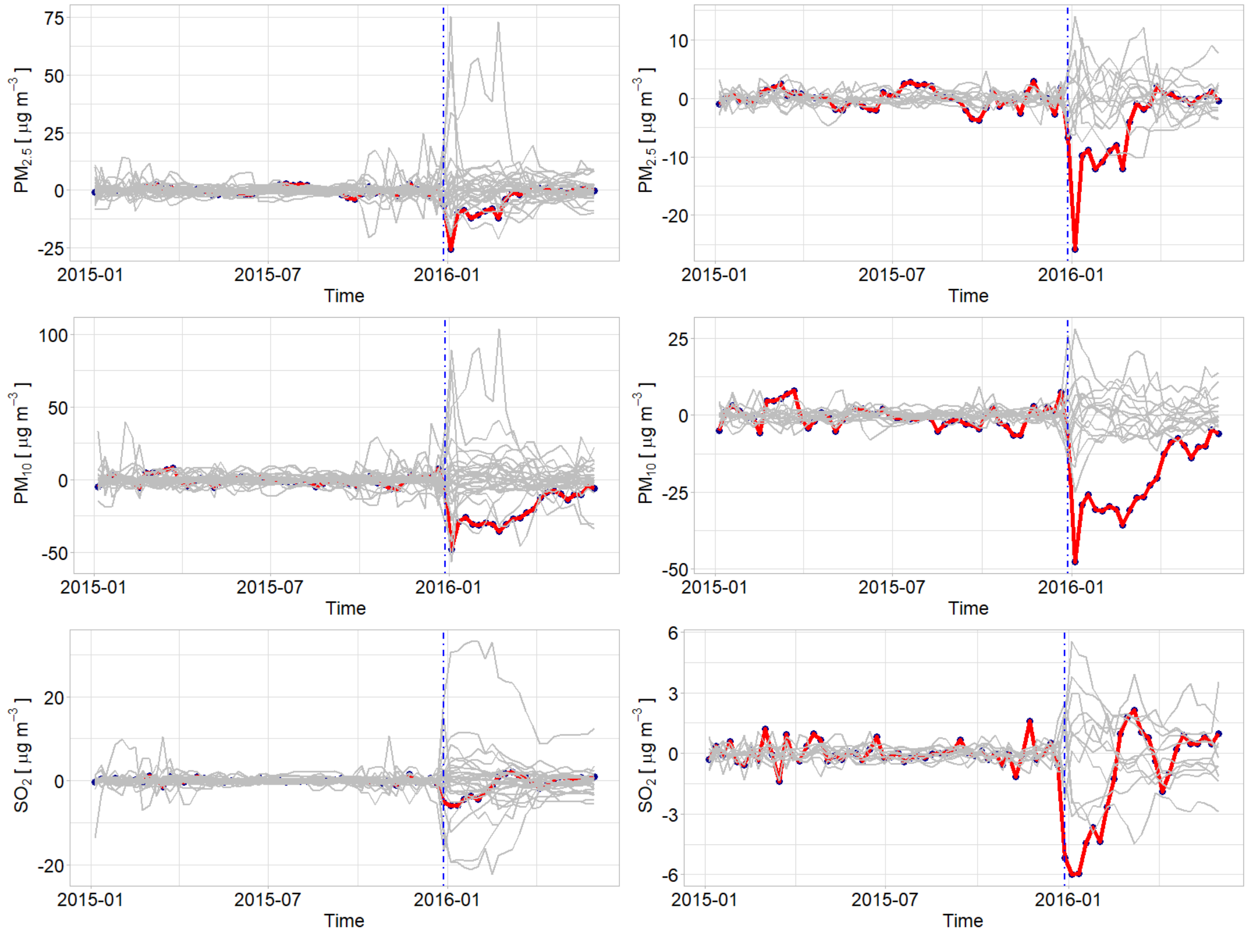
Note: See appendix for the results of each city with confidence interval.

Figure 4.7: The heterogeneity effects of inspection on weather normalised pollutants in 8 Hebei cities.



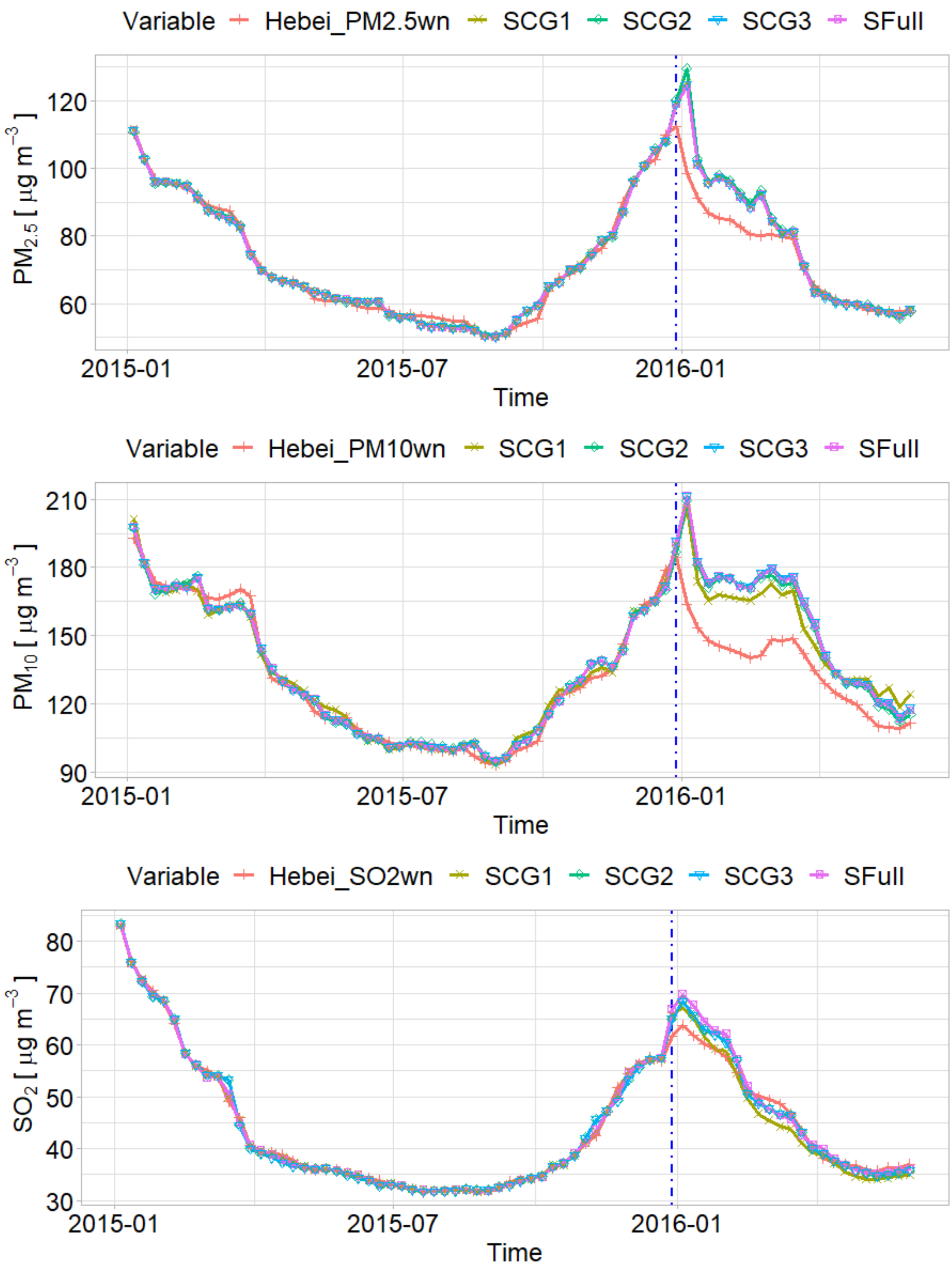
Note: We backdate 6 months and assume Hebei was inspected during July 2015 (instead of January 2016), the vertical lines refer to 1 July 2015. The shaded area indicates the point-wise 95% confidence interval using Jackknife+ procedure.

Figure 4.8: The in-time placebo test results of weather normalised $PM_{2.5}$, PM10, NO_2 and SO_2 in Hebei.



Note: The left figures plot the results of all 28 cities (Hebei + 27 control cities), the right figure plots the results after dropping the cities with worse pre-match than the main Hebei results. For $PM_{2.5}$wn, we drop cities with pre-policy match of Ridge ASCM worse than Hebei (i.e., L2 imbalance > 11.492 (Hebei)): Changchun, Changsha, Chengdu, Chongqing, Harbin, Hohhot, Jinan, Nanchang, Nanjing, Shenyang, Taiyuan, Urumqi, Xian, Yinchuan and Zhengzhou. For PM10wn, we drop cities with pre-policy match of Ridge ASCM worse than Hebei (i.e., L2 imbalance > 24.429 (Hebei)): Changchun, Harbin, Jinan, Lanzhou, Lhasa, Shenyang, Taiyuan, Urumqi, Xian, Yinchuan and Zhengzhou. For SO2wn, we drop cities with pre-policy match of Ridge ASCM worse than Hebei (i.e., L2 imbalance > 3.750 (Hebei)): Changchun, Guangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Nanchang, Shenyang, Taiyuan, Urumqi, Wuhan, Xining, Yinchuan and Zhengzhou.

Figure 4.9: The in-place placebo test results of weather normalised $PM_{2.5}$, PM10 and SO_2 in Hebei.



Note:

Appendix Table C.3 defines each synthetic control group (SCG).

Figure 4.10: The results of alternative control group tests on $PM_{2.5}$wn, PM10wn and SO_2wn in Hebei.

Table 4.1: A list of input variables (predictors) used in this study

Time variables	
day_julian	Day of the year, represents the seasonal variation
date_unix	Number of seconds since 1970-01-01, represents pollutant emission trend
weekday	Day of the week, represents the weekly variation
hour	Hour of the day, represents the hourly variation
Meteorological variables	
Surface observed data:	
temp	Temperature (degrees Celsius)
wd	Wind direction (m/s)
ws	Wind speed (in degrees, 90 is from the east)
rh	Relative humidity (%)
ERA5 Reanalysis data:	
blh	Boundary layer height (m)
tcc	Total cloud cover (Dimensionless)
ssr	Surface net solar radiation (J $m^(-2)$)
sp	Surface pressure (Pa)
tp	Total precipitation (m);
cluster	72-hour back-trajectory clusters (category)

Table 4.2: Random forest model performance in Shijiazhuang

Pollutants	FAC2	MB	MGE	NMB	NMGE
$PM_{2.5}$	0.8	1.147	29.663	0.013	0.339
PM10	0.853	-2.318	45.398	-0.016	0.306
NO_2	0.904	-0.841	11.793	-0.018	0.252
SO_2	0.855	-0.946	15.645	-0.019	0.314
Pollutants	RMSE	r	COE	IOA	
$PM_{2.5}$	43.306	0.867	0.228	0.614	
PM10	60.429	0.853	0.091	0.545	
NO_2	15.752	0.869	0.347	0.674	
SO_2	22.722	0.861	0.352	0.676	

Note : FAC2 (fraction of predictions with a factor of two), MB (mean bias), MGE (mean gross 224 error), NMB (normalized mean bias), RMSE (Root-mean-square deviation), NMGE (normalized mean gross error), r (correlation coefficient), COE (Coefficient of Efficiency), IOA (Index of Agreement).

Table 4.3: The impact of CEIP on air quality in Hebei (levels)

	(1)	(2)	(3)	(4)
(a) Variables	$PM_{2.5}$	PM10	NO_2	SO_2
Inspection	-11.380***	-15.214***	0.933	0.329
	(1.898)	(1.969)	(0.923)	(0.581)
R-squared	0.589	0.614	0.719	0.619
Weather Control	N	N	N	N
(b) Variables	$PM_{2.5}$wn	PM10wn	NO_2wn	SO_2wn
Inspection	-2.479**	-6.171***	1.450**	1.709**
	(1.155)	(1.385)	(0.697)	(0.635)
R-squared	0.798	0.859	0.913	0.796
Weather Control	N	N	N	N
(c) Variables	$PM_{2.5}$	PM10	NO_2	SO_2
Inspection	-8.405***	-17.288***	2.286**	-2.765**
	(1.243)	(1.691)	(0.870)	(1.255)
R-squared	0.643	0.643	0.781	0.725
Weather Control	Y	Y	Y	Y
(d) Variables	$PM_{2.5}$wn	PM10wn	NO_2wn	SO_2wn
Inspection	-2.349**	-8.276***	1.851**	0.090
	(0.891)	(1.272)	(0.718)	(1.011)
R-squared	0.833	0.891	0.926	0.858
Weather Control	Y	Y	Y	Y
City FE	Y	Y	Y	Y
Week FE	Y	Y	Y	Y
Observations	2,072	2,072	2,072	2,072
City Numbers	28	28	28	28

Note : *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Robust standard errors in parentheses, standard errors are clustered at the city level.

Table 4.4: The estimated impact of CEIP on air quality in Hebei using different methods

	(1)	(2)	(3)	(4)
ML+ Ridge ASCM	$PM_{2.5}$wn	PM10wn	NO_2wn	SO_2wn
(a1) Max. change (Levels)	-25.89	-47.80	-2.62	-6.00
(a2) Max. change (%)	-20.81%	-22.61%	-4.32%	-8.59%
(a3) Avg. ATT (Levels)	-4.69	-20.80	-1.11	-1.26
(a4) Avg. change (%)	-5.94%	-13.25%	-2.19%	-2.62%
DID (Obs.)	$PM_{2.5}$	PM10	NO_2	SO_2
(b1) Inspection (Levels)	-8.41***	-17.29***	2.29**	-2.77**
(b2) Inspection (Logs)	-15.22%	-16.76%	5.12%	-6.90%
DID (WN)	$PM_{2.5}$wn	PM10wn	NO_2wn	SO_2wn
(c1) Inspection (Levels)	-2.35**	-8.28***	1.85**	0.090
(c2) Inspection (Logs)	-4.84%	-7.05%	3.49%	2.12%

Note 1: *** p<0.01, ** p<0.05, * p<0.1.

Note 2: As shown in the first columns of Figures 5 and 6, the largest reduction for all four weather normalised pollutants occurred in the week immediately after inspection (i.e., the week starts 4 January 2016), thus the maximum changes (in levels) of weather normalised pollutants (i.e., row a1) due to CEIP inspection were calculated by (P_{hebei} - P_{syn_hebei}), and the maximum changes (in %) (i.e., row a2) were calculated by (P_{hebei} - P_{syn_hebei})/ (P_{syn_hebei}) * 100, where P_{hebei} and P_{syn_hebei} refer to the weather normalised pollutant values in Hebei and Synthetic Hebei on the week of 4 January 2016.

Note 3: Row (a3) represents the Average ATT (Average Treatment Effect on Treated) estimation of our main results (running Ridge ASCM using weather normalised value), the 95% CI lower and upper bond of the Average ATT are: $PM_{2.5}$wn "-4.69 (-9.61, -0.22)", PM10wn "-20.80 (-30.91, -8.11)", NO_2wn "-1.11 (-4.15, 0.42)" and SO_2wn "-1.26 (-4.42, 1.33)".

Note 4: Row (a4) represents the average change of the weather normalised value due to CEIP, which was calculated by (A_{hebei} - A_{syn_hebei})/ (A_{syn_hebei}) * 100, where A_{hebei} indicate the average weather normalised pollutant values in Hebei, and A_{syn_hebei} indicates the average weather normalised value in Synthetic Hebei (i.e., the "business as usual" level had Hebei not been inspected) after inspection (the period between Dec 28 2015 and 30 May 2016), for example, A_{syn_hebei} of $PM_{2.5}$wn were calculated by (A_{hebei} - (-4.69)), where "-4.69" was the avg. ATT estimated by Ridge ASCM using weather normalised $PM_{2.5}$ (i.e., column 1 and row a3). Note 5: The results in rows (b1) (b2) and (c1) (c3) come from the DID estimation (see Table 4.3 for detailed results in levels and Table C.1 in appendix for detailed results in logs).

Table 4.5: Sources and health effects of our four pollutants

Pollutant	Sources	Health Effects
Nitrogen Dioxide (NO_2)	Combustion processes, mainly for power generation, heating and motor vehicles	Respiratory difficulties, reduced lung function
Sulphur Dioxide (SO_2)	The burning of sulphur-containing fossil fuels, mainly from power generation, domestic heating and transport	Respiratory difficulties, irritation of the eyes
Coarse Particulate Matter (PM10)	Road transport and the burning of fuels for industrial, commercial and domestic uses	Respiratory difficulties and cardiovascular disease
Carbon Monoxide (CO)	Incomplete burning of fossil fuels in transport and industrial processes	Cardiovascular disease

Source: WHO (2018).

Table 4.6: Summary of previous literature on the PM10 mortality association

Author	Study region	Period	Key finding
Chen et al. (2012)	16 Chinese cities	1996-2008	“A 10 μg/m^3 increase in 2-day moving-average PM10 was associated with 0.35% increase of total mortality"
Chen et al. (2013)	16 Chinese cities	1996-2008	“A 10 μg/m^3 increase in PM10 concentrations corresponded to a 0.65% increase of total mortality"
Lu et al. (2015)	Meta analysis on 1464 studies covering 127 Chinese cities	Before December 2013	“In terms of short-term effects, the combined excess risks of total non-accidental mortality is 0.36% for each 10 μg/m^3 increase PM10.
Yin et al. (2017)	38 Chinese cities	2010-2013	“A 10μg/m^3 change in concurrent day PM10 concentrations was associated with a 0.44% increase in daily number of deaths%"

Table 4.7: The potential lives saved due to falling PM10 attributable to CEIP in Hebei and China

Region:	Hebei	China
Mortality effects:	15 μg/m^3	15 μg/m^3
Chen et al. (2012): 0.35%	565	12,828
Chen et al. (2013): 0.65%	1,048	23,824
Lu et al. (2015): 0.36%	581	13,195
Yin et al. (2017): 0.44%	710	16,127

Note: We assume the CEIP was effective in significantly reducing on average 15 μg/m^3 PM10 concentrations in Hebei for 3 months, and we time the mortality effects from previous literature with 1.5 to measure the change of 15 μg/m^3 PM10 concentrations. For china as a whole we assume the inspection would lead to reduction of PM10 concentrations for 15 μg/m^3 and last for 3 months as well. According to Hebei and National Bureau of Statistics, in 2015, the resident population of Hebei is 74.25 million and monthly mortality is 0.04825% , while the whole China has 1.375 billion population and monthly mortality is 0.05925%. A worked example: lives saved in Hebei from a 15 μg/m^3 reduction in PM10 using Chen et al.'s (20112) mortality estimates are calculated as 3((74,249,200 * 0.0004825)*(0.0035*1.5) = 565.

Chapter Five

Winter heating and air pollution in China

5.1 Introduction

The winter heating policy has been implemented in China for over fifty years. Each year when the winter season comes, northern Chinese cities will turn on the heating system and provide heating services to its urban residents. The heating service is provided through large-scale and centralised heating boilers and heated water or steam that is transported through pipelines that are connected to each household (Fan, He and Zhou, 2020). The heating system usually operates for around 4 months (between November to March) and then switches off at the end of the winter season. Since the heating system relies primarily on coal burning and is technically inefficient, one unintended consequence is that it contributes significantly to the poor air quality across northern Chinese cities which also results in a series of negative impacts to human health.

There is already a volume of literature investigating the impacts of the winter heating policy on air pollution. Using data between 1981 to 1993, Almond et al. (2009) found that the annual average levels of Total Suspended Particulates (TSP) are around 300 g/m^3 higher in northern cities than southern cities, due to the winter heating policy. Moreover, Ebenstein, Fan, Greenstone, He and Zhou (2017) estimated that China's winter heating policy led to PM10 concentration levels being 41.7 g/m^3 higher in northern cities. Similarly, Fan, He and Zhou (2020) documented that air quality in northern Chinese cities declines immediately after the operation of the heating system and found that, on average, turning on the heating system immediately increases the Air Quality Index (AQI) by 40 points in northern Chinese cities.

Despite the winter heating policy's impact on air quality, no large scale regulations have been implemented to tackle these problems by the Chinese central government until recently. In 2017 the Chinese central government introduced and implemented the 'Clean winter heating plan for Northern China (2017-2021)' which aimed to achieve low emissions and energy conservation by replacing the existing coal-based inefficient heating system with a series of 'coal-to-gas' and 'coal-to-electricity' policies. Following the implementation, it is important to evaluate the effectiveness of the clean winter heating plan. Several papers have attempted to do this, for example, Wang, Su et al. (2020) demonstrate that, compared to the winter in 2016, the clean heating strategies on average effectively reduced PM2.5 concentrations by 14% in Beijing (and surrounding regions) in winter 2017. Similarly, using data between 2017 and 2019, Zhang, Li and Wu (2020) found that the clean winter heating plan on average reduced AQI, PM2.5, NO_x and SO_2 by 20.4%, 18.59%, 34.1% and 68.4% respectively in 35 treatment cities.

Building upon these studies, this paper examines the impact of the winter heating system on air quality, as well as the effectiveness of the clean winter heating strategies in northern Chinese cities. More specifically, this paper has the following two objectives. Firstly, to use superior techniques to the previous literature to better understand the impact of turning-on and turning-off winter heating on local air quality levels. Secondly, to use these same techniques to more precisely quantify the effectiveness of the 'Clean winter heating plan for Northern China (2017-2021)' on air quality levels.

The previous literature that studies the impact of winter heating and the effectiveness of the clean winter heating plan on air quality, relies on traditional statistical models (e.g., econometric techniques as used in Almond et al. (2009), Zhang, Li and Wu (2020) and Fan, He and Zhou (2020)) or air quality models (as used in Wang, Su et al. (2020)). Each of these approaches has the challenge of decoupling the complex relationship between meteorology and observed pollutant concentrations. Since observed pollution concentrations are significantly influenced by meteorological conditions, it is necessary to strip away the effect of these conditions in order to reveal any policy effect on pollution (Grange, Carslaw et al., 2018). A failure to remove the effects of meteorological conditions in a convincing manner means any estimated policy effects will be unreliable. The key contributions of this study are therefore, first to apply advanced machine learning based weather normalisation techniques to remove meteorological effects to allow us to more precisely reveal the impact of winter heating and the effectiveness of the clean heating policy. Second, to do so using data up to 2021 thereby allowing us to show the effect of the clean heating policy over a longer time period than the previous studies.

We adopt a weather normalisation technique known as a random forest based machine learning model that removes the impact of meteorological variables on observed air pollutant concentrations to enable us to identify the policy effects. The weather normalisation technique was developed by Grange, Carslaw et al. (2018) and then widely used in a number of studies recently (Vu et al., 2019; Cole, Elliott and Liu, 2020; Shi et al., 2021; Dai et al., 2021). In air quality analysis, The random forest based machine learning algorithm has several advantages over conventional statistical analysis and air quality modeling (Grange, Carslaw et al., 2018). For instance, the random forest model is non-parametric and does not require the assumption of normality, homoscedasticity and independence among the input variables. Moreover, the random forest model is able to handle the complicated scenarios such as collinearity, non-linearity and interaction effects among different meteorological variables, which is challenging for traditional statistical models (Grange, Carslaw et al., 2018). Similarly air quality models (e.g. chemical transport models) can create high uncertainties when used to evaluate air pollution control interventions due to their dependence on the accuracy and availability of the emission inventory (Vu et al., 2019).

Using surface observed air quality data between January 2015 and May 2021, we run the weather normalisation model for 29 Chinese capital cities. Of these 29 cities, 15 are northern cities which represent the key target of our study, whereas the remaining 14 southern cities are used as a comparison group.

After decoupling the impact of meteorological conditions on air pollutant concentrations, our findings suggest that, during the winter period in 2015 (Nov 2015 to Mar 2016) and 2016 (Nov 2016 to Mar 2017), the turning on of the winter heating system immediately and dramatically increased the air pollution level (especially for SO_2) across northern Chinese cities. Similarly, the turning off of the winter heating system led to a sudden drop of air pollution. These results are consistent with Fan, He and Zhou (2020). More importantly, the immediate deterioration (improvement) in air quality as the winter heating system turned on (off) became less evident after 2017, and gradually disappeared in the winter period in 2019 (Nov 2019 to Mar 2020) and 2020 (Nov 2020 to Mar 2021). This evidence suggests the effectiveness of the clean heating plan which was implemented progressively since 2017. Additionally, through exploring the partial dependence relationship between winter heating (indicated by the HDD5 index) and air pollution, we precisely quantify the contribution of winter heating on air pollution level across all cities in each winter period. Indeed, substantial reductions of the contribution are detected in 11 (out of 15) northern cities. Lastly, we employ the ASCM approach on weather normalised pollution levels and quantify the contribution of winter heating on air pollution level in cities surrounding Beijing-Tianjin (referred to "2+26" cities throughout the chapter) and the rest 10 northern cities (referred to "North" cities throughout the chapter) as a whole. Our results suggest that, following a series of regulations implemented under the "Clean winter heating plan for Northern China (2017-2021)" since 2017, air quality (especially for SO_2 and CO)has largely improved in northern Chinese cities, demonstrating that the large scale clean heating plan was reasonably effective in mitigating air pollution problems in China.

5.2 Policy background

5.2.1 Winter heating policy in China

The proposal and implementation of the Chinese winter heating policy dates back to the period between the 1950s and the 1980s. The outdoor temperatures were relatively low in winter periods in China (especially in northern cities) which significantly affected the daily lives of residents. To help improve living standards, the Chinese central government decided to provide a winter indoor heating system to keep people warm. The principal idea of the winter heating system is to provide free or low cost (heavily subsidised) heating services through building centralised heating infrastructures in urban areas, where the large scale/capacity coal-fired boilers were connected to each household through hot water pipelines and radiators that can spread heat within homes and offices (Fan, He and Zhou, 2020).

According to Almond et al. (2009), due to limited budgets and resources at that time, the central government decided to build up the centralised winter heating system only in Northern Chinese cities.[1] Traditionally, China can be divided into northern and southern

[1] China was experiencing the "planed economy" period between the 1950s and 1980s.

parts through a geographic boundary line that connects the Qinling Mountains and the Huai River from west to east, therefore the winter heating policy is also known as the 'Huai River Policy' (Ebenstein, Fan, Greenstone, He and Zhou, 2017)[2]. Besides limited budgets and resources (which are primary factors), the central government also took the following factors into consideration when it decided to provide the winter heating system only for Northern cities. Firstly, the 'Qinling Mountains and Huai River' line is also known as the 'January 0 °C line' where the average temperature along the Huai River line is around 0 °C in January, while below 0 °C in the cities north of this line and above 0 °C in cities south of the line. Thus the northern cities are colder in winter and have a longer winter period. Secondly, compared to southern cities, northern cities have relatively abundant resources of coal (which is the single fuel source for the heating system) and are more industrialised and urbanised (during the planed economy period), thus northern cities have advantages in building up the heating infrastructure. However, up to now, the centralised free indoor winter heating system is still operating only in northern cities but not in southern cities.

It is worth noting that, the winter heating system in China is primarily coal-based and technically inefficient (Fan, He and Zhou, 2020) [3]. The combustion of coal would result in the emission of a number of pollutants such as $PM_{2.5}$, PM10, SO_2, NO_2, CO, etc., and lead to a resultant decrease in air quality. Since the heating coal-fired boilers were specifically designed for the winter heating system, and only operating during the winter period, it has been identified as the key contributor to the high levels of air pollution in Northern Chinese regions (especially during winter) and the health impacts of exposure to such high pollution levels are also well documented. For example, Almond et al. (2009) show that between

[2]Figure D.1 in the appendix illustrates the location of northern (centralised heating) and southern (non-centralised heating) cities in China

[3]The Chinese central government stated that in 2017, the coal fired heating area still accounts for 83% of the total heating area in Northern cities (Source: http://www.gov.cn/xinwen/2017-12/20/content$_5$248855.htm).

1981 and 1993, the winter heating policy led to the average yearly concentrations of total suspended particulate (TSP) being 300 g/m^3 higher in northern cities than those in southern cities. Similarly, using data from 1981 to 2000, Chen, Ebenstein et al. (2013) suggested that on average, the TSP concentrations were 184 μg/m^3 (55%) higher in northern Chinese cities due to the winter heating policy, and the long term exposure to an extra 100 μg/m^3 of total suspended particulates (TSPs) reduced life expectancy at birth by 3 years. Additionally, using weekly data between 2014 and 2015 in 114 Northern Chinese cities, Fan, He and Zhou (2020) found that turning-on the winter heating system led to an increase in the Air Quality Index (AQI) of 40 points. They also estimate that a 10 point increase in the AQI and a 10 μg/m^3 increase in $PM_{2.5}$ increased the weekly mortality by 2.2% and 2.5%, respectively.

5.2.2 Clean winter heating policy in China

Within the 'Clean winter heating plan for Northern China (2017-2021)', a series of policies and regulations have been implemented by the Chinese government to tackle the adverse impacts of the winter heating system on air quality and human health. According to the central government, the core principle of the clean winter heating strategy is to replace the coal-based and inefficient heating system with a system that incorporates clean energy sources, efficient heat transmission and distribution, and energy-efficient buildings with the aim of achieving ultra low emissions and energy conservation [4]. More specifically, great emphasis was put into the clean heating sources that include natural gas, electricity and solar energy (we well as geothermal, biomass, clean coal etc.). The clean winter heating plan targeted 15 provinces in northern China and covers the period between 2017 and 2021. [5]

The plan contains three key strategies to promote clean heating. The first is to select

[4]See http://www.gov.cn/xinwen/2017-12/20/content_5248855.htm for detailed information.

[5]The 15 provinces include Hebei, Henan, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shandong, Shaanxi, Gansu, Ningxia, Xinjiang, Qinghai, as well as Beijing and Tianjin.

heating sources in accordance with the local conditions, for example different cities could promote and adopt different clean heating sources (natural gas, electricity, solar energy, etc.) to supply their heating system, according to their resource abundance. The second is to comprehensively improve the efficiency of the heating supply network, for example targeted cities are required to maintain, update and optimize their old and inefficient heating pipe network, to become more energy efficient, automatically and intelligently operated. The third is to effectively reduce the heating energy consumption of households, for example by improving the energy efficiency of buildings across northern cities, to optimize the indoor radiator efficiency, etc.

Additionally, the clean winter heating plan proposed the following key measures to ensure the implementation of the plan. Central government will increase the funding support toward the renewable energy and clean heating development projects, the government and state owned electricity and gas enterprises are required to strengthen natural gas development and accelerate the construction of relevant infrastructure, targeted cities are required to optimise their electricity market through tiered pricing and peak-time pricing, in order to reduce the cost of electric heating.

More specifically, targeted cities are also required to reform the existing centralised heating system, for example to expand the area of central heating and eliminate small boilers that do not meet environmental requirements. Coal fired boiler stations in urban areas must achieve ultra-low emission standard, and install automatic monitoring facilities for air pollution sources. For boilers that fail to meet the requirements, alternative measures shall be formulated to clarify the shutdown and elimination plan and cancel subsidies. Regulation and monitoring will be strengthened for heating boiler stations, for example all coal-fired co-generation units with updating potentials in China must achieve ultra-low emission by 2020, energy inefficient boilers will be replaced with upgraded boilers, the quality of all the heating related equipment will be addressed and promoted, government will heavily regulate the sales of unqualified raw coals, the inferior dispersed coal is not allowed to enter into the market.

The implementation and enforcement of the aforementioned policies and regulations since 2017 were successful and effective according to a recent government report. The state council published the ‘Energy in China’s New Era’ white book in December 2020, illustrating the following key points. By 2020, 49.3 billion yuan hads been invested by central government to implement the clean winter heating plan, the clean heating rate across northern cities reached 55% (an 21% improvement compared to the 2016 level), the dispersed coal were replaced by 100 million tons, the area of clean heating across northern China reached 11.6 billion square meters (an increase of 5.1 billion square meters compared to 2016 level). By April 2020, 26.77 million households had been transformed into clean heating households in 43 northern cities. By 2020, a total number of 890 million KW coal-fired power units reached ultra-low emission, accounting for 86% of the total installed capacity of coal-fired power, and more than 750 million KW coal-fired power units implemented energy-saving transformation, moreover, more than 200,000 small coal-fired boilers are eliminated, and the coal-fired boilers below 35 steam tons / hour in key targeting areas are basically cleared.[6]

5.3 Data

5.3.1 Air pollution data

We collect hourly concentrations of six pollutants ($PM_{2.5}$, PM10, SO_2, NO_2, CO and O_3) for 31 capital cities from the ‘China National Environmental Monitoring Centre website’, covering the period between 1 January 2015 to 31 May 2021. Table A1 (appendix 4) shows

[6] See http://www.gov.cn/zhengce/2020-12/21/content_5571916.htm for more detailed information.

the official date that the winter heating policy starts and ends each year in these 29 cities. Since most of the winter heating periods were between November to March, we include the data until May to allow 1-2 months after the heating ends. [7]

5.3.2 Meteorological data

To apply the weather normalisation method, following Cole, Elliott and Liu (2020) and Shi et al. (2021), we match our air quality data with hourly-city level meteorological variables. Using the 'worldMet' R packgage (https://github.com/davidcarslaw/worldmet), we collect surface observed meteorological variables (air temperature, relative humidity, wind direction, wind speed, and pressure) from the National Oceanic and Atmospheric Administration (NOAA). In addition, the boundary layer height (blh), total cloud cover (tcc), surface net solar radiation (ssr), surface pressure (sp) and total precipitation (tp) were collected from the ERA5 reanalysis data set (from the European Centre for Medium-Range Weather Forecasts: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview). The hourly data of "72-hour backward air mass trajectories" for each corresponding city and time period were acquired and calculated using the "HYSPLIT model(Hybrid Single-Particle Lagrangian Integrated Trajectory)" from NOAA.

5.3.3 Additional predictors

In addition to the above variables, we also include two more predictors, the Chinese lunar day and the Heating Degree Days index (HDD5).

[7] 29 capital cities include: Beijing, Changchun, Changsha, Chengdu, Chongqing, Fuzhou, Guangzhou, Guiyang, Hangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Nanchang, Nanjing, Nanning, Shanghai, Shenyang, Shijiazhuang, Taiyuan, Tianjin, Urumqi, Wuhan, Xian, Xining, Yinchuan, Zhengzhou.

In many previous studies that have applied the machine learning method, the standard predictor accounting for the seasonal term refers to the day-julian (day of the year), which was based on the Gregorian dates that were widely used in many western countries. However, according to Dai et al (2021), the day-julian can not fully capture the holiday effects that are specifically related to the traditional Chinese lunar calendar, such as the Chinese lunar Spring Festival (CSF), which could have a significant impact on the emission strengths due to changing transportation and industrial activities. We therefore include the Chinese lunar day (day_lunar) as an additional predictor for the machine learning model to account for the emission changes resulting from Chinese lunar holidays. [8]

We collect the Heating Degree Days index with base temperature at 5°C (HDD5) from https://www.degreedays.net developed by BizEE Software, and match it with air pollution and meteorological data for all 29 cities. The heating degree day measures the extent to which the outside air temperature was lower than a specific base temperature (or balance point) and for how long. It therefore provides an indication of the likely required energy consumption for heating during cold periods which are extensively used in building related energy consumption analyses. [9]

5.4 Methods

In this chapter, we firstly apply our weather normalisation techniques to de-weather the data and then take an ASCM approach to estimate the causal impact of the winter heating on air quality levels. Our ML-ASCM approach weather normalises the data using a machine learning technique to decouple the effects of meteorological conditions from observed air

[8] Dai et al (2021) indicate that the performance of machine learning models improves once the Chinese lunar day was included.

[9] See detailed explanation at https://www.degreedays.net/introduction.

pollution concentrations and then uses the de-weathered air pollution concentrations in a Ridge ASCM (Augmented Synthetic Control Method) model to isolate the causal impact of the heating policy.

5.4.1 Random forest-based weather normalisation method

The random forest (RF) based weather normalisation technique was conducted using the "rmweather" R package developed by Grange, Carslaw et al. (2018). Compared with the decision tree model, the random forest model reduces the variance and overcomes the overfitting problems associated with single decision tree models (Breiman, 2001). In our study, we conduct the weather normalisation for 6 pollutants in 29 cities separately. More specifically, our set of explanatory variables ("predictors") include meteorology variables (air temperature, relative humidity, wind direction, wind speed, pressure, boundary layer height, total cloud cover, surface net solar radiation, surface pressure and total precipitation), time variables (date_unix, day_julian, lunar_day, weekday, hour), hourly air mass backward trajectory clusters (72-hour back-trajectory clusters), HDD5. These were used to build the RF model and to predict the six air pollutant concentrations. By default, the observed data was randomly sampled into a training set (70%) and a test set (30%). The training set was used to train the random forest model and the test set to test the model performance. The number of trees to grow to make up the forest is 300 (n_trees = 300), the number of times to sample the whole data and then predict is 300 (n_samples = 300), the number of variables to possibly split at in each node is 3 (mtry=3), and the minimum size of terminal nodes for the model is 5 (min_node_size=5).

Given that the major objective of this study is to quantify the short term impact of the winter heating period on air quality levels, we follow the weather normalisation procedure by Shi et al. (2021) who document the impact of Covid-19 lockdowns in 11 global cities. The random forest model for each pollutant is built using the above discussed predictors and model parameters for 29 cities. More importantly, we normalise weather for each winter heating period of each city separately using data between 1 September and 31 May for five consecutive years. [10]

The original weather normalization technique was developed by Grange et al. (2018), the core concept of which is to simply apply the random forest model (built using all the predictors) to predict pollutant concentrations using the re-sampled predictor data set. Taking 5 May 2015 of Beijing for instance, all the time (excluding “date_unix”) and meteorological variables at any day within the original dataset period are randomly selected to feed into the random forest to predict the concentration on 5 May 2015. This procedure is usually repeated e.g. 500 times and the final weather normalised concentration for this date is given by averaging these 500 predictions. It is worth noting that Grange et al. (2018) not only de-weather but also de-trend the observed concentration level (where de-weather refers to the prediction using the replaced meteorological variables and de-trend refers to using the replaced time variables), which makes it difficult to compare the emission strength (weather normalised pollutant value) year by year.

In this study, we use the modified weather normalisation technique proposed by Vu et al (2019) and Shi et al (2021). There are two key differences between the original Grange et al (2018) method and these modified weather normalisation methods. Firstly, using the modified technique we apply the random forest model to predict pollutant concentrations only using the re-sampled meteorological data set (therefore only de-weathers but does not de-trend). Secondly, rather than de-weather the whole data period (e.g., 1 January 2015 to 31 May 2020), we only de-weather the winter heating period year by year and do so for six

[10]Taking Beijing for example, we perform the weather normalisation model using the data periods 1 Sep 2015 to 31 May 2016, 1 Sep 2016 to 31 May 2017, 1 Sep 2017 to 31 May 2018, 1 Sep 2018 to 31 May 2019 and 1 Sep 2019 to 31 May 2020, respectively.

separate years. The key reason for de-weathering each winter period separately is that the weather conditions differ a lot between summer (when there is no winter heating) and winter (when there is winter heating) and these differences may also vary across years. Using only the winter heating period data to conduct weather normalisation can potentially reduce the uncertainty (brought by summer period weather conditions), and more importantly, enable us to detect the short term changes in emission strengths due to the starting and ending of winter heating, and coal-to-gas/ clean heating policies implemented in specific cities and specific years (Shi et al., 2021).

Lastly, compared with other 'black-box' machine learning algorithms, one advantage of the decision tree based random forest model is that it allows researchers to investigate the marginal effect (i.e., the response relationship) of a specific predictor on an outcome variable holding all other predictors fixed at their average levels(Grange and Carslaw, 2019). Here we make use of the partial dependence plot (extracted from the RF model built for each pollutant in each city and each heating season) to quantify the impact of winter heating and the clean heating policy on local air pollution.

5.4.2 The Augmented Synthetic Control method

Isolating the effectiveness of air pollution control strategies has always been challenging since there are many confounding factors that would mask and complicate the direct causal policy impacts. This challenge can be potentially addressed using causal inference methods in Economics, such as Difference-in-Difference (DID) and Synthetic Control Method (SCM). To investigate the causal impacts of both winter clean heating and clean heating policies on air quality in northern China, we utilize the recently developed Augmented Synthetic Control Method (ASCM).

The SCM was proposed to estimate the effects of policy intervention which are implemented at aggregate levels like cities (provinces/countries), the principal idea of the SCM is based on the "comparative case study", where the effectiveness of an intervention can be identified through comparing the movement of the outcome variable between a treatment group (e.g., cities exposed to intervention) and a control group (similar cities to treatment group but not exposed to intervention). A weighted average combination of a set of control units is used to construct the counterfactual scenario of in treatment group it were if not exposed to the intervention, if the "synthetic unit" can closely mimic the evolution of outcome variable for the treatment unit before policy intervention, then the difference (after the intervention) of outcome movement between treatment unit and synthetic unit can be regarded as the causal impact of the intervention (Abadie, Diamond and Hainmueller, 2010; Abadie, Diamond and Hainmueller, 2015).

The ASCM applied in this chapter extends from original SCM and uses an outcome model (Ridge regression model) to estimate the bias due to imperfect pre-intervention fit, and then directly corrects the original SCM estimate using the outcome model. The ASCM approach was implemented by the "augsynth" R package, and a detailed description of the method can be found in Ben-Michael, Feller and Rothstein (2021). In our study, to understand the impact of central heating on air quality in northern China, 15 northern Chinese cities excluding "2+26" cities are assigned as a treatment group (exposed to central heating) and 15 southern Chinese cities (not exposed to central heating) as a control group. Similarly, to understand the impact of clean winter heating, 6 Chinese major cities from "2+26" cities (exposed to clean heating) were assigned as a treatment group and the other northern cities as a treatment group. The geographic locations of the treatment and control groups are shown in Fig. D.1. The Ridge ASCM were used on both observed and deweathered air pollutant concentrations on a year-by-year base to isolate the effects from winter heating in each heating season. To allow for sufficient pre-treatment period in each year, the pre-treatment period is from the end date (1 June) of last heating period to the start date (range from October-November) of the current heating period in each city and post-treatment period is the current heating period (from Oct-Nov to 31 May next year). The policy impacts (synthetic difference / ASCM estimate) are derived from the difference in air quality between each treatment unit and its corresponding synthetic counterfactual unit, the synthetic difference in "North" group and "2+26" cities group are calculated using population-weighted-average synthetic difference from 15 cities and 6 cities, respectively.

5.5 Results

5.5.1 Variable importance

The random forest approach enables us to interpret and explain the relationship among different variables as well as their importance in predicting the output variable (Grange, Carslaw et al., 2018). For example, Figures 5.1 to 5.3 report the importance of each variable in the random forest model on SO_2, $PM_{2.5}$ and PM_{10} of Beijing in the 2015 winter heating season, respectively. As shown, the pollutant emission trend represented by date_unix, the seasonal variation represented by day_julian and the emission changes from lunar holidays represented by day_Lunar are the three most important variables in predicting SO_2, $PM_{2.5}$ and PM_{10} concentrations, suggesting that local and regional emissions still dominate concentration levels of the three pollutants (Grange, Carslaw et al., 2018). HDD5 plays the most important roles in predicting SO_2 besides the above variables, suggesting the primary contribution of winter heating to SO_2, which further demonstrates the importance of HDD5 as the indicator of winter heating. The meteorological variables including surface pressure (sp), relative humidity (RH) and temperature (temp) are all moderately important in predicting the three pollutant concentrations.

[Figure 5.1 about here]

[Figure 5.2 about here]

[Figure 5.3 about here]

5.5.2 Impact of winter heating on local air quality in Northern Chinese cities

The impact of the turning on and off of winter heating on local air quality levels is reported in Figures 5.4 to 5.7. More specifically, Figures 5.4 and 5.5 show the impact of turning on and off winter heating on SO_2 levels in 15 northern cities between 2015 and 2021 (six winter heating seasons), respectively. In Figure 5.4, the horizontal axis refers to the 50 days before and after the turning on/off of the winter heating system, the vertical axis represents the weather normalised SO_2 concentration levels. The vertical dotted line refers to the official turning on/off date within each city, which varies in each city and across different years. The local government of each city has an official on/off date, however the actual on/off date might be different from the official date, and depend on the specific temperature. For example, if the outdoor temperature of a city was below 5 °C for 5 previous consecutive days prior to the official turn-on date , the local government can turn on the heating system in advance. Similarly, local government can also postpone the switch off of the winter heating system if the outdoor temperature remains low after the official turn-off date.

[Figure 5.4 about here]

[Figure 5.5 about here]

According to Figure 5.4, sharp increases of weather normalised SO_2 around the turn-on date are detected in the majority of the 15 Northern cities that include Beijing, Changchun, Harbin, Hohhot, Shenyang, Shijiazhuang, Taiyuan, Yinchuan and Zhengzhou, especially in 2015 and 2016. Similarly, according to Figure 5.5, relatively sharp reductions of weather normalised SO_2 around the turn-off date are detected in Beijing, Changchun, Harbin, Hohhot, Shenyang, Shijiazhuang, Taiyuan, Xian, Yinchuan and Zhengzhou in 2015 and 2016, suggesting that the turning-on and off of the winter heating system had a substantial impact on SO_2 concentrations in these cities in 2015 and 2016. In addition, the abrupt increase or decrease in some cities occurred prior to the vertical line, confirming that the actual turning on/off date were not identical to the official date.

It is noticeable that the sharp changes of weather normalised SO_2 around the vertical lines in most of the 15 northern cities are less evident after 2017, suggesting that compared with 2015 and 2016, the turning on and off of the winter heating system led to a notably smaller change in weather normalised SO_2 after 2017, especially in 2019 and 2020. The dramatically reduced impact of the winter heating on SO_2 can be primarily attributed to the implementation of the Clean winter heating plan for Northern China (2017-2021).

For comparison, Figures 5.6 and 5.7 plot the weather normalised SO_2 before and after the hypothetical turning on and turning off dates in 14 southern cities. We set 15 Nov and 15 March as the hypothetical turning on and off dates for the southern cities since these two dates are the average dates of turning on and off the heating system in northern cities (Almond et al., 2009). Since the centralised winter heating system does not exist in southern cities, it is not surprising that we do not find abrupt increases and decreases around the vertical dotted line, especially in cities like Changsha, Chengdu, Chongqing, Fuzhou, Hangzhou, Hefei, Kunming, Nanchang and Wuhan, where the weather normalised SO_2 were relatively flat around the vertical line across the six years. Such evidence further demonstrates that the impact of winter heating on SO_2 levels is only evident in northern cities. [11]

[11]Guiyang shows very similar trend with northern cities, where the sharp increase and decrease of weather

[Figure 5.6 about here]

[Figure 5.7 about here]

To better illustrate how the winter heating system interacts with SO_2, temperature and the HDD5 index, Figure 5.8 plots the daily trend of observed and weather normalised SO_2, HDD5 and temperature in Beijing for six winter heating seasons. The horizontal axis represents the time period between e.g., Beijing_2015 refers to the time period between Sep 2015 to May 2016, and the vertical axis represents the value of the four different trends. As we can see, the daily observed SO_2 is volatile while weather normalised SO_2 is relatively stable, suggesting that the use of observed SO_2 to understand the impact of winter heating on air pollution is challenging and separating the meteorological impacts on observed concentrations allows us to illustrate the changes in emission strength. Moreover, taking Beijing_2015 as an example, it can be observed that as temperature gradually reduced to 0 °C the HDD5 remains at 0, when temperature intersects with HDD5 the winter heating system turns on and the weather normalised SO_2 begins to rise sharply, when temperature intersects HDD5 at the end of the heating season, the heating system turns off and the weather normalised SO_2 falls sharply. Nevertheless, Figure 5.8 confirms the role played by HDD5 in the turning on and off of the winter heating system and thus HDD5 will be used as a key indicator to further investigate and quantify the impact of heating on air pollution levels.

[Figure 5.8 about here]

---

normalised SO_2 can be detected around the vertical dotted line, this is due to the large scale residential usage of coal burning during winter period, as well as the industrial usage of coal for local metal industry.

5.5.3 The effectiveness of the Clean Winter Heating Plan for Northern China (2017-2021) on local air quality

Figure 5.9 illustrates the HDD5-SO_2 partial dependence response plots for Beijing in six heating seasons. The vertical axis refers to SO_2 levels and the horizontal axis refers to HDD5 levels. Figure 5.9 (a) presents plots for each of our six years separately. As shown in Figure 5.9 (a), for the 2015 winter heating season (Sep 2015 to May 2016) in Beijing, HDD5 is (non linearly) positively related with (observed) SO_2 concentration levels (holding all other predictors fixed at average level), as HDD5 increase from 0 to 5 SO_2 increases sharply from 11.3 to 16.3 μg/m^3, which then gradually increases to 17.3 μg/m^3 as HDD5 increases from 5 to 15. Lastly as HDD5 increases from 15 to 20, the SO_2 remains stable around 17.2 μg/m^3. Similar trends can also be detected throughout Beijing 2016 to Beijing 2020. To sum up, during the 2015 winter heating season, from the turning on to the daily normal operation of the winter heating system, concentrations of SO_2 increased from 11.3 to 17.3 μg/m^3.

[Figure 5.9 about here]

On the other hand, Figure 5.9 (b) presents the six years' plots together within one figure, which allows for easier comparison across years. According to Figure 5.9 (b), for the 2015 and 2016 heating seasons, the winter heating system contributed to large increases of SO_2 in Beijing, while the contribution of winter heating to SO_2 has fallen dramatically since 2017 (compared to 2015 and 2016), which further reduced to only 0.75 (4.00-3.25) μg/m^3 in the 2020 winter season. These results show clear evidence of the effectiveness of the Clean winter heating plan implemented in 2016 and 2017 in Beijing.

[Figure 5.10 about here]

Following the example of HDD5-SO_2 response in Beijing, Figures 5.10 and 5.11 illustrate the HDD5-SO_2 response plots in 15 northern and 14 southern cities, respectively.

According to Figure 5.10, the majority of the northern cities (Beijing, Changchun, Harbin, Hohhot, Shenyang, Shijiazhuang, Taiyuan, Tianjin, Xian, Xining, Yinchuan and Zhengzhou) share similar trend with Beijing, i.e., the winter heating contributed to a substantial and sharp increase of SO_2 in 2015 and 2016, whereas such contributions reduced dramatically since 2017. As for Lanzhou, the winter heating contributed to the largest increase of SO_2 in 2017 (rather than 2015 or 2016).

[Figure 5.11 about here]

On the other hand, Figure 5.11 shows that, for the majority of southern cities (Changsha, Chengdu, Chonqing, Hangzhou, Hefei, Kunming, Nanchang, Nanjing, Nanning, Shanghai and Wuhan), the heating season of each year did not contribute to a large and abrupt rise of SO_2. Indeed, the HDD5-SO_2 response plots were flat and relatively stable. One exception is Guiyang, which shows a trend similar to the northern cities with SO_2 being positively related to the HDD5. This can be partially explained by the large scale residential usage of coal burning during the winter period, as well as the industrial usage of coal for the local metal industry.

[Figure 5.12 about here]

Furthermore, Figures 5.12 reports the difference between the maximum and minimum values of SO_2 for each year's HDD5-SO2 response plot in 15 northern cities. As can be seen, winter heating contributes to the largest increase of SO_2 in 14 cities (except for Lanzhou) either in 2015 or 2016. For example, it contributed to 41.34 μg/m^3, 24.98 μg/m^3, 20.32 μg/m^3, 16.35 μg/m^3 and 12.18 μg/m^3 increase of SO_2 in Taiyuan (2016), Yinchuan (2015), Shijiazhuang (2016), Shenyang (2015) and Hohhot (2016), respectively. 2017 represents a clear threshold where the impact of the winter heating on the increase of SO_2 reduces dramatically in most cities since 2017. For example, the contribution of winter heating to the increase of SO_2 reduced from 41.34 μg/m^3 (2016) to 21.39 μg/m^3 (2017) in Taiyuan, from

20.32 μg/m^3 (2016) to 10.27 μg/m^3 (2017) in Shijiazhuang, from 14.55 μg/m^3 (2016) to 6.98 μg/m^3 (2017) in Shenyang, from 19.09 μg/m^3 (2016) to 6.84 μg/m^3 (2017) in Yinchuan, from 12.18 μg/m^3 (2016) to 5.77 μg/m^3 (2017) in Hohhot, from 9.62 μg/m^3 (2016) to 3.01 μg/m^3 (2017) in Urumqi, from 6.29 μg/m^3 (2016) to 2.38 μg/m^3 (2017) in Beijing.

[Figure 5.13 about here]

Conversely, Figure 5.13 reports the difference of the maximum and minimum value of SO_2 levels for each year's HDD5-SO2 response plot in 14 southern cities. Besides Guiyang, the winter heating period is associated with no discernible change of SO_2 for all six heating seasons, as we would expect.

[Figure 5.14 about here]

[Figure 5.15 about here]

[Figure 5.16 about here]

[Figure 5.17 about here]

[Figure 5.18 about here]

[Figure 5.19 about here]

[Figure 5.20 about here]

[Figure 5.21 about here]

Additionally, we conduct the same analysis for the our other pollutants. Figures 5.14 to 5.21 show the difference of the maximum and minimum value of NO_2, $PM_{2.5}$, PM10 and CO levels for each year's HDD5-NO_2, HDD5-$PM_{2.5}$, HDD5-PM10 and HDD5-CO response plots in 15 northern and 14 southern cities, respectively. We do not find strong evidence of the impact of the clean heating plan on NO_2, since the winter heating is not the dominant source of NO_2 emissions. The impact of the clean heating plan on $PM_{2.5}$ is evident in Beijing, Changchun, Harbin, Shenyang, Shijiazhuang, Taiyuan, Tianjin, Zhengzhou, substantial and large drop of the contribution of heating to $PM_{2.5}$ after 2016 are detected in these cities (see Figure 5.16). Such impacts can also be observed for CO in Beijing, Harbin, Hohhot, Shijiazhuang, Taiyuan, Tianjin and Zhengzhou.[12]

[Figure 5.22 about here]

Finally, to provide a comprehensive picture, we divide northern cities into two subgroups, including the “2+26” and “North” cities, the location of which are illustrated in Figure D.1 of Appendix. [13] we apply the ML-ASCM approach and present the results in Figures 5.22, 5.23 and 5.24, respectively.

Figure 5.22 shows the weekly values for both observed and deweathered PM2.5, PM10, CO, SO2, NO_2 concentrations and SO2/NO_2 ratios from 2015 to 2021 in the three regions. Note that the SO2/NO_2 ratio is an indicator for the relative importance of stationery and traffic sources on air pollution. According to Figure 5.22, clear seasonal cycles – a nearly uniform pattern during non-heating periods and a unimodal pattern during heating periods (broadly refer to 15 Nov to 15 March) - were observed for all the air pollutants in the northern China though observed values show greater fluctuation than weather-normalized ones. The start and end of the seasonal peaks coincided well with the start and end of the heating

[12]The incomplete burning of coal is the major driver of the emission of SO_2, $PM_{2.5}$ and CO, which explains why the impacts of winter heating are evident on these pollutants, whereas coal burning is not the major contribution of NO_2 and PM10

[13]“2+26”cities include Beijing, Tianjin, Shijiazhuang, Taiyuan, Jinan and Zhengzhou, where stricter clean heating policies were implemented since 2017. “ North” cities refers to the rest northern cities excluding “2+26”cities, including Harbin, Changchun, Shenyang, Hohhot, Yichuan, Xian, Lanzhou, Xining, Lhasa and Urumqi, they were exposed to the winter heating police.

periods. In addition, percentage changes in CO, SO2, and NO_2 between the heating and non-heating periods during 2015-2021 in southern China are 22.5% (22.3%), 16.8% (20.6%) and 24.7% (16.7%), respectively, which are smaller than 53.7% (65.6%), 84.4% (92.0%) and 31.1 (27.7%) in "2+26 cities" and much smaller than 62.4% (52.3%), 153% (123.1%) and 34.7% (25.0%) in northern cities. This confirms that more enhanced wintertime air pollution in northern than southern China was related to heating-related emissions. In addition, the impacts of winter heating on air quality in "2+26" cities become less evident than those in the other northern cities.

[Figure 5.23 about here]

[Figure 5.24 about here]

Figures 5.23 illustrates the difference (of $PM_{2.5}$ and SO_2) in observed, deweathered and synthetic counterfactual (estiamted from the ASCM model) between each treatment and control group, while Figure 5.24 illustrates the results for PM10, CO and NO_2 NO_2. We found different effects of heating polices from observed, deweathered and synthetic differences in air pollutant concentrations between paired groups during heating, non-heating and full-year periods. Large bias exists in estimating the true air quality effects from different winter heating policies using observed or deweathered differences between paired groups. The observed differences in air quality comprised of differences in both emission and meteorology between the groups. Although the weather normalization technique can remove the confounding effects of meterological conditions on air pollutant concentrations, deweathered differences cannot be attributed to the heating policies due to the imbalance during pre-treatment periods, e.g., deweathered differences between "2+26" cities and the other northern cities are imbalanced (i.e., 0) during non-heating periods (Fig. 5.22B). As shown in Fig. 5.23 the synthetic differences in air pollutant concentrations between the control and the treatment group during non-heating periods are close to 0, indicating that the ASCM successfully removed the bias from imbalance in pre-treatment outcomes. The ACSM corrected concentration is then used as the counterfactual, from which we can then derive the causal effect. Furthermore, synthetic differences in observed concentrations between paired groups were not useful because of weather effects and poor performance of fitting during the pre-treatment periods. We also tried to build the counterfactual with the observed concentrations but the large weather effects made it impossible to obtain an ideal pre-treatment fitting. Among the three metrics, only synthetic differences with deweathered concentrations between paired groups are causal effects of the heating policies.

[Table 5.1 about here]

Further more, Table 5.1 reports the average treatment effects estimated from ML-ASCM approach on SO_2wn, $PM_{2.5}$wn and PM10wn respectively, the causal contribution of winter heating policy on SO_2 reduced from 27.4 μg/m^3 (95% CI: 13.8, 42.9) in 2014 (June 2014 to May 2015) to 6.5 μg/m^3 (95% CI: 1.0, 12.1) in 2017 (June 2014 to May 2015) and eventually 0.2 μg/m^3 (95% CI -2.9, 3.1) in 2020 (June 2020 to May 2021). The result also demonstrate the effectiveness of clean heating policy in "2+26" cities as the contribution from winter heating on SO_2 were largely reduced after 2017 and close to 0 in 2020. Conversely, the contribution of winter heating on $PM_{2.5}$ and PM10 remained consistent from 2014 to 2020, suggesting the effectiveness of clean winter heating policy were less evident on these two pollutants. Similarly, Table 5.2 reports the average treatment effects estimated from ML-ASCM approach on NO_2wn, COwn and PM10wn respectively. Compared to SO_2wn, the effectiveness of clean winter heating on NO_2wn were less evident, since NO_2 is primarily contributed by vehicle emissions. As shown in Table 5.2, the contribution of winter heating on COwn reduced from 0.5 mg/m^3 (95% CI: -0.1, 1.) in 2014 to 0 mg/m^3 (95% CI: -0.2, 0.2) in 2020, demonstrating the effectiveness of clean heating on CO.

[Table 5.2 about here]

5.6 Conclusion and Discussions

The winter heating policy has been implemented in China for over 5 decades. On one hand, the free (or heavily subsidized) heating system (largely coal-based) has greatly improved the welfare of residents living in northern part of China in terms of helping them to cope with the extreme cold during the winter period. On the other hand, the policy also brings about severe degradation of air quality and human health in these northern Chinese cities (Ebenstein, Fan, Greenstone, He and Zhou, 2017). To tackle this problem, in 2017 the Chinese central government implemented the Clean winter heating plan for Northern China (2017-2021) which implemented coal-to-gas and coal-to-electricity strategies and a series of regulations aiming to achieve emission reductions and energy conservation.

Using hourly air pollution data from surface observations and meteorological data, this study firstly employs a machine learning based weather normalisation technique that allows to decouple the impacts of weather conditions on observed air pollutant concentrations, and then apply the Augmented synthetic control to isolate the causal impacts of winter heating on air pollution in China under the quasi-experimental design. Our study firstly documents the impact of the turning on and off of the winter heating system on air quality. Second we quantify the contribution of the winter heating system to air pollution in 29 Chinese capital cities, based on the partial dependence plot within the random forest model. Finally we estimates the causal impact of winter heating on 6 weather normalised pollutants for each year from 2014 to 2021. Our results confirm the immediate increase and reduction of air pollution when the heating system turns on and off, respectively, in 15 northern cities. Similar patterns are not detected in 14 southern cities suggesting these Northern patterns are indeed driven by the winter heating policy. Our findings demonstrate that the winter heating system makes a major contribution to air pollution levels (represented mainly by SO_2, $PM_{2.5}$ and CO) in northern Chinese cities during the winter period in 2015 and 2016. However, since the implementation of the clean winter heating plan, the contribution of winter heating on air pollution levels reduced substantially. Again, such trends are not detected in southern cities.

Since identifying the effectiveness of air pollution control policies can be very challenging, it is of great importance to take meteorological variables into account and to remove their impacts from pollutant concentrations in air quality trend analysis. Based on the weather normalisation machine learning model, our results demonstrate that the Clean winter heating plan for Northern China (2017-2021) played a key role in mitigating air pollution emissions. Our results also suggest that the adoption of coal-to-gas and coal-to-electricity strategies can be effective in supporting the transition of China's clean winter heating target.

Chapter 5 Figures and Tables

Figure 5.1: Variable importance of SO_2 in Beijing

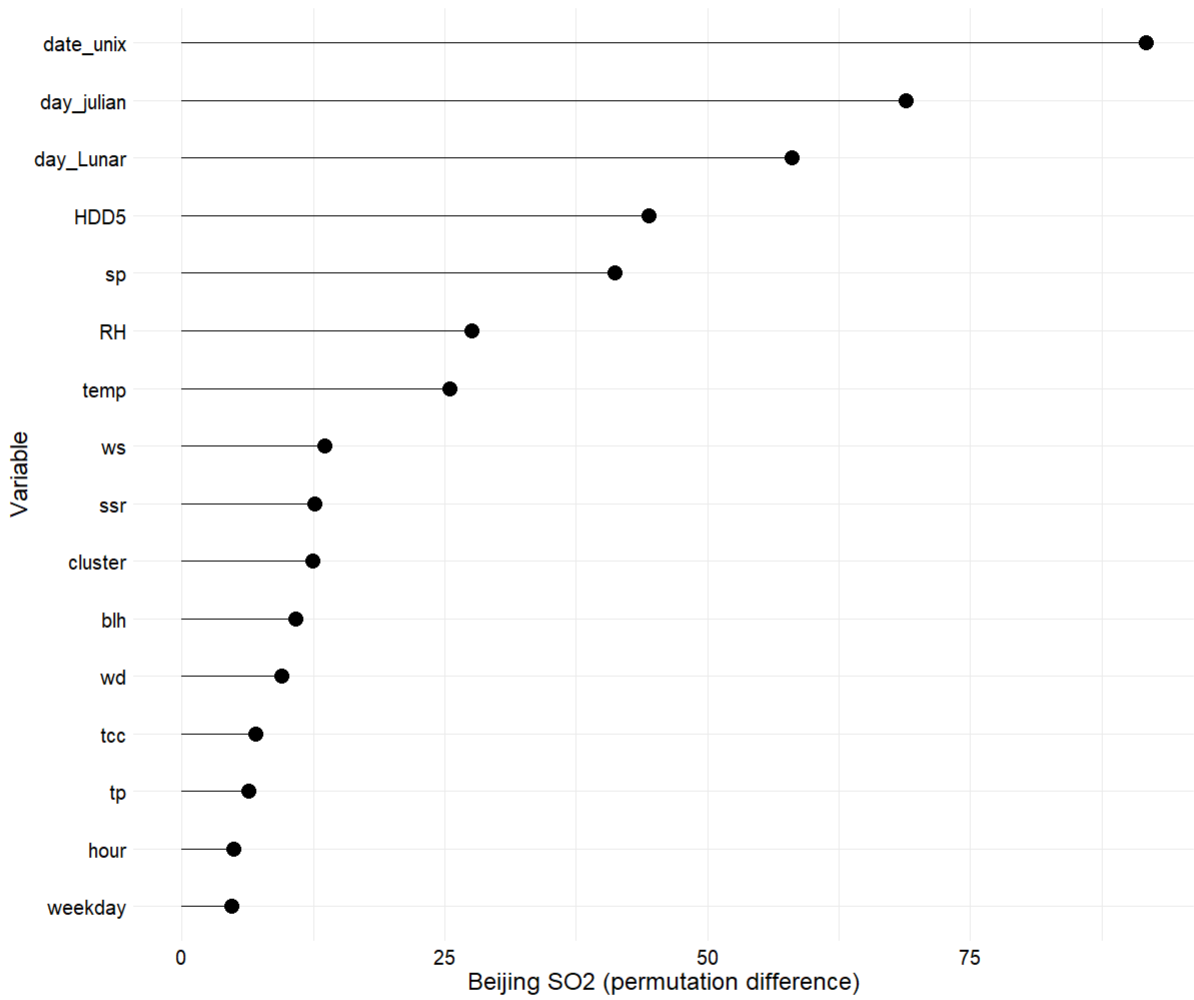


Figure 5.2: Variable importance of $PM_{2.5}$ in Beijing

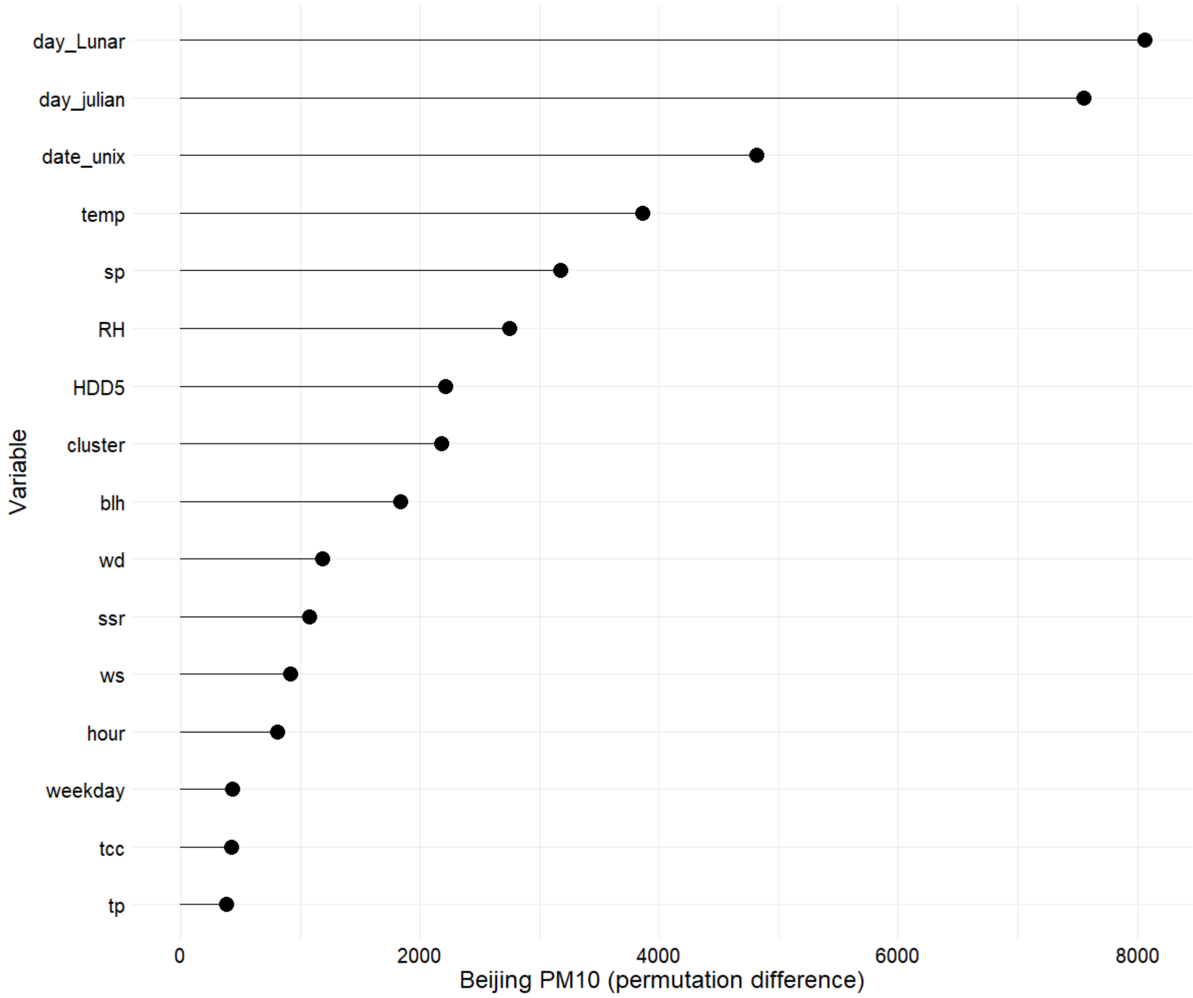


Figure 5.3: Variable importance of PM10 in Beijing

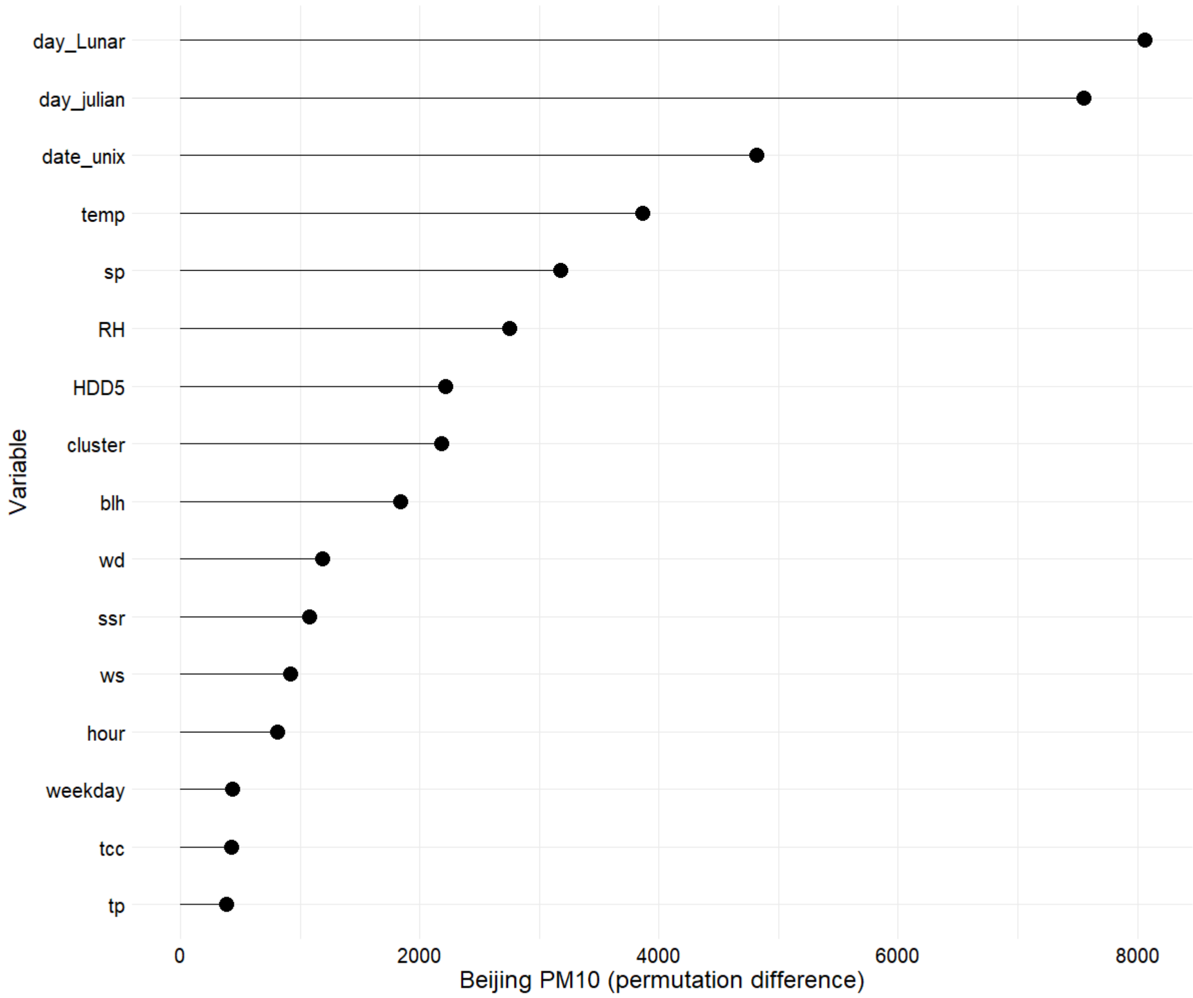
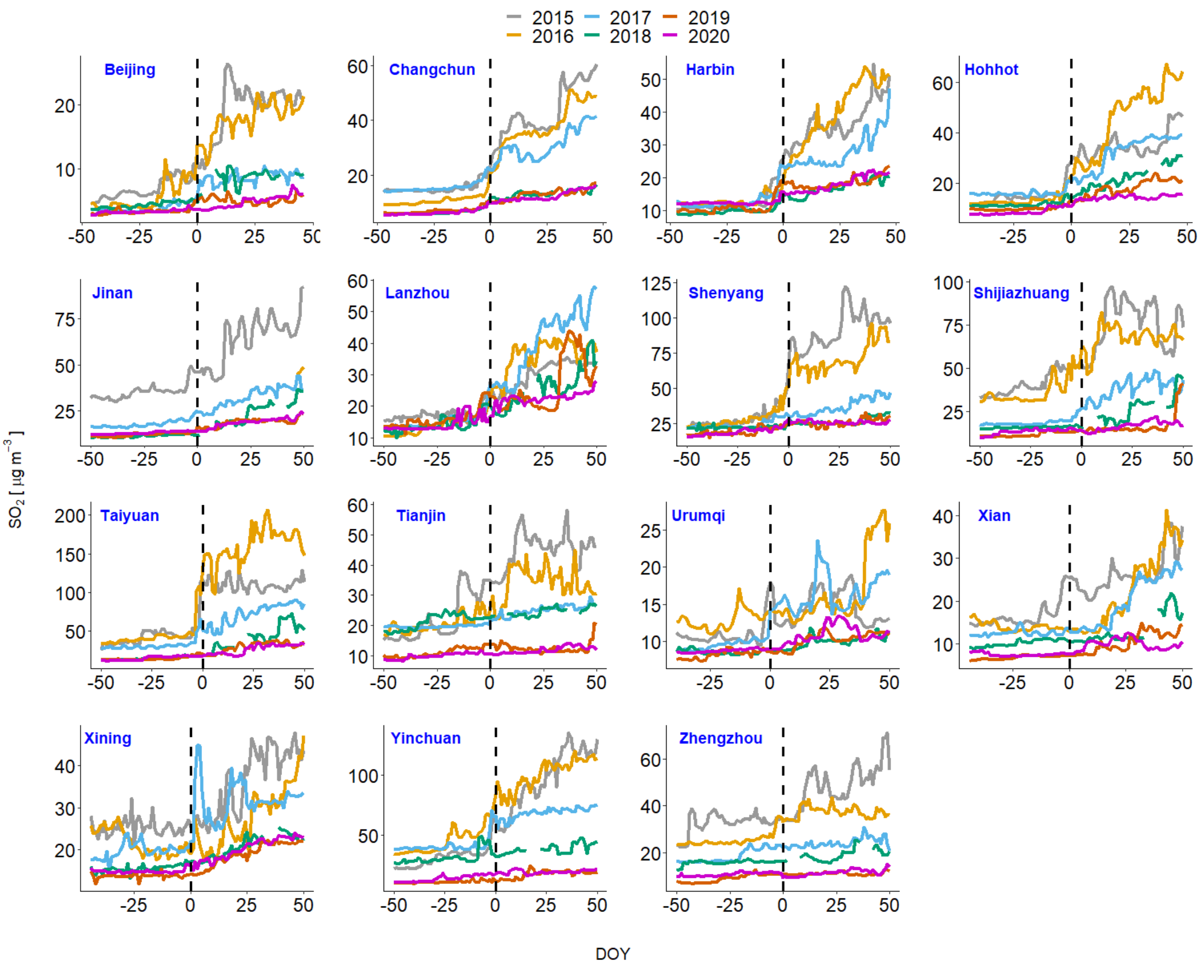
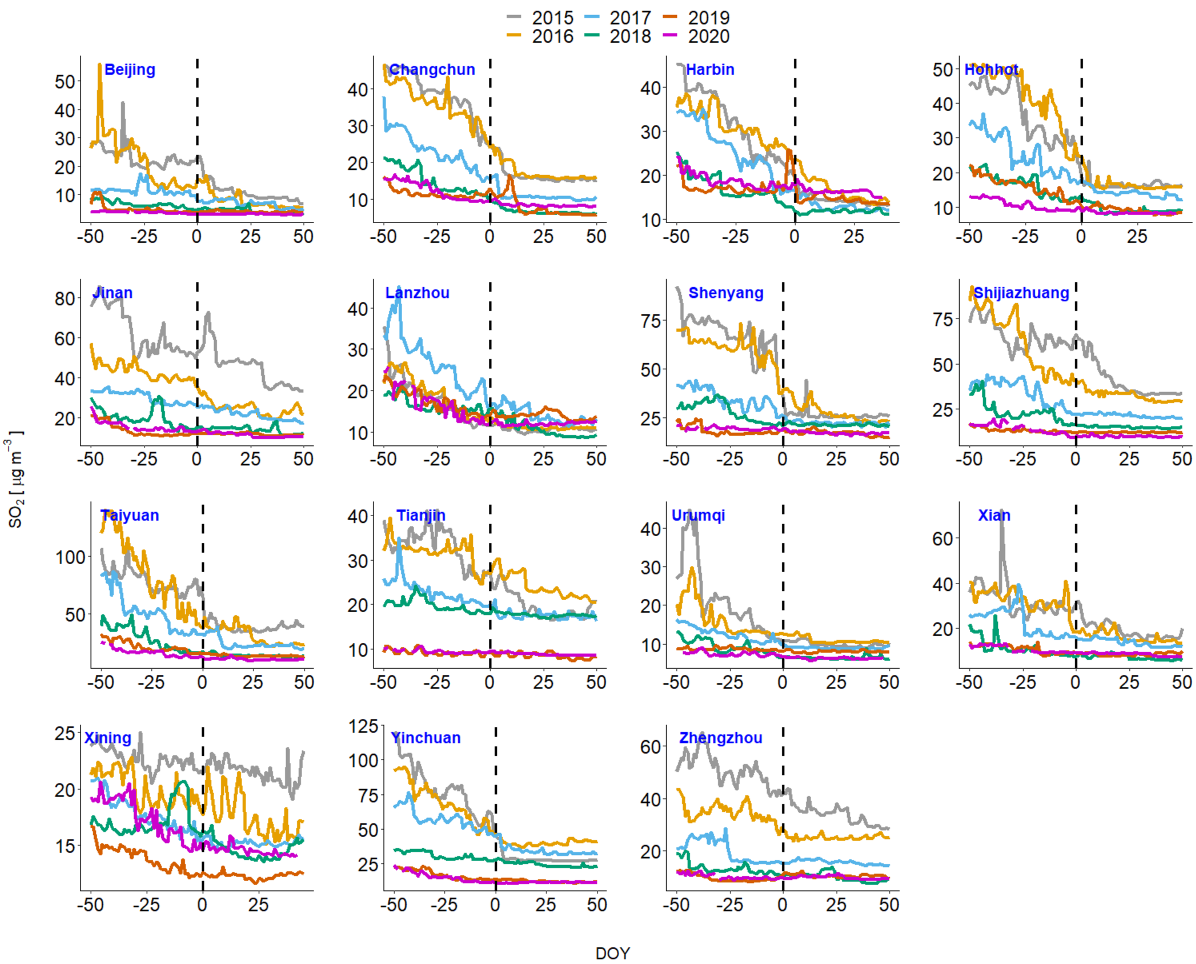


Figure 5.4: Weather normalised concentrations of daily SO_2 in 15 northern Chinese cities before and after the winter heating turn on between 2015 and 2021.



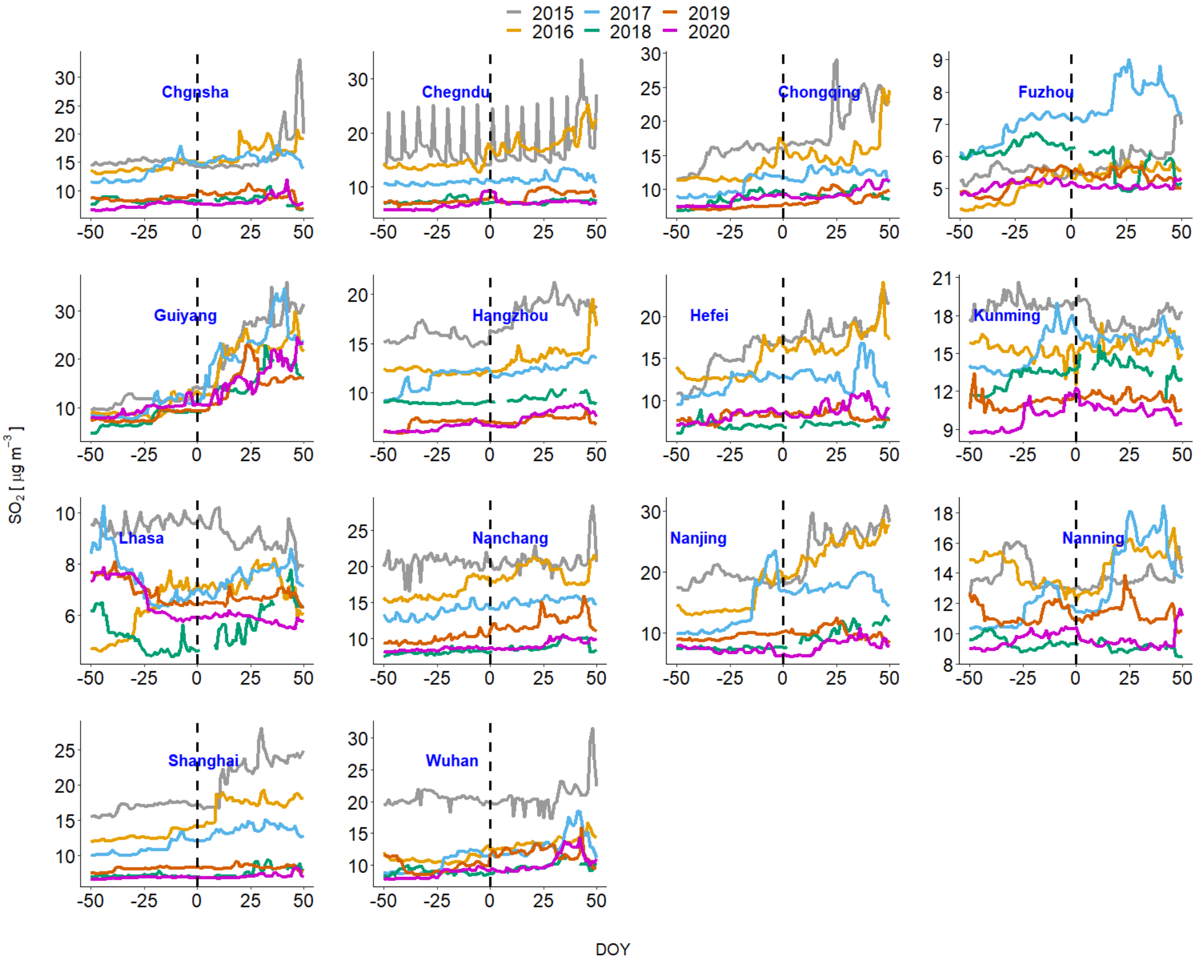
Note: The horizontal axis refers to the (50) days before and after the winter heating turn on date (indicated by the vertical lines),the vertical axis refers to the weather normalise concentrations of daily SO_2, six lines refers to the winter heating season in each year, i.e., 2015 refers to “Sep 2015 to May 2016“.

Figure 5.5: Weather normalised concentrations of daily SO_2 in 15 northern Chinese cities before and after the winter heating turn off between 2015 and 2021.



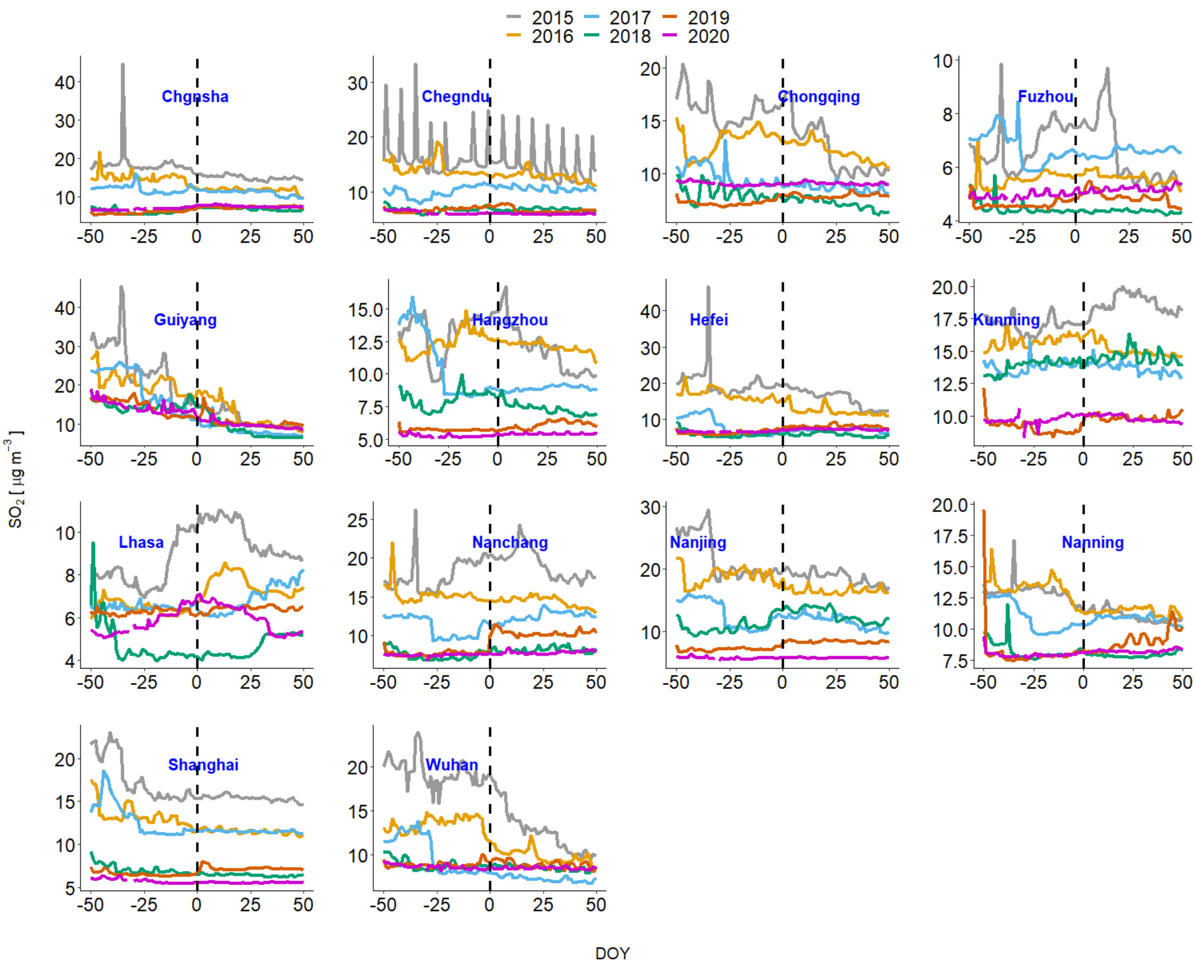
Note: The horizontal axis refers to the (50) days before and after the winter heating turn off date (indicated by the vertical lines),the vertical axis refers to the weather normalise concentrations of daily SO_2, six lines refers to the winter heating season in each year, i.e., 2015 refers to “Sep 2015 to May 2016“.

Figure 5.6: Weather normalised concentrations of daily SO_2 in 14 southern Chinese cities before and after the hypothetical winter heating turn on date between 2015 and 2021.



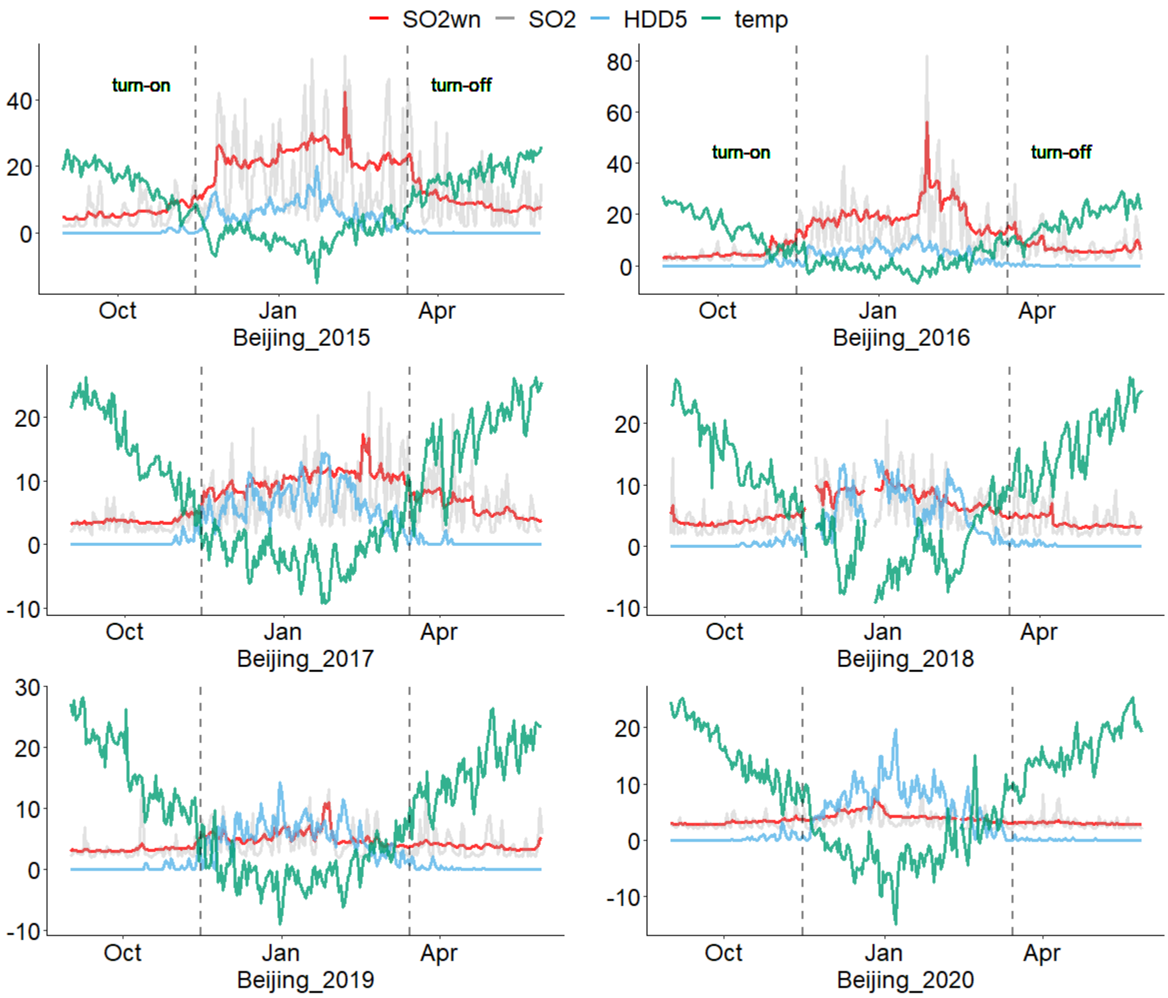
Note: The horizontal axis refers to the (50) days before and after the hypothetical winter heating turn on date (the vertical line in each figure refers to the date “15 Nov“ of each year),the vertical axis refers to the weather normalise concentrations of daily SO_2, six lines refers to the winter heating season in each year, i.e., 2015 refers to “Sep 2015 to May 2016“.

Figure 5.7: Weather normalised concentrations of daily SO_2 in 14 southern Chinese cities before and after the hypothetical winter heating turn off date between 2015 and 2021.



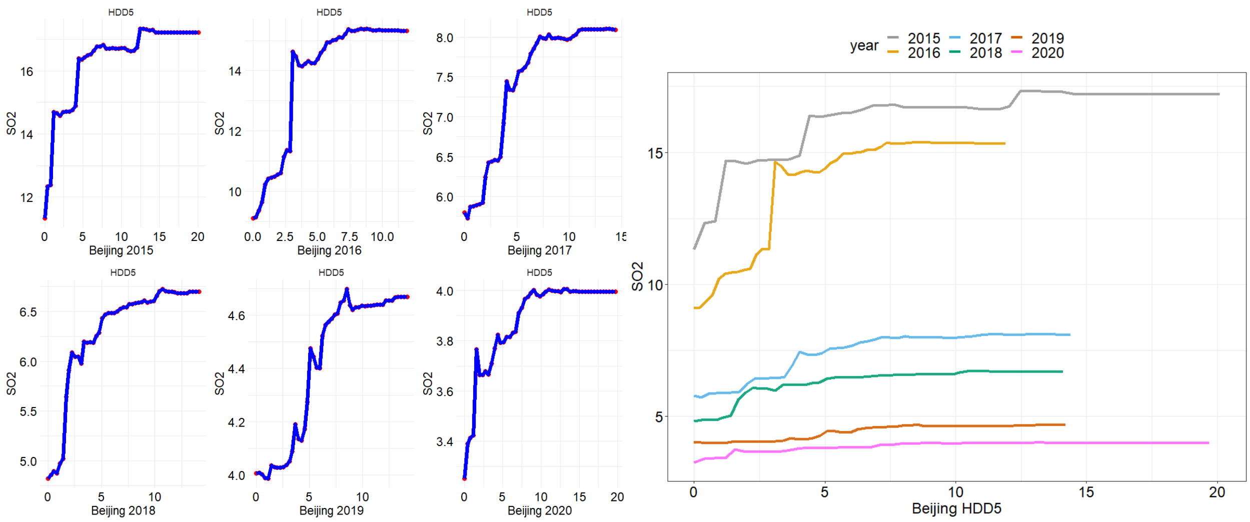
Note: The horizontal axis refers to the (50) days before and after the hypothetical winter heating turn off date (the vertical line in each figure refers to the date "15 Mar" of each year),the vertical axis refers to the weather normalise concentrations of daily SO_2, six lines refers to the winter heating season in each year, i.e., 2015 refers to "Sep 2015 to May 2016".

Figure 5.8: The trend plot of observed and weather normalised SO_2, HDD5 and temperature in Beijing between 2015 to 2021



Note: The horizontal axis refers to time period, the left and right vertical lines refer to the winter heating turn on and off date in each year, SO_2 and SO_2wn are daily observed and daily weather normalised concentrations of SO_2, respectively.

Figure 5.9: The HDD5-SO_2 response plot of Beijing in six heating seasons

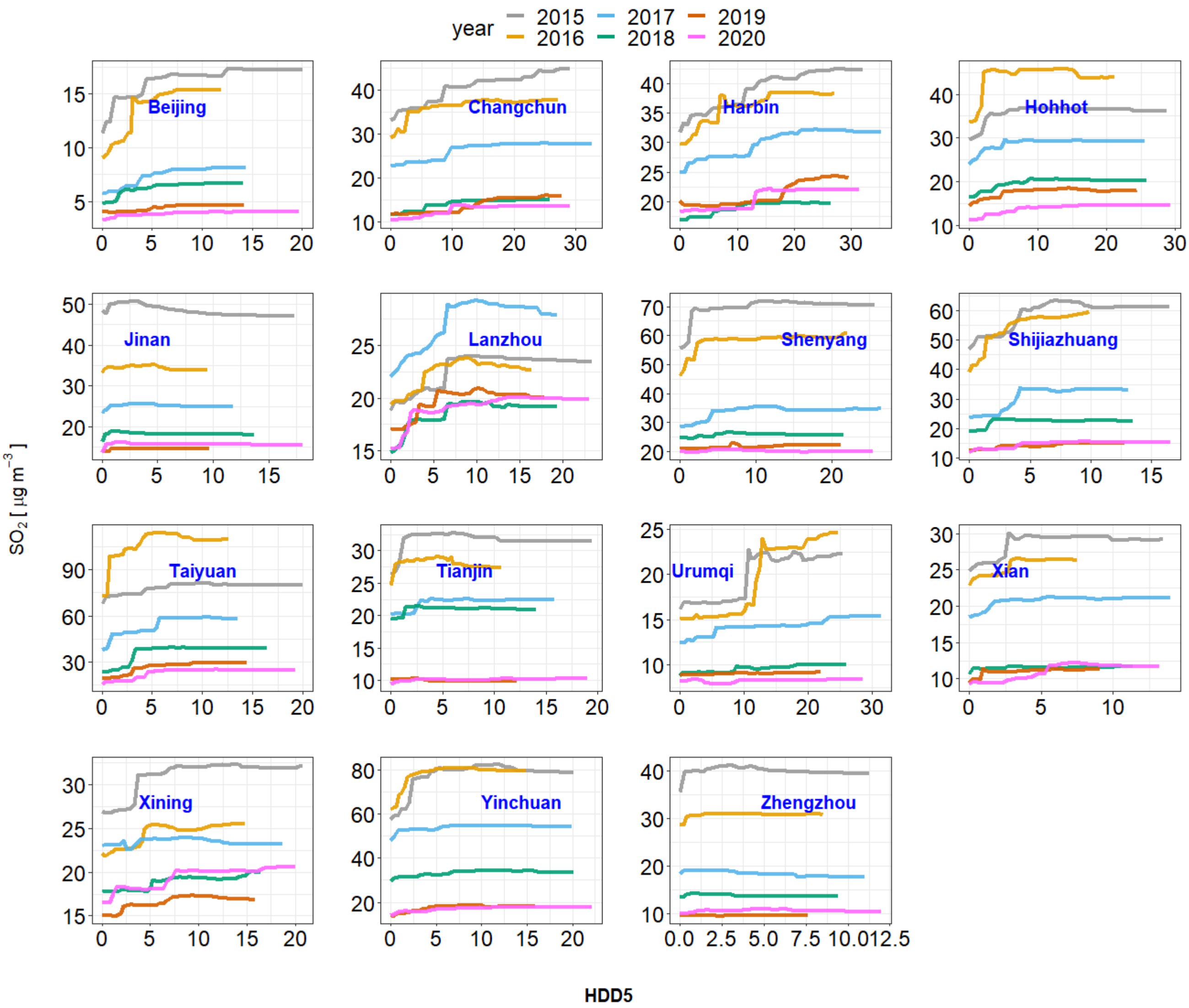


(a) 6 heating seasons separately

(b) 6 heating seasons in one plot

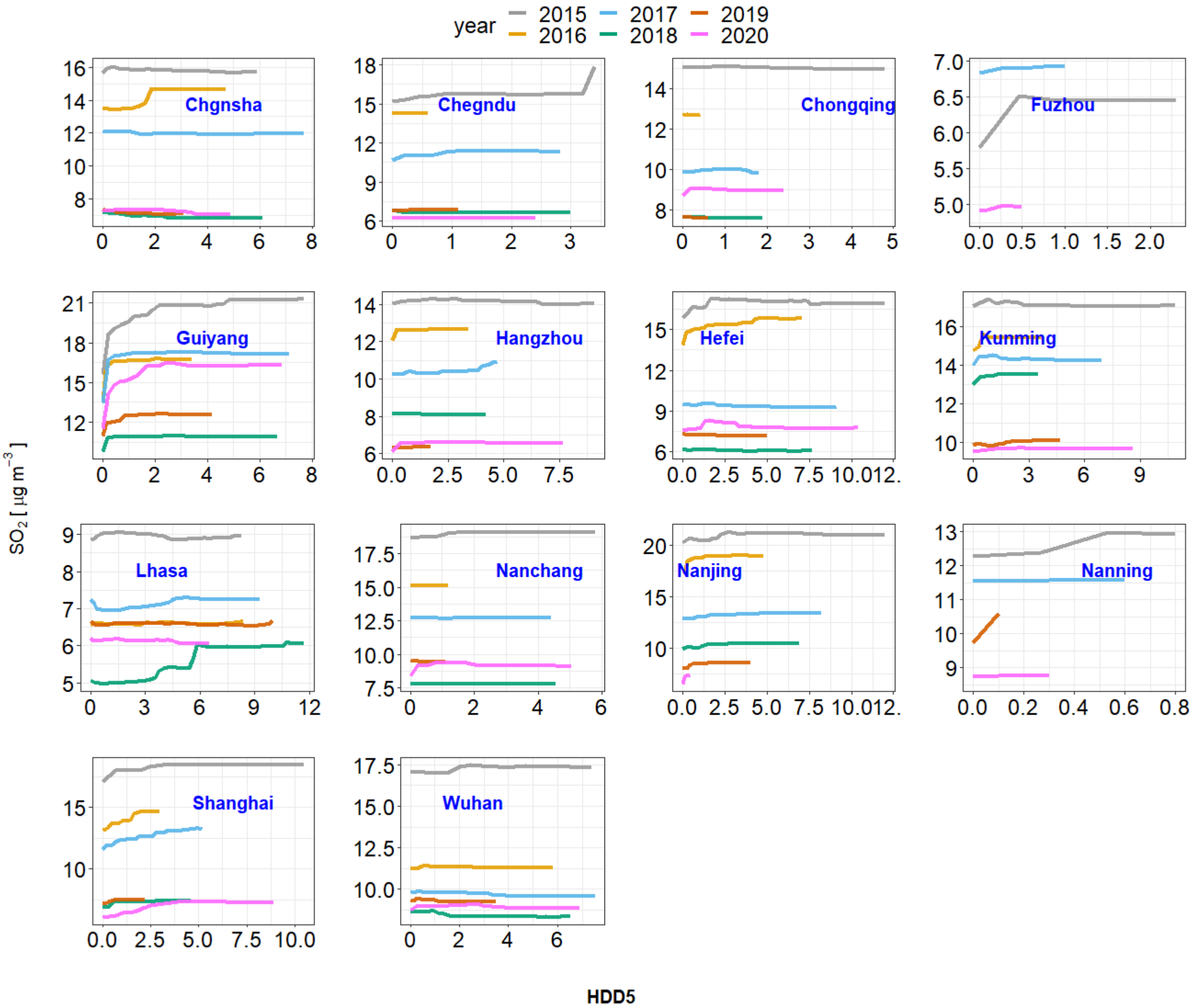
Note: The HDD5-SO_2 response was taken from the partial dependence plot within the ML model

Figure 5.10: The HDD5-SO_2 response plots in 15 northern cities between 2015 and 2021



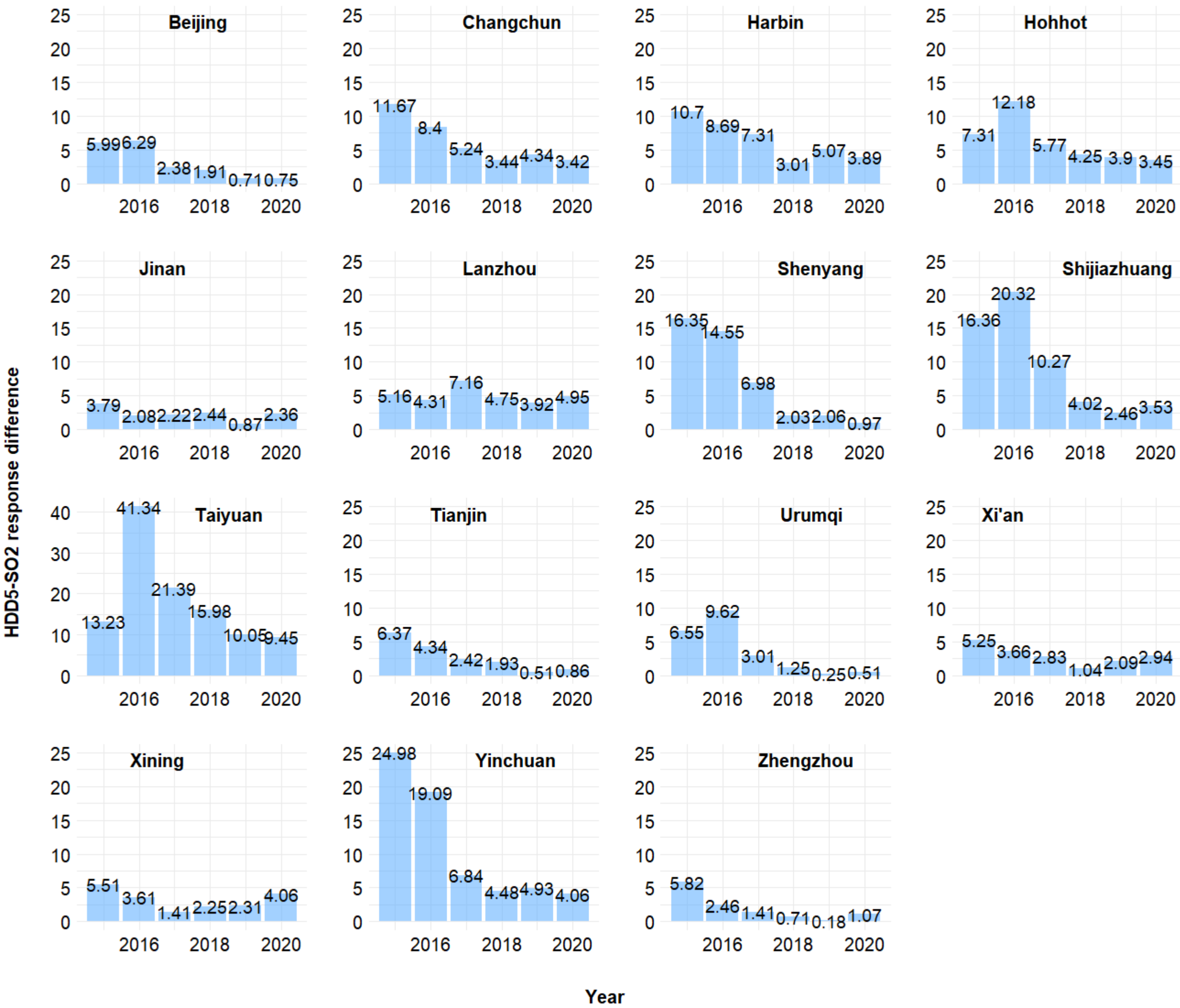
Note: The horizontal axis refers to the (50) days before and after the winter heating turn on date (indicated by the vertical lines),the vertical axis refers to the weather normalise concentrations of daily SO_2, six lines refers to the winter heating season in each year, i.e., 2015 refers to “Sep 2015 to May 2016“.

Figure 5.11: The HDD5-SO_2 response plots in 14 southern cities between 2015 and 2021



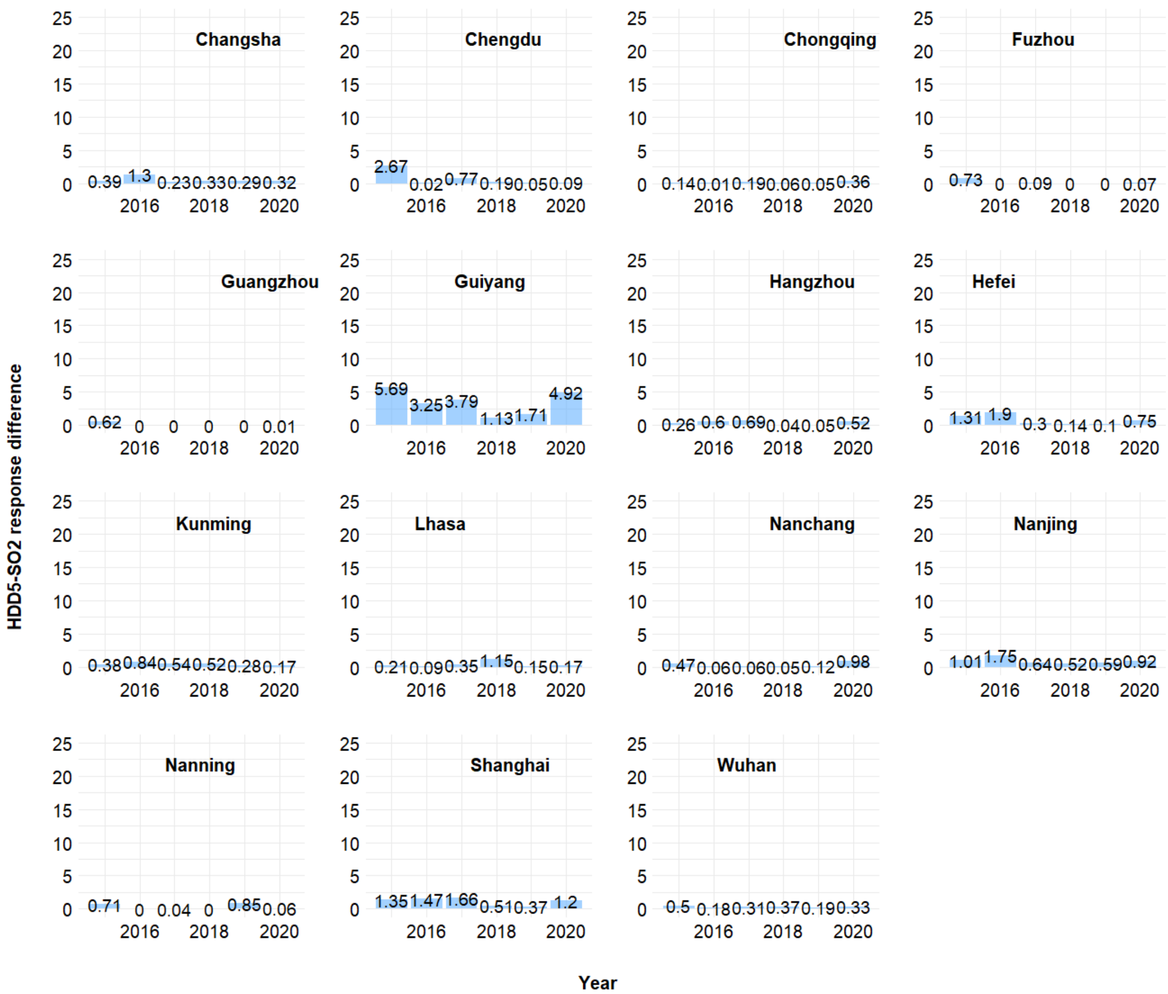
Note: The horizontal axis refers to the HDD5 value and the vertical axis refers to the SO_2 concentration levels. Lhasa belongs to northern city but did not have winter heating until winter season in 2020, thus included in the southern (non-heating) cities

Figure 5.12: The HDD5-SO_2 response difference plots in 15 northern cities between 2015 and 2021



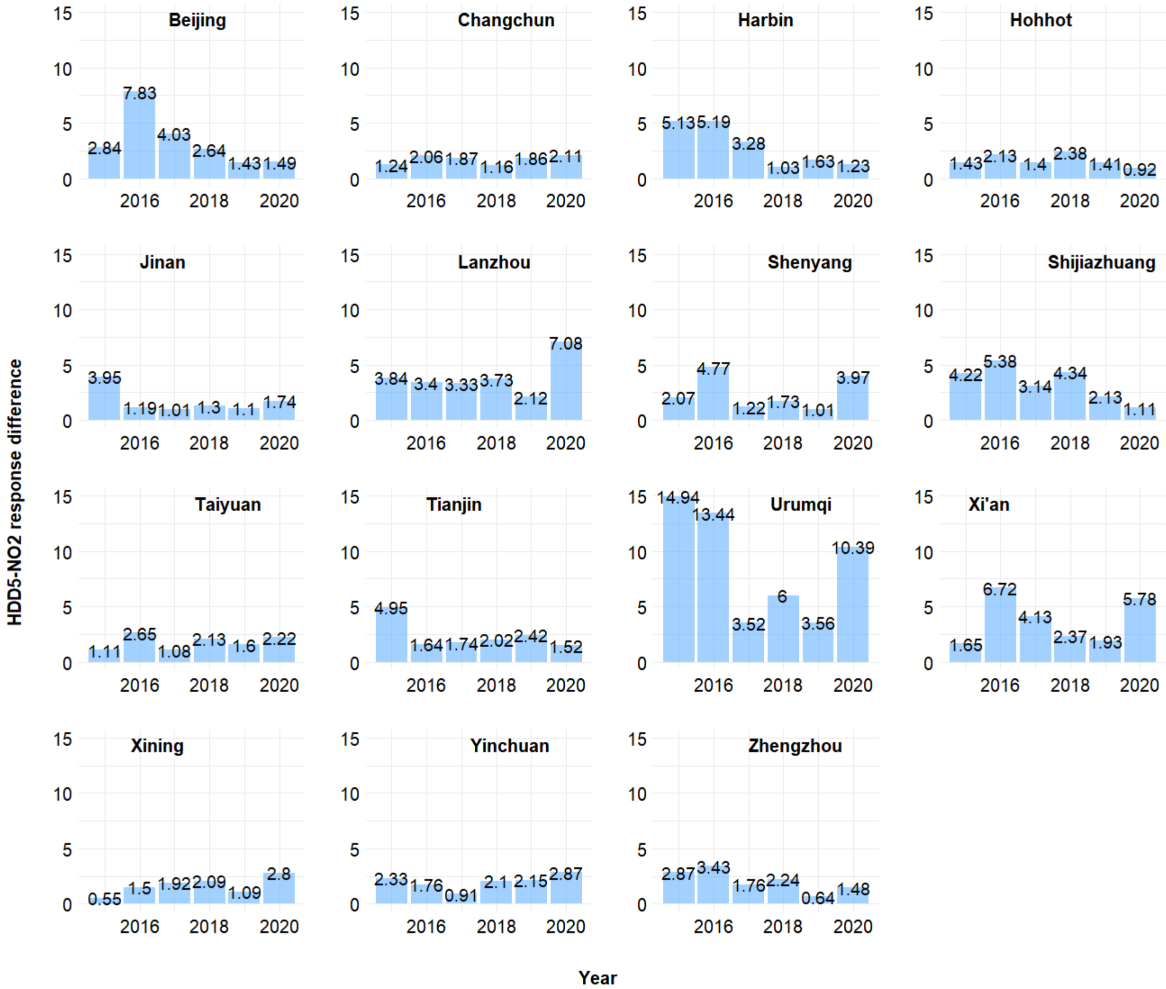
Note: The vertical axis refers to the difference of the maximum and minimum value of SO_2 levels of each year's HDD5-SO2 response plot

Figure 5.13: The HDD5-SO_2 response difference plots in 14 southern cities between 2015 and 2021



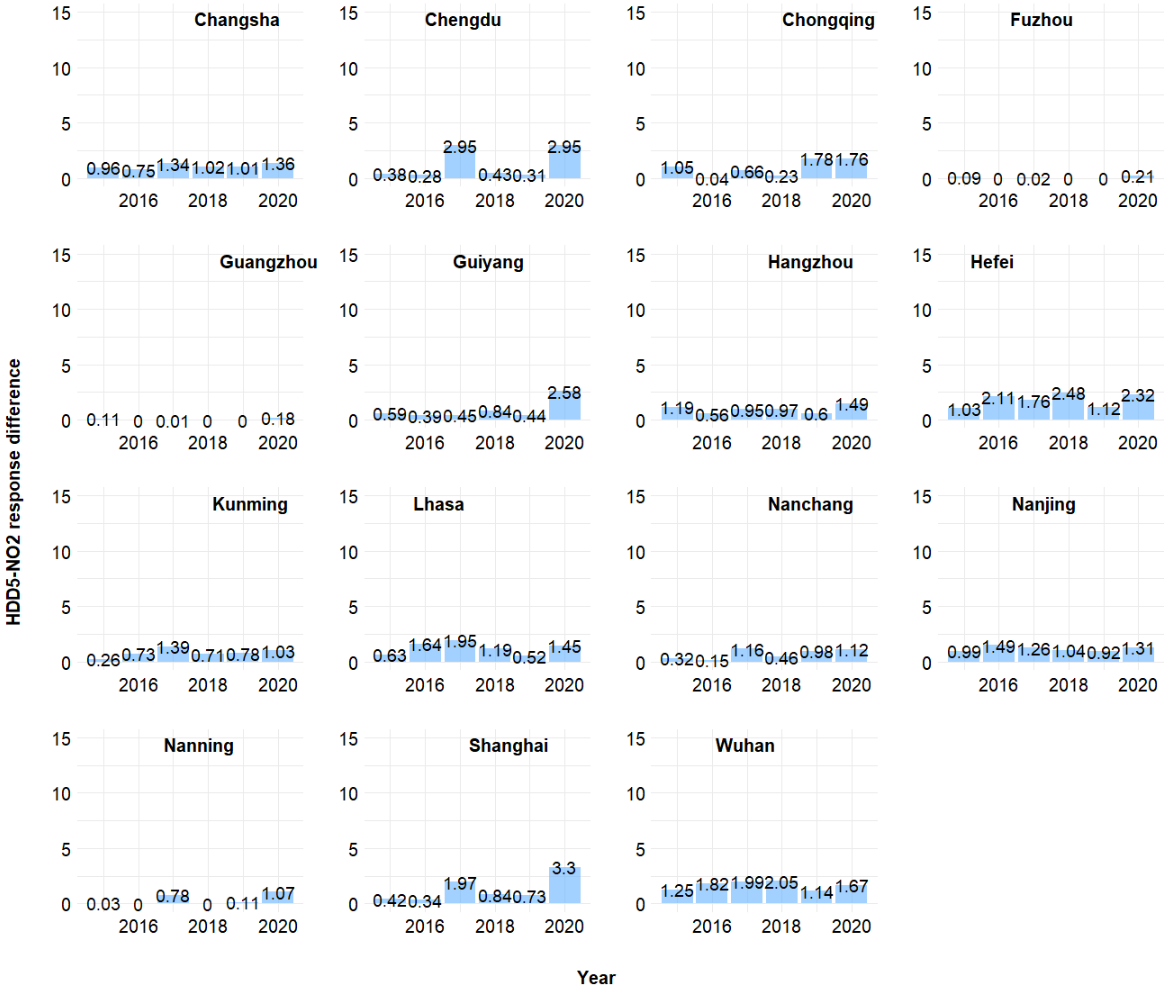
Note: The vertical axis refers to the difference of the maximum and minimum value of SO_2 levels of each year's HDD5-SO2 response plot.

Figure 5.14: The HDD5-NO_2 response difference plots in 15 northern cities between 2015 and 2021



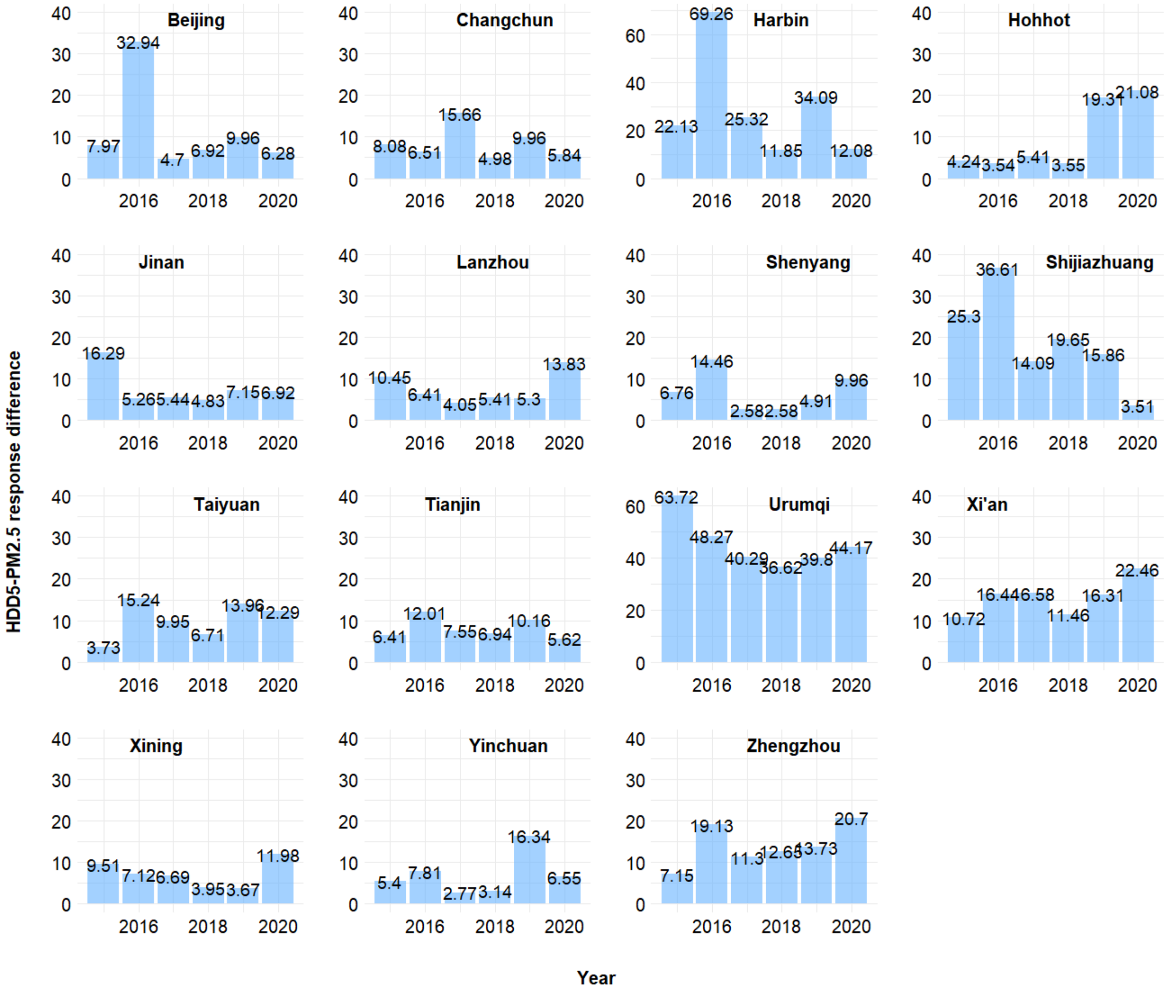
Note: The vertical axis refers to the difference of the maximum and minimum value of NO_2 levels of each year's HDD5-NO_2 response plot.

Figure 5.15: The HDD5-NO_2 response difference plots in 14 southern cities between 2015 and 2021



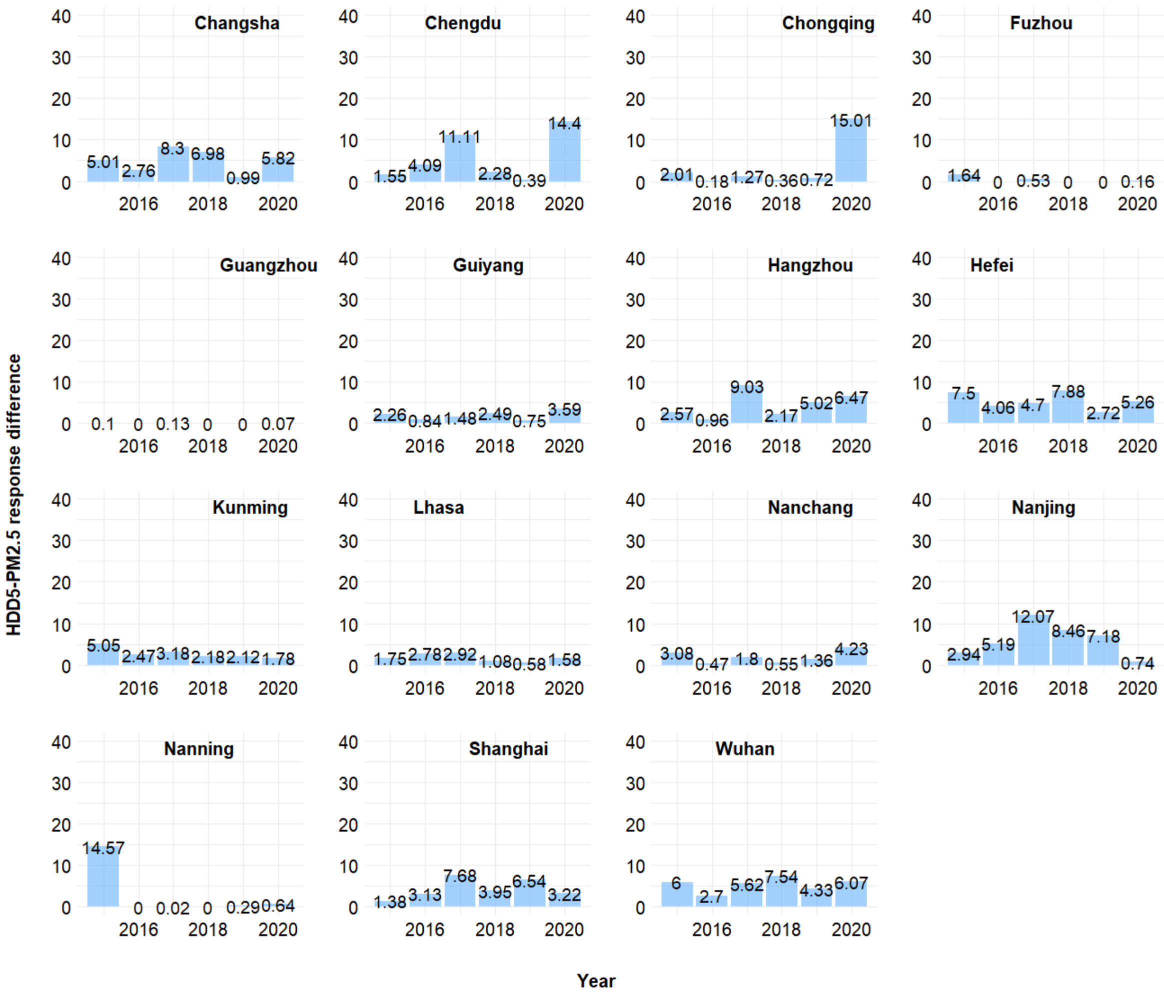
Note: The vertical axis refers to the difference of the maximum and minimum value of NO_2 levels of each year's HDD5-NO_2 response plot.

Figure 5.16: The HDD5-$PM_{2.5}$ response difference plots in 15 northern cities between 2015 and 2021



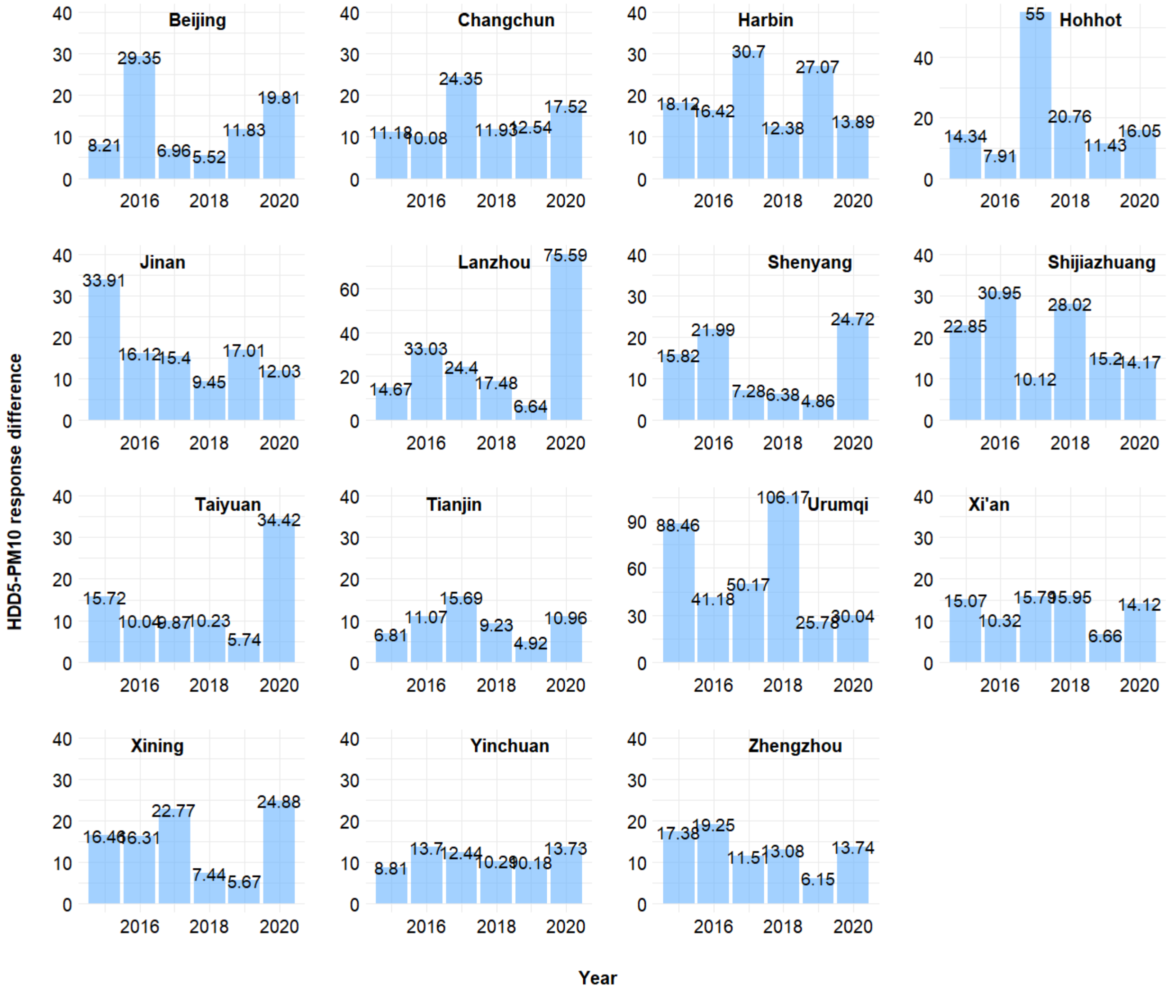
Note: The vertical axis refers to the difference of the maximum and minimum value of $PM_{2.5}$ levels of each year's HDD5-$PM_{2.5}$ response plot.

Figure 5.17: The HDD5-$PM_{2.5}$ response difference plots in 14 southern cities between 2015 and 2021



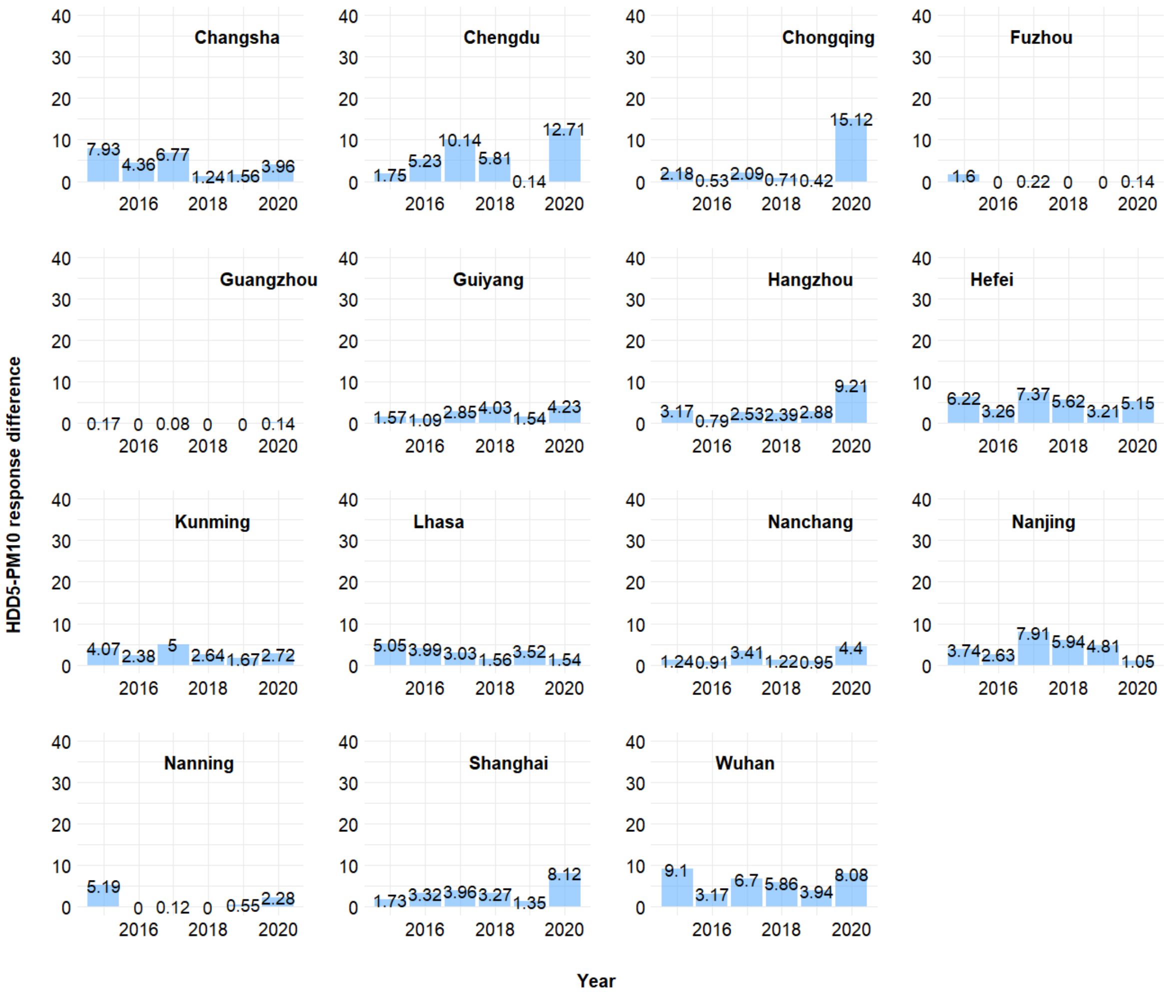
Note: The vertical axis refers to the difference of the maximum and minimum value of $PM_{2.5}$ levels of each year's HDD5-$PM_{2.5}$ response plot.

Figure 5.18: The HDD5-PM10 response difference plots in 15 northern cities between 2015 and 2021



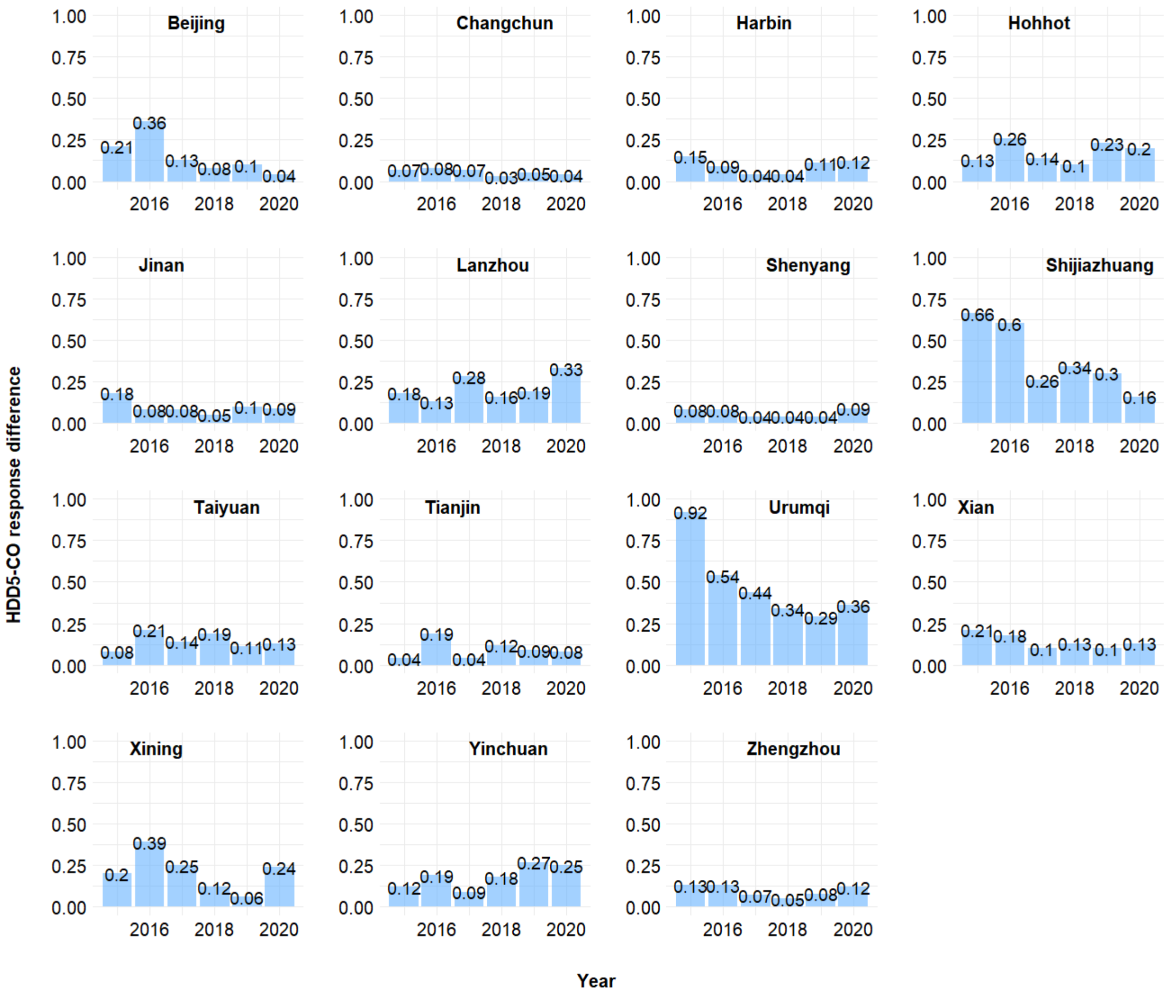
Note: The vertical axis refers to the difference of the maximum and minimum value of PM10 levels of each year's HDD5-PM10 response plot.

Figure 5.19: The HDD5-PM10 response difference plots in 14 southern cities between 2015 and 2021



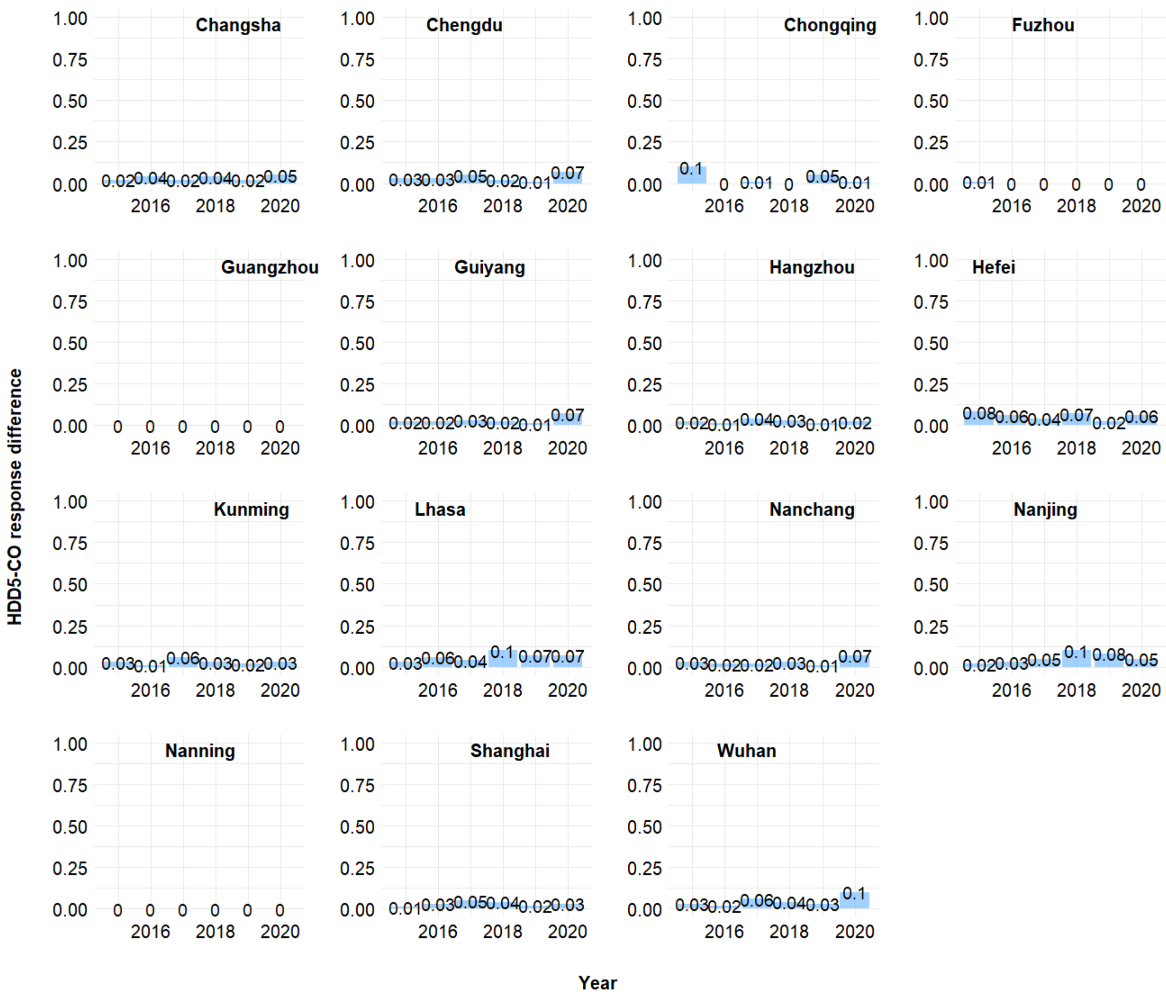
Note: The vertical axis refers to the difference of the maximum and minimum value of PM10 levels of each year's HDD5-PM10 response plot.

Figure 5.20: The HDD5-CO response difference plots in 15 northern cities between 2015 and 2021



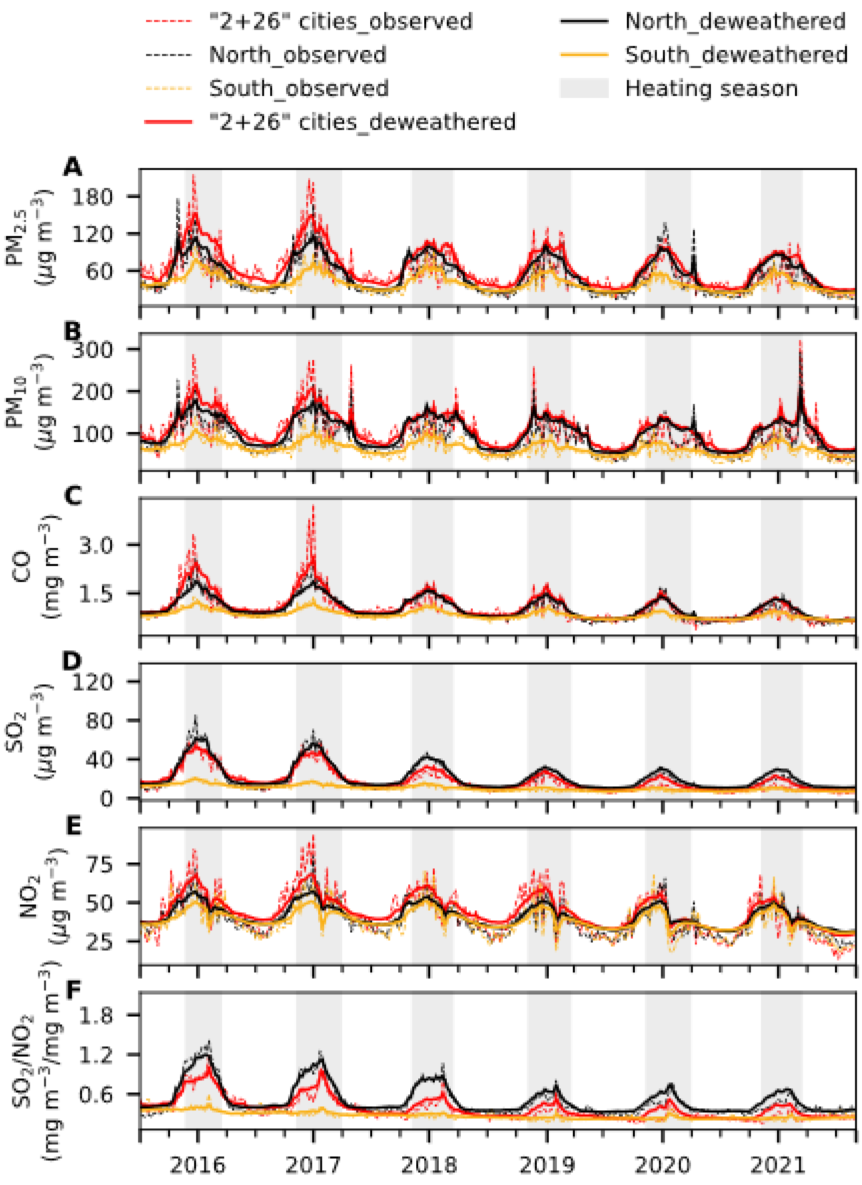
Note: The vertical axis refers to the difference of the maximum and minimum value of CO levels of each year's HDD5-CO response plot.

Figure 5.21: The HDD5-CO response difference plots in 14 southern cities between 2015 and 2021



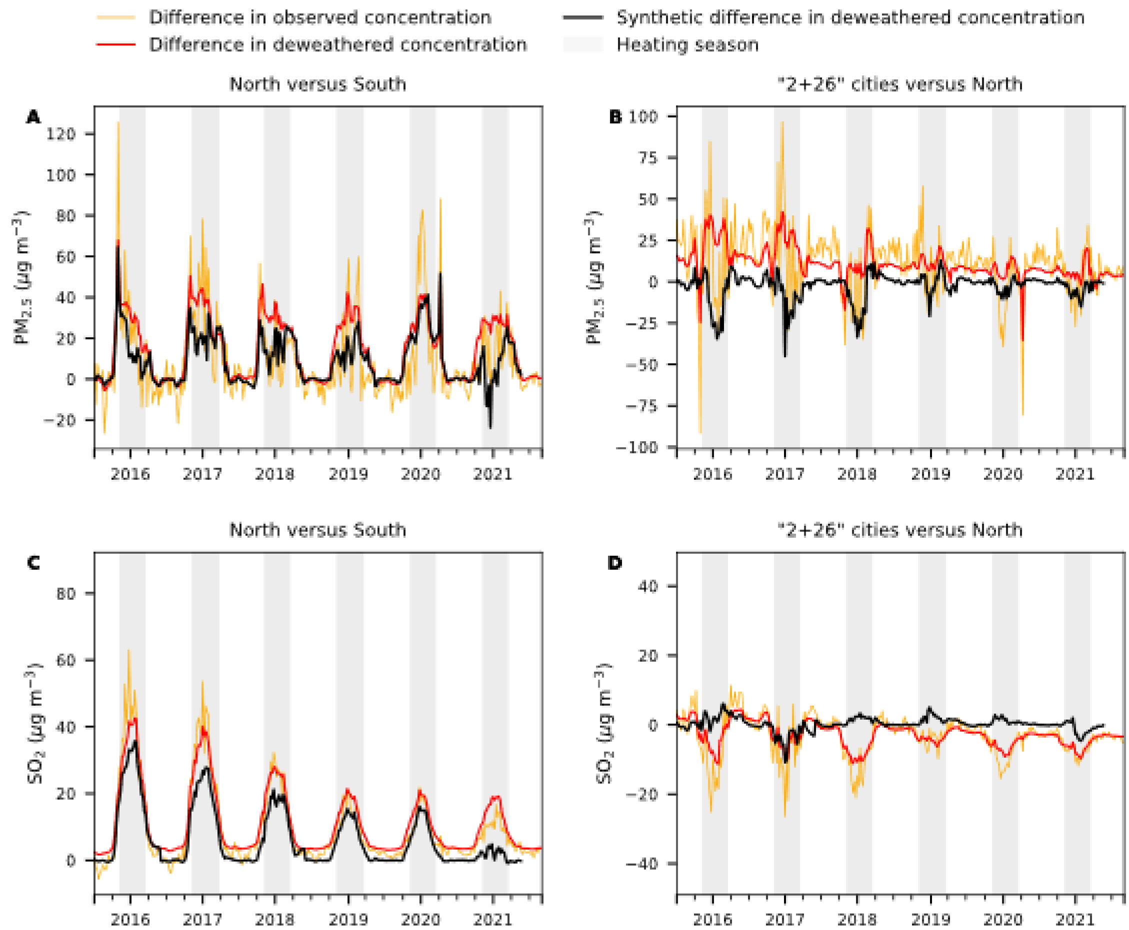
Note: The vertical axis refers to the difference of the maximum and minimum value of CO levels of each year's HDD5-CO response plot.

Figure 5.22: The weekly observed and deweathered PM2.5, PM10, CO, SO2, NO_2 concentrations and SO2/NO_2 ratios from 2015 to 2021 in the three regions.



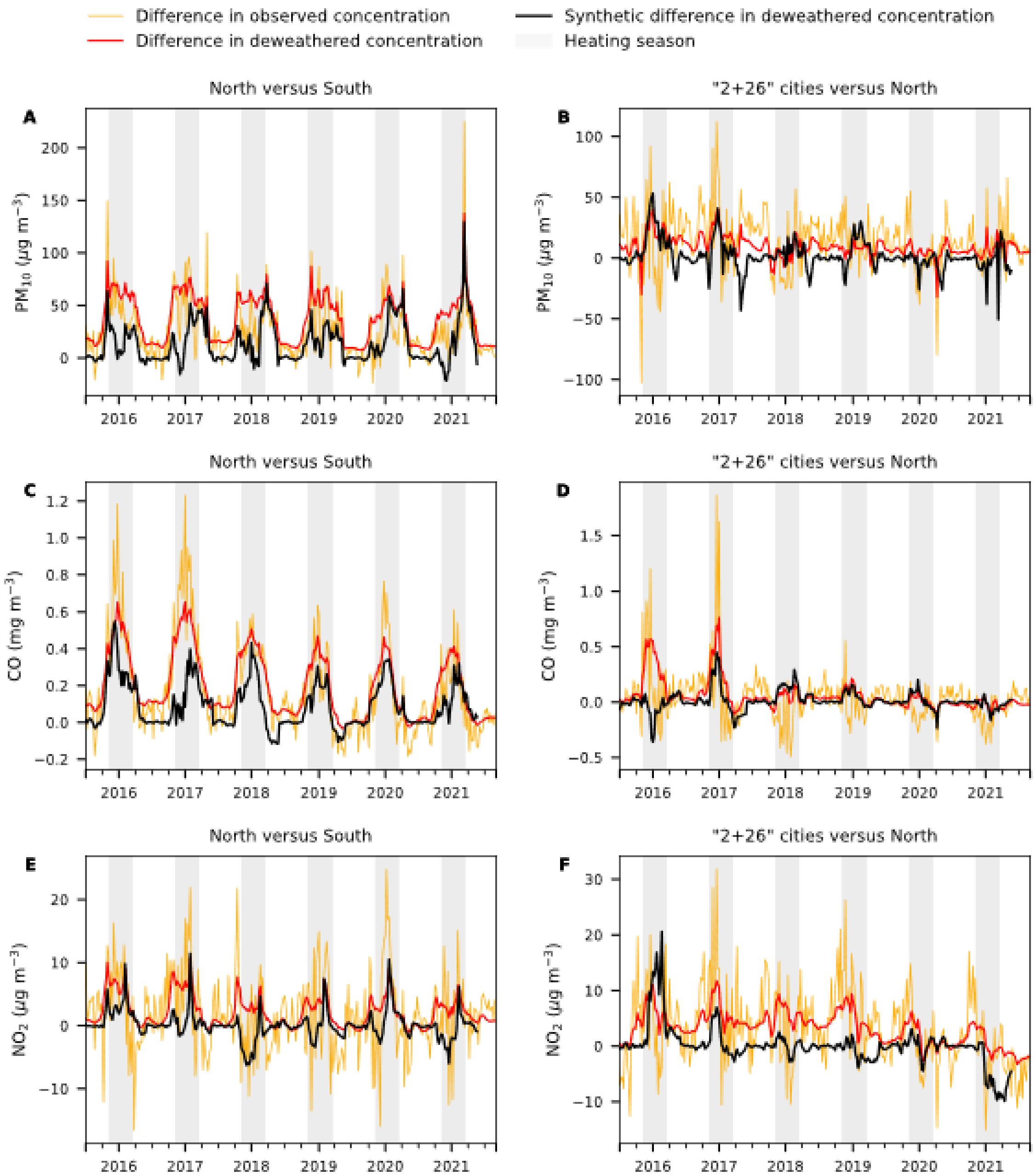
Note: The shaded area (i.e., heating season) is the period from 15th November of the current year to 15th March of the next year.

Figure 5.23: Observed, deweathered and synthetic differences in weekly PM2.5 and SO2 for different regions



Note:The shaded area (i.e., heating season) is the period from 15th November of the current year to 15th March of the next year.

Figure 5.24: Observed, deweathered and synthetic differences in weekly PM10, CO and NO_2 for different regions



Note:The shaded area (i.e., heating season) is the period from 15th November of the current year to 15th March of the next year.

Table 5.1: The causal impact of winter heating on SO2wn, $PM_{2.5}$wn and PM10wn in “2+26” and “North” in each heating season

		(1)	(2)	(3)
Time	Region	SO2wn (μg/m^3)	$PM_{2.5}$wn (μg/m^3)	PM10wn (μg/m^3)
2014	2+26	27.4 (13.8, 42.9)	8.8 (-36.6, 59.2)	32.3 (-37.5, 104.8)
	North	18.4 (9.6, 27.2)	26.4 (-11.9, 61.8)	6.1 (-19.3, 32.3)
2015	2+26	19.8 (11.5, 28.1)	-2.0 (-36.4, 34.8)	10.4 (-44.9, 62.0)
	North	16.2 (12.4, 18.8)	5.7 (-32.1, 47.8)	8.1 (-27.6, 53.8)
2016	2+26	12.5 (4.2, 23.2)	10.4 (-36.7, 62.5)	22.2 (-45.7, 96.9)
	North	-4.2 (-9.1, 1.1)	-3.2 (-23.1, 23.6)	-4.7 (-21.5, 10.3)
2017	2+26	6.5 (1.0, 12.1)	-10.9 (-35.3, 11.7)	-4.9 (-48.2, 38.8)
	North	2.8 (0.4, 5.8)	6.1 (-9.9, 25.6)	-12.6 (-23.6, 18.2)
2018	2+26	5.3 (1.4, 8.6)	2.0 (-24.1, 30.3)	5.8 (-19.8, 34.2)
	North	-3.4 (-4.2, -1.4)	-0.4 (-4.1, 5.1)	3.3 (-2.2, 11.5)
2019	2+26	2.1 (-0.4, 4.8)	11.2 (-12.4, 35.6)	22.5 (-25.8, 71.4)
	North	-2.5 (-4.1, -1.3)	5.7 (-1.9, 16.3)	26.2 (10.3, 45.4)
2020	2+26	0.2 (-2.9, 3.1)	-2.7 (-30.4, 26.1)	19.3 (-27.5, 67.9)
	North	-4.1 (-5.2, -2.3)	7.6 (-1.5, 19.0)	17.4 (6.7, 29.3)

Note: 95% CI in the brackets. 2014 refers to “June 2014 to May 2015”, 2015 refers to “June 2015 to May 2016”, 2016 refers to “June 2016 to May 2017”, 2017 refers to “June 2017 to May 2018”, 2018 refers to “June 2018 to May 2019”, 2019 refers to “June 2019 to May 2020”,2020 refers to “June 2020 to May 2021”.

Table 5.2: The causal impact of winter heating on NO_2 wn, COwn and O3wn in “2+26” and “North” in each heating season

		(1)	(2)	(3)
Time	Region	NO_2 wn ($\mu g/m^3$)	COwn (mg/m^3)	O3wn ($\mu g/m^3$)
2014	2+26	2.6 (-6.3, 11.2)	0.5 (-0.1, 1.0)	-9.2 (-49.0, 30.3)
	North	-2.1 (-4.7, 0.9)	0.1 (0.0, 0.3)	-3.6 (-6.3, -0.2)
2015	2+26	3.3 (-7.4, 13.6)	0.5 (0.0, 0.9)	-1.6 (-39.0, 34.4)
	North	-1.9 (-3.5, 0.8)	0.2 (0.1, 0.4)	-6.5 (-8.3, -3.5)
2016	2+26	-1.6 (-20.0, 19.9)	0.3 (-0.1, 0.7)	13.3 (-22.2, 44.3)
	North	-0.9 (-2.3, 0.9)	0.0 (-0.1, 0.1)	6.0 (3.0, 9.1)
2017	2+26	-6.8 (-19.9, 7.2)	0.0 (-0.2, 0.3)	-2.6 (-56.4, 51.9)
	North	1.3 (-1.1, 2.7)	0.0 (-0.1, 0.1)	1.9 (-1.0, 4.6)
2018	2+26	2.3 (-12.3, 17.4)	0.1 (-0.2, 0.3)	2.0 (-48.8, 59.9)
	North	-0.6 (-1.6, 0.6)	-0.1 (-0.1, 0.0)	2.5 (-1.0, 6.0)
2019	2+26	-1.4 (-12.2, 9.4)	0.1 (-0.1, 0.3)	4.1 (-54.6, 62.9)
	North	0.5 (-0.5, 1.7)	0.1 (0.0, 0.2)	2.1 (-1.3, 5.4)
2020	2+26	-4.6 (-17.3, 9.5)	0.0 (-0.2, 0.2)	7.0 (-35.6, 50.3)
	North	1.8 (0.1, 2.9)	0.1 (0.1, 0.2)	3.0 (0.7, 5.2)

Note: 95% CI in the brackets.2014 refers to “June 2014 to May 2015”, 2015 refers to “June 2015 to May 2016”, 2016 refers to “June 2016 to May 2017”, 2017 refers to “June 2017 to May 2018”, 2018 refers to “June 2018 to May 2019”, 2019 refers to “June 2019 to May 2020”,2020 refers to “June 2020 to May 2021”.

Chapter Six

Conclusions

This thesis has evaluated the effect of three government policies on air quality in China; the environmental inspection policy, the Covid-19 lockdown policy, and the clean winter heating policy. In terms of the research methodology, this thesis has used state of the art econometric techniques with machine learning models and argued that the combination of both methods can produce more reliable and rigorous analyses.

This thesis has four chapters (independent papers) that looked at each of these policies separately.

Chapter 2 utilised econometric techniques to investigate the effectiveness of China's Environmental Protection Inspection program implemented in 2016, while Chapter 3 combined a machine learning algorithm with econometric techniques to quantify the impact of the Wuhan Covid-19 lockdown on concentrations of four air pollutants in 2020. Chapter 4 applied similar methods as chapter 3 to evaluate the inspection policy (from chapter 2) again, and finally Chapter 5 used the machine learning model to understand the impact of the winter heating system and the clean winter heating policy (implemented since 2017) on air quality in China.

Using an air quality index (AQI) and its 6 constituent pollutants, and focusing on the first inspection (January 2016) in Hebei province that acts as a quasi-natural experiment, Chapter 2 found that, relative to a control group, there was a small decrease in the recorded AQI in the six months before the inspection period, a fairly large drop during the month of the inspection, and a longer-term effect of up to five months after the inspection team left Hebei. The results are robust to various sensitivity checks and placebo tests.

Chapter 3 quantified the impact of the Wuhan Covid-19 lockdown on concentrations of four air pollutants using a two-step approach. We first used machine learning to remove the confounding effects of weather conditions on pollution concentrations, we then use the Augmented Synthetic Control Method to estimate the impact of the lockdown on weather normalised pollution relative to a control group of cities that were not in lockdown. We found NO_2 concentrations fell by as much as 24 μg/m^3 during the lockdown (a reduction of 63% from the pre-lockdown level), while PM10 concentrations fell by a similar amount but for a shorter period. The lockdown had no discernible impact on concentrations of SO_2 or CO.

Chapter 4 examined how effective the recent Central Environmental Inspection Policy (CEIP) was in reducing local air pollution at the regional and city level. Using similar methods as chapter 3, we found that the Hebei inspection led to a substantial, but short term, reduction in concentrations of PM2.5, PM10 and SO_2 in Hebei as a whole. The weather normalised value of PM2.5 and PM10 fell by as much as 25.89 μg/m^3 (20.81%) and 47.80 μg/m^3 (22.61%) immediately after the inspection, although pollution levels gradually returned to previous levels within 3 months.

Finally, Chapter 5 estimated the impact of the winter heating policy and the effectiveness of the clean winter heating plan on air quality in Chinese capital cities. After decoupling the impact of meteorological conditions on observed air pollutant concentrations, this chapter confirmed the immediate increase and reduction of air pollution when the heating system turns on and off, respectively, in 15 northern cities. Most importantly, we found that the adverse effects of the winter heating policy have considerably reduced since 2017 due to the shift towards cleaner sources of energy resulting from the clean winter heating plan.

To sum up, this thesis provides an evaluation of three selected policies in China and demonstrates that the increased attention from both central and local government as well as the general public contributed a lot to the air quality improvement recently. This thesis also adds to the literature documenting China's recent 'war on pollution (Greenstone, He et al., 2021)' since 2014, confirming the effectiveness of the policies and regulations implemented by the Chinese government to protect the environment. China's experience of tackling pollution can provide valuable lessons for many developing countries in the world.

Regarding future research, there are several directions that might be interesting to explore. For example, understanding the negative consequences of air pollution remains important for both academic researchers and policy makers. While the impact of air pollution on human physical health is well known, it is also important to look at the air pollution impact on other measures of social welfare, for example, the mental health, productivity, human capital accumulation, ecosystem service, natural capital, etc. As the data quality and availability of these measurements improve gradually, isolating both short term and long term consequences of air pollution on various measurements can demonstrate the 'true' cost of air pollution, which may have been previously underestimated within the existing literature (Greenstone, He et al., 2021).

More importantly, interacting Econometrics with machine learning models can be very fruitful for future studies. As a statistical analysis tool, machine learning is becoming increasingly used by social scientists, especially economists. The research methods that combine both machine learning and Econometrics can be powerful tools and there are already machine learning models that can address causal inference within Econometrics, suggesting a useful potential avenue for future research.

Appendix One

Chapter 2 Appendices

Appendix A.1 The detailed operations and measures of the CEIP inspection [1]

1. Enterprise production.

1.1. Industry type and main products, products and capacity of the previous month, whether production lines are under operation;

1.2. Whether there exist illegal construction and whether the investigation results are consistent with the Environmental Impact Assessment;

1.3. Whether there are newly produced waste gas, waste water, solid waste and other pollutants;

2. Implementation of environmental protection in enterprises.

2.1. Examine whether the project performs the EIA (Environmental impact assessment) procedures according to the law, check the EIA documents and the review of the EIA, whether the EIA documents involves counterfeiting;

2.2. Examine whether the nature of the project, the scale of production, the location, or the measures used for pollution control are consistent with the EIA and the approval doc-

[1](source: http://www.tiandeyi.com/)

uments; Whether the enterprise' pollution control facilities are put into production without the acceptance of the environmental protection department;

2.3. Examine whether the enterprise has been approved by the Environmental Protection Completion acceptance, whether relevant procedures are complete, after the project being put into operation; Examine whether there exist pollution control facilities being put into production without acceptance by local environmental protection departments;

2.4. Examine production workshop: whether proper anti-corrosion treatment has been conducted and regularly maintained in Raw materials related to acid, alkali, and other corrosive workshop floor; whether there exist leakage fields during the production process have;

2.5. Examine and inspect the application of a pollution discharge license, the execution of a discharge declaration and the payment of a discharge fee;

3. Inspection of waste gas pollution

Check and examine the running state, historical operation state, processing capacity and quantity of the waste gas treatment facilities; Make sure emission standards are met and pollution control measures are strengthened.

4. Air pollution prevention and control facilities

4.1. Examine for dust-removing, desulphurization, denitrification, and other gaseous pollutants purification systems;

4.2. Examine whether the waste gas emission meet the standard;

4.3. Examine whether polluter has been building new exhaust cylinder in the prohibited areas;

4.4. Examine whether the height of the exhaust cylinder meets the requirements of the national or local pollutant emission standards;

4.5. Examine whether the sampling hole and sampling monitoring platform are set on the exhaust pipes;

4.6. Examine whether the exhaust outlet settings meet the requirements, including height, sampling port and sign; whether online monitoring facilities are installed and implemented according to requirements;

5. Unorganized emission source

5.1. For unorganized emission of toxic and harmful gases, dusts and soot, Examine whether conditional emission is organized, check whether emitters have been regulated;

5.2. Examine whether the coal yard, stock yard, goods yard take measures to prevent or control the waste air pollution according to the requirements or set up the anti-dust equipment;

5.3 Monitor at the boundary of the enterprise to check whether the unorganized emission meets the requirements of the relevant environmental protection standards.

6. Inspection on waste gas emitted from Boiler, Petrochemical, Chemical Combustion (key industries of pollution control)

6.1. Examine whether measures taken by chemical and petrochemical enterprises to deal with continuously combustible organic waste gas meets requirements and whether the way of recycling or incineration is reasonable;

6.2. Examine the verification procedures and performance of the equipment for boiler combustion, examine the operation conditions of the combustion equipment. Examine the conditions of controlling for sulfur dioxide and nitrogen oxides;

6.3. Examine whether waste gas emissions, dust, and odor emissions meet the requirements of the related pollution emission standards;

6.4. Examine the environmental protection measures for the transport, handling and storage of volatile toxic, harmful gas and dust;

7. Warnings for other enterprises

7.1. The enterprises are not allowed to dismantle, idle or shut down the pollution treatment facilities and places without the approval of the environmental protection department;

7.2. The enterprises are prohibited to violate the setting regulation of emission outlet, to secretly install emission pipe and modify monitoring data with falsification;

7.3. The enterprises were not allowed to construct in prohibited areas, meanwhile the project construction should comply with planning.

The MEP's AQI approach is based on the Chinese Ambient Air Quality Standards (CAAQS). The sub-AQI of each pollutant is calculated using Ea. (1) with the monitored pollutant concentrations:

$$AQI_i = \frac{(AQI_{i,j} - AQI_{i,j-1})}{(C_{i,j} - C_{i,j-1})} \times \left(C_{i,m} - C_{i,j-1}\right) + AQI_{i,j-1}, \text{ j>1}$$

$$AQI_i = AQI_{i,1}\frac{C_{i,m}}{C_{i,1}}, \text{ j=1} \qquad (1)$$

where AQI_i is the index of pollutant i; $C_{i,m}$ is the monitored ambient concentration of pollutant i; j is the health category index so that $C_{i,j}$ and $C_{i,j-1}$, which are the upper limit concentrations for the jth and j-1th health categories, encompass $C_{i,m}$; $AQI_{i,j}$ and $AQI_{i,j-1}$ are the AQI values of pollutant i corresponding to $C_{i,j}$ and $C_{i,j-1}$, respectively.

Table 1.1 gives the health categories and the corresponding ranges of AQI values and pollutant concentrations. The overall AQI is the maximum of the sub-AQI of all pollutants, as shown in Eq. (2):

$$AQI = \max\,(AQI_1, AQI_2, \ldots, AQI_n), \text{ n=1, 2, …, 6.} \qquad (2)$$

Figure A.1: Calculating the AQI

Table A.1: China five-year plan pollution reduction targets and achievements
(10th Five-Year Plan for environmental protection: major targets vs. performance)

No.	Indicator	2000 Base	2005 Target	2005 Actual	Attained/not attained
1	SO_2 emissions (mil. tons)	19.95	18	25.49	Not attained
2	Emissions of smoke and dust (mil. tons)	11.65	11	11.83	Not attained
3	Industrial dust (mil. tons)	10.92	9	9.11	Not attained
4	COD (mil. tons)	14.45	13	14.14	Not attained
5	Industrial solid waste (mil. tons)	31.86	29	16.55	attained
6	Reuse rate of industrial water (%)		60	75	attained
7	Industrial SO_2 (mil. tons)	16.13	14.5	21.68	Not attained
8	Emissions of industrial smoke and dust (mil. tons)	9.53	8.5	9.49	Not attained
9	Industrial COD (mil. tons)	7.05	6.5	5.55	attained
10	Comprehensive use rate of industrial solid waste (%)	51.8	50	56.1	attained
11	Percent of cities meeting Grade II national standard (%)	36.5	50	54	attained
12	Urban sewage treatment rate (%)	34.3	45	52	attained
13	Green coverage of urban built-up areas (%)	28.1	35	33	Not attained
14	Percentage of land area in nature reserves (%)	9.9	13	15	attained

Source: SEPA, 2007

Table A.2: Major targets in 11th Five-Year Plan for environmental protection

No.	Indicator	2005 Actual	2010 Target	% Change to achieve target
1	COD (mil. tons)	14.14	12.7	-10%
2	SO_2 emissions (mil. tons)	25.49	22.95	-10%
3	Percentage of river sections under national monitoring program failing to meet Grade V National Surface Water Quality Standard (%)	26.1	22	-4.1 percentage points
4	Percentage of sections of 7 major rivers under national monitoring program meeting Grade III National Surface Water Quality Standard (%)	41	43	2 percentage points
5	Number of days in which urban air quality of key cities is superior to Grade II Grade II National Air Quality Standard exceeding 292 days (%)	69.4	75	5.6 percentage points

Source: SEPA, 2007

Table A.3: Major targets in 12th13th Five-Year Plan for environmental protection: targets vs. performance

No.	Item	12th Five-Year Plan Targets (Compared to 2010)	12th Five-Year Plan Achievements (Compared to 2010)	13th Five-Year Plan's Achievements (Compared to 2015)
1	Energy Intensity (Energy Consumption per Unit of GDP)	-16%	-18.20%	-15%
2	Carbon Intensity (Carbon Emission per Unit of GDP)	-17%	-20%	-18%
3	Non-Fossil Fuel Percentage	11.40%	12%	15%
4	Sulfur Dioxide (SO_2)	-8%	-18%	-15%
5	Nitrogen Oxides (NOX)	-8%	-18.60%	-15%
6	Ammonia Nitrogen	-10%	-13%	-10%
7	Chemical Oxygen Demand (COD)	-10%	-12.90%	-10%
8	Forest Coverage	21.70%	21.63%	23.04%

Source:National Bureau of Statistics of China, 2016

Appendix Two

Chapter 3 Appendices

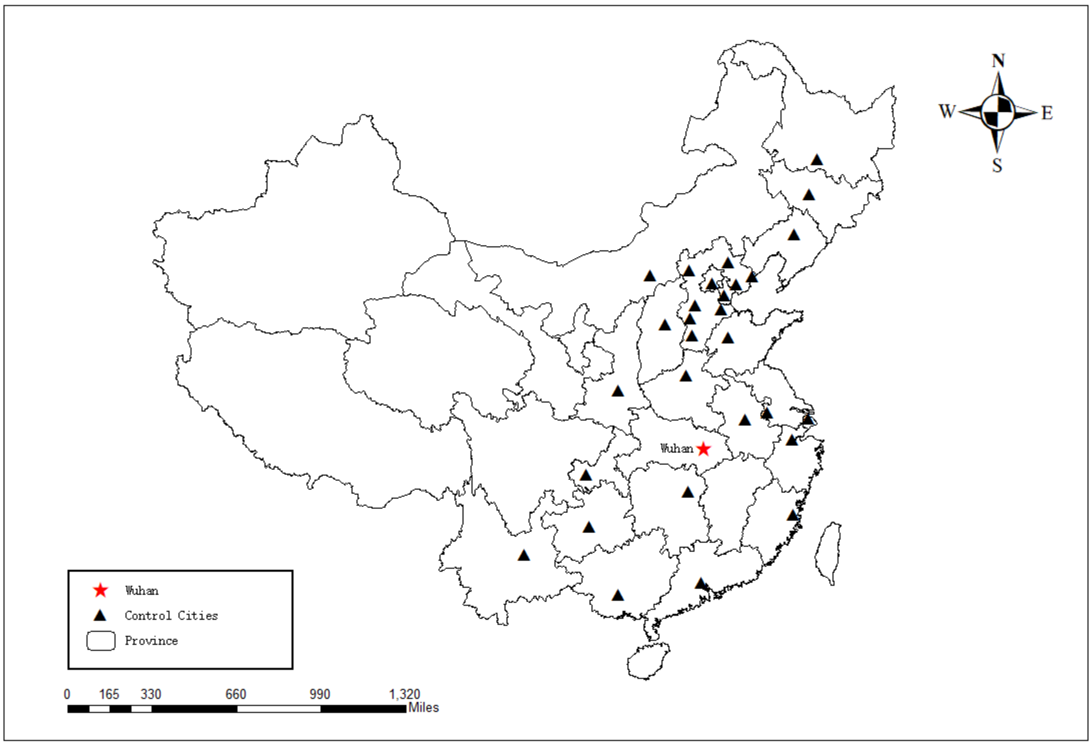


Figure B.1: The location of Wuhan and the 29 control cities

Table B.1: Meteorological monitoring station information used in the research

City	Station name	Station code	Latitude	Longitude	Elevation (m)
Wuhan	TIANHE	574940-99999	30.8	114.2	34.4
Shijiazhuang	SHIJIAZHUANG	536980-99999	38.1	114.5	105
Zhengzhou	XINZHENG	570830-99999	34.5	113.5	151
Kunming	YUANMOU	567630-99999	25.7	101.8	1120
Beijing	BEIJING - CAPITAL INTERNATIONAL AIRPORT	545110-99999	40.1	116.4	35.4
Shanghai	SHANGHAI	583620-99999	31.4	121.3	4
Guangzhou	BAIYUN INTL	592870-99999	23.4	113.2	15.2
Chongqing	JIANGBEI	575160-99999	29.7	106.4	416
Tianjin	TIANJIN	545270-99999	39.1	117.1	5
Shenyang	SHENYANG	543420-99999	41.7	123.4	43
Hefei	LUOGANG	583210-99999	31.8	117	32.9
Changsha	CHANGSHA	576870-99999	28.1	112.6	120
Jinan	JINAN	548230-99999	36.7	116.6	58
Changchun	LONGJIA	541610-99999	44	126	215
Guiyang	LONGDONGBAO	578160-99999	26.5	106.5	1139
Xian	JINGHE	571310-99999	34.4	109	411
Fuzhou	PINGTAN	589440-99999	25.5	119.3	31
Hangzhou	XIAOSHAN	584570-99999	30.2	120.3	7
Taiyuan	WUSU	537720-99999	37.7	112.4	785
Harbin	HARBIN	509530-99999	45.9	126.3	1186
Huhehaote	BAITA	534630-99999	40.9	111.5	1084
Nanning	WUXU	594310-99999	22.6	108.2	128
Nanjing	LUKOU	582380-99999	31.7	118.5	14.9
Chengde	CHENGDE	544230-99999	40.6	117.6	423
Tangshan	TANGSHAN	545340-99999	39.4	118.1	29
Cangzhou	POTOU	546180-99999	38.1	116.6	13
Xingtai	XINGTAI	537980-99999	37	114.3	184
Baoding	BAODING	546020-99999	38.5	115.3	17
Qinhuangdao	QINGLONG	544360-99999	40.4	119.4	228
Zhangjiakou	ZHANGJIAKOU	544010-99999	40.8	114.5	726

Source: NOAA (2016)

Table B.2: Control groups used in the main analysis and sensitivity analysis

Syn_full	Full sample, using the 29 cities as the control group	Shijiazhuang, Zhengzhou, Kunming, Beijing, Shanghai, Guangzhou, Chongqing, Tianjin, Shenyang, Hefei, Changsha, Jinan, Changchun, Guiyang, Xian, Fuzhou, Hangzhou, Taiyuan, Harbin, Huhehaote, Nanning, Nanjing, Chengde, Tangshan, Cangzhou, Xingtai, Baoding, Qinhuangdao, Zhangjiakou
Syn_CG1	Capital cities only	Shijiazhuang, Zhengzhou, Kunming, Beijing, Shanghai, Guangzhou, Chongqing, Tianjin, Shenyang, Hefei, Changsha, Jinan, Changchun, Guiyang, Xian, Fuzhou, Hangzhou, Taiyuan, Harbin, Huhehaote, Nanning, Nanjing
Syn_CG2	Northern Chinese cities only	Shijiazhuang, Zhengzhou, Beijing, Tianjin, Shenyang, Jinan, Changchun, Xian, Taiyuan, Harbin, Huhehaote, Chengde, Tangshan, Cangzhou, Xingtai, Baoding, Qinhuangdao, Zhangjiakou
Syn_CG3	Cities that never locked down between December 2019 and March 2020	Tianjin, Changsha, Chengde, Tangshan, Cangzhou, Xingtai, Baoding, Changchun, Qinhuangdao, Hangzhou, Zhangjiakou, Taiyuan, Huhehaote
Syn_CG4	Cities that locked down after 3 February 2020	Shijiazhuang, Zhengzhou, Kunming, Beijing, Shanghai, Guangzhou, Chongqing, Shenyang, Hefei, Jinan, Guiyang, Xian, Fuzhou, Harbin, Nanning, Nanjing

The lists of cities are collected by the authors from news, social media and official government announcements.

Table B.3: Selected socio-demographic data in Wuhan and 29 control cities in 2018

Variable	Wuhan	Average of the 29 control cities
Annual Average Population (10,000 persons)	869	844
Natural Growth Rate (%) of population	8.09	6.76
Total Land Area of Administrative region (sq. km)	8,569	18,440
Gross Regional Product (Current Prices) (100,000 yuan)	14.85m	9.01m
Per Capita Gross Regional Product (Yuan)	135,136	86,763
Secondary Industry as a Percentage of Gross Regional Product	42.96	37.07
Number of Industrial Enterprises	2,651	2,235
Number of Buses and Trolley Buses at Year-end (unit)	9,710	6,422
Total Annual Number of Passengers Transported by Buses and Trolley Buses (10,000 person-times)	145,246	83,271
Number of Taxis at Year-end (unit)	17,885	12,816
Highway Passenger Traffic (10,000 persons)	8,867	12,344
Highway Freight Traffic (10,000 tons)	38,634	30,789
Annual Electricity Consumption (10,000 kwh)	0.58m	0.49m
Electricity Consumption for Industrial use	0.280m	0.257m
Household Electricity Consumption for Urban and Rural Residential	996,309	600,036

Source: China City Statistical Yearbook (2019).

Appendix Three

Chapter 4 Appendices

Appendix C.1 The detailed operations and measures of the CEIP inspection [1]

1. Enterprise production.

1.1. Industry type and main products, products and capacity of the previous month, whether production lines are under operation;

1.2. Whether there exist illegal construction and whether the investigation results are consistent with the Environmental Impact Assessment;

1.3. Whether there are newly produced waste gas, waste water, solid waste and other pollutants;

2. Implementation of environmental protection in enterprises.

2.1. Examine whether the project performs the EIA (Environmental impact assessment) procedures according to the law, check the EIA documents and the review of the EIA, whether the EIA documents involves counterfeiting;

2.2. Examine whether the nature of the project, the scale of production, the location, or the measures used for pollution control are consistent with the EIA and the approval doc-

[1] source: http://www.tiandeyi.com/

uments; Whether the enterprise' pollution control facilities are put into production without the acceptance of the environmental protection department;

2.3. Examine whether the enterprise has been approved by the Environmental Protection Completion acceptance, whether relevant procedures are complete, after the project being put into operation; Examine whether there exist pollution control facilities being put into production without acceptance by local environmental protection departments;

2.4. Examine production workshop: whether proper anti-corrosion treatment has been conducted and regularly maintained in Raw materials related to acid, alkali, and other corrosive workshop floor; whether there exist leakage fields during the production process have;

2.5. Examine and inspect the application of a pollution discharge license, the execution of a discharge declaration and the payment of a discharge fee;

3. Inspection of waste gas pollution

Check and examine the running state, historical operation state, processing capacity and quantity of the waste gas treatment facilities; Make sure emission standards are met and pollution control measures are strengthened.

4. Air pollution prevention and control facilities

4.1. Examine for dust-removing, desulphurization, denitrification, and other gaseous pollutants purification systems;

4.2. Examine whether the waste gas emission meet the standard;

4.3. Examine whether polluter has been building new exhaust cylinder in the prohibited areas;

4.4. Examine whether the height of the exhaust cylinder meets the requirements of the national or local pollutant emission standards;

4.5. Examine whether the sampling hole and sampling monitoring platform are set on the exhaust pipes;

4.6. Examine whether the exhaust outlet settings meet the requirements, including height, sampling port and sign; whether online monitoring facilities are installed and implemented according to requirements;

5. Unorganized emission source

5.1. For unorganized emission of toxic and harmful gases, dusts and soot, Examine whether conditional emission is organized, check whether emitters have been regulated;

5.2. Examine whether the coal yard, stock yard, goods yard take measures to prevent or control the waste air pollution according to the requirements or set up the anti-dust equipment;

5.3 Monitor at the boundary of the enterprise to check whether the unorganized emission meets the requirements of the relevant environmental protection standards.

6. Inspection on waste gas emitted from Boiler, Petrochemical, Chemical Combustion (key industries of pollution control)

6.1. Examine whether measures taken by chemical and petrochemical enterprises to deal with continuously combustible organic waste gas meets requirements and whether the way of recycling or incineration is reasonable;

6.2. Examine the verification procedures and performance of the equipment for boiler combustion, examine the operation conditions of the combustion equipment. Examine the conditions of controlling for sulfur dioxide and nitrogen oxides;

6.3. Examine whether waste gas emissions, dust, and odor emissions meet the requirements of the related pollution emission standards;

6.4. Examine the environmental protection measures for the transport, handling and storage of volatile toxic, harmful gas and dust;

7. Warnings for other enterprises

7.1. The enterprises are not allowed to dismantle, idle or shut down the pollution treatment facilities and places without the approval of the environmental protection department;

7.2. The enterprises are prohibited to violate the setting regulation of emission outlet, to secretly install emission pipe and modify monitoring data with falsification;

7.3. The enterprises were not allowed to construct in prohibited areas, meanwhile the project construction should comply with planning.

Figure C.1: The location of Hebei cities and control cities

Table C.1: The impact of CGEP inspection on air quality in Hebei (logs)

	(1)	(2)	(3)	(4)
(a) Variables	l_$PM_{2.5}$	l_PM10	l_NO_2	l_SO_2
Inspection	-0.192***	-0.138***	0.021	0.025
	(0.023)	(0.016)	(0.023)	(0.032)
R-squared	0.672	0.666	0.761	0.737
Weather Control	N	N	N	N
(b) Variables	l_$PM_{2.5}$wn	l_PM10wn	l_NO_2wn	l_SO_2wn
Inspection	-0.051***	-0.055***	0.026	0.041**
	(0.015)	(0.009)	(0.016)	(0.019)
R-squared	0.879	0.906	0.930	0.924
Weather Control	N	N	N	N
(c) Variables	l_$PM_{2.5}$	l_PM10	l_NO_2	l_SO_2
Inspection	-0.152***	-0.168***	0.051**	-0.069**
	(0.022)	(0.018)	(0.021)	(0.033)
R-squared	0.720	0.717	0.817	0.827
Weather Control	Y	Y	Y	Y
(d) Variables	l_$PM_{2.5}$wn	l_PM10wn	l_NO_2wn	l_SO_2wn
Inspection	-0.048**	-0.071***	0.035**	0.021
	(0.014)	(0.009)	(0.017)	(0.021)
R-squared	0.899	0.925	0.937	0.952
Weather Control	Y	Y	Y	Y
City FE	Y	Y	Y	Y
Week FE	Y	Y	Y	Y
Observations	2,072	2,072	2,072	2,072
City Numbers	28	28	28	28

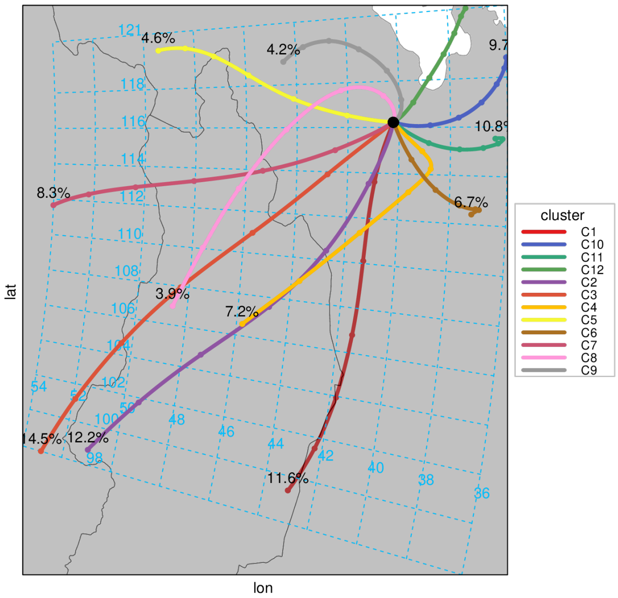
Note : *** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses, standard errors are clustered at the city level.

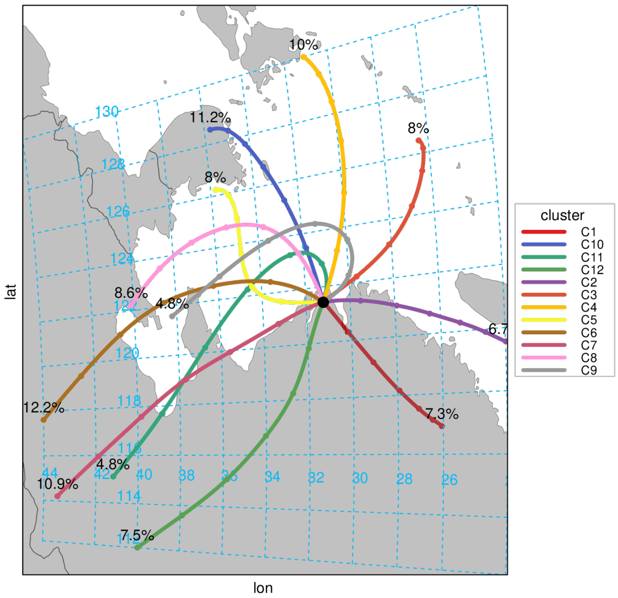
Table C.2: The estimated impact of CEIP on air quality in Hebei using alternative weighting methods

	(1)	(2)	(3)	(4)
Average treatment effect on treated (ATT)	Hebei	Hebei	Hebei	Hebei
Weighting method	Arithmetic mean (main)	Population weighted	GDP weighted	Industrial output weighted
$PM_{2.5}$wn	-4.69	-6.67	-4.26	-4.29
	(-9.61, -0.22)	(-13.19, -0.17)	(-10.00, -2.12)	(-10.11, 2.47)
PM10wn	-20.80	-23.94	-20.61	-20.62
	(-30.91,-8.11)	(-36.71, -7.47)	(-33.90, -6.02)	(-34.56, -5.28)
NO_2wn	-1.11	-1.56	0.18	1.30
	(-4.42, 1.33)	(-5.42, 1.94)	(-3.28, 3.23)	(-2.09, 4.66)
SO_2wn	-1.26	-1.79	-0.87	-0.69
	(-4.15, 0.42)	(-4.76, 0.74)	(-2.51, 1.14)	(-2.45, 1.24)

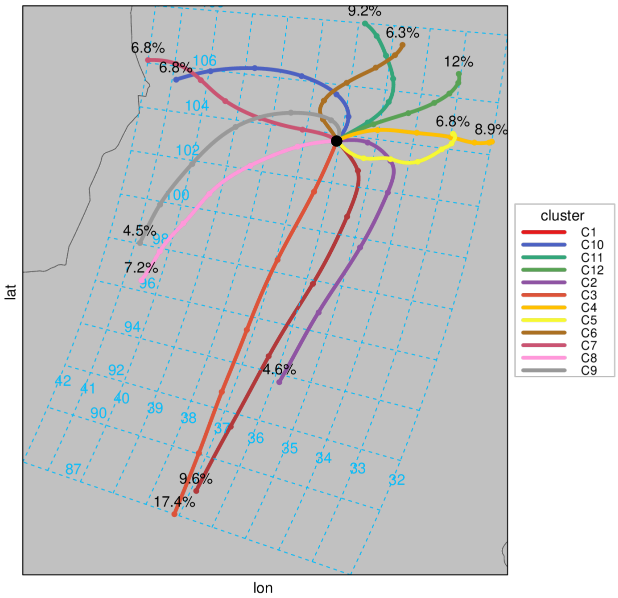
Note : Lower and upper bound of the average-ATT estimation in parentheses. Column 1 calculates the arithmetic mean of four weather normalised pollutants in eight Hebei cities, while columns 2, 3 and 4 refer to the population, GDP and industrial output weighted average of four weather normalised pollutants in eight Hebei cities. Data regarding population (Household Registered Population at year-end (10,000 persons) in 2015), GDP (Gross Regional Product, Current Prices (10,000 yuan) in 2015) and industrial output (Gross Industrial Output Value, current price, 10,000 yuan in 2015) come from China City Statistical Year Book 2016.



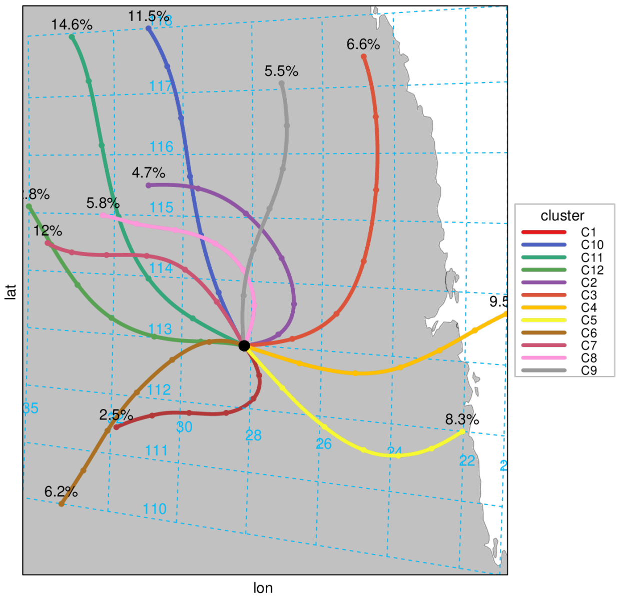
(a) Beijing (lat:40.22, lon:116.23)



(b) Shanghai (lat:31.21, lon:121.45)

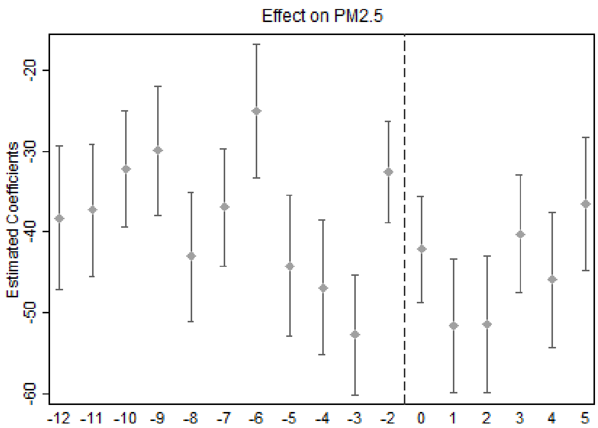


(c) Lanzhou (lat:36.10, lon:103.72)

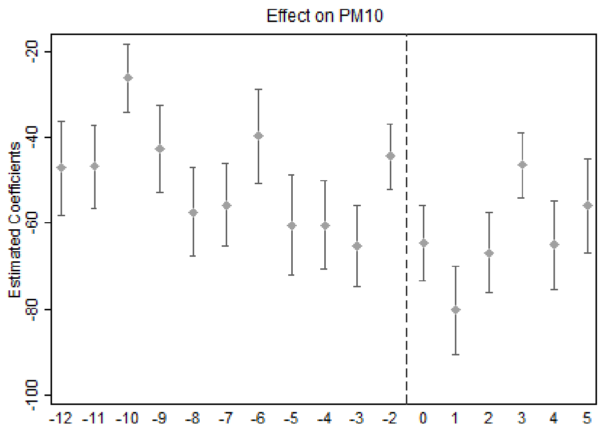


(d) Changsha (lat:28.20, lon:112.98)

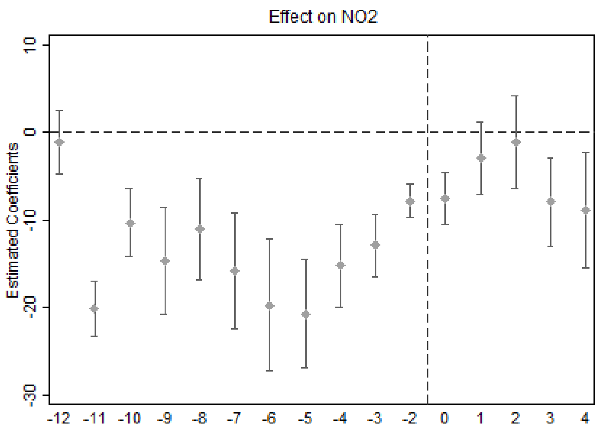
Figure C.2: The plots of 72-hour back-trajectory clusters in 4 selected cities



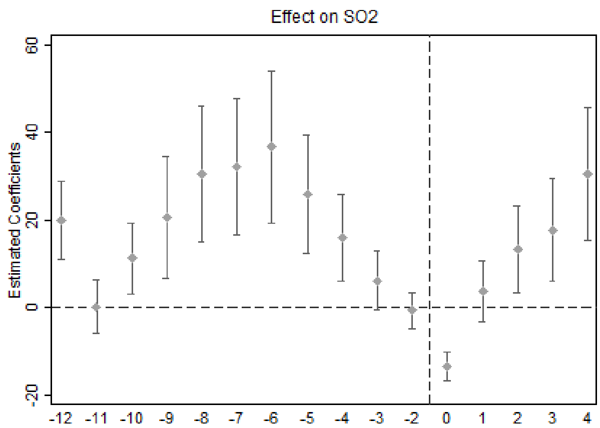
(a) $PM_{2.5}$ (every 1-month window)



(b) PM10 (every 1-month window)



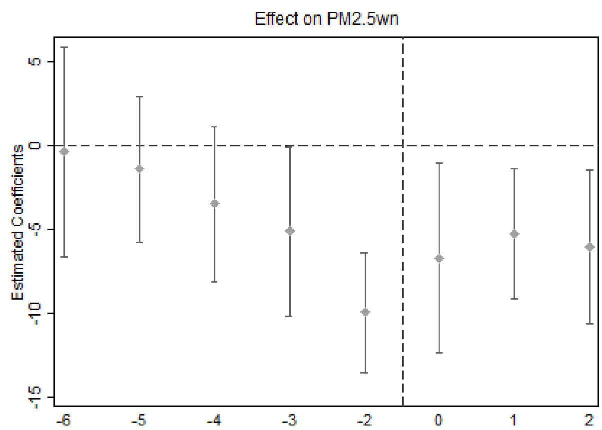
(c) NO_2 (every 1-month window)



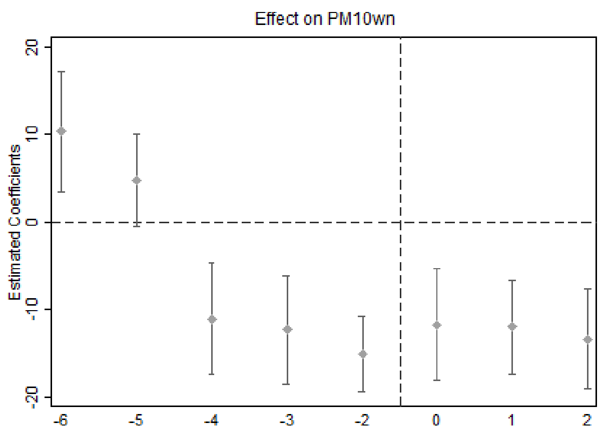
(d) SO_2 (every 1-month window)

Note: The error bars refer to the 90% confidence intervals of the estimated coefficients.

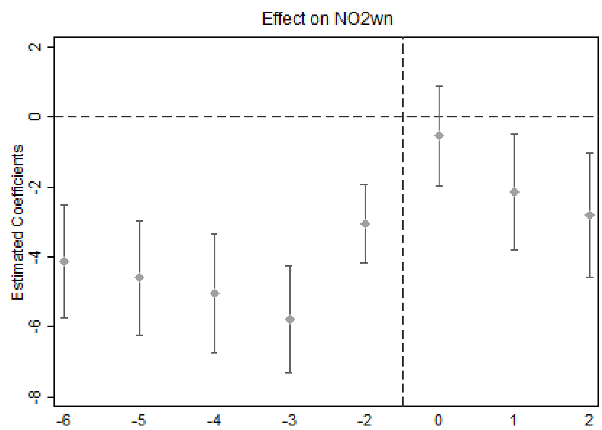
Figure C.3: The plot of pre-inspection parallel trend tests on observed pollutants



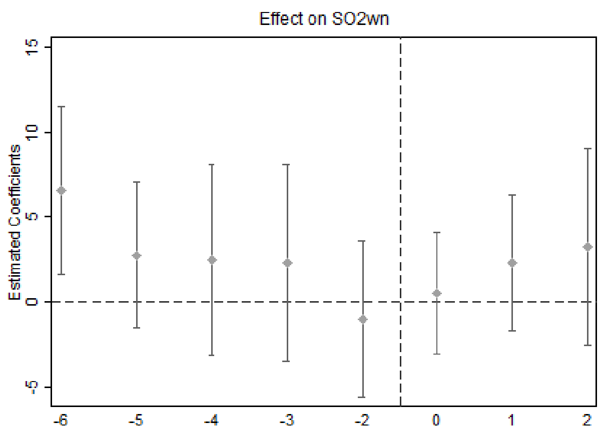
(a) $PM_{2.5}$ (every 2-month window)



(b) PM10 (every 2-month window)



(c) NO_2 (every 2-month window)



(d) SO_2 (every 2-month window)

Note: The error bars refer to the 90% confidence intervals of the estimated coefficients.

Figure C.4: The plot of pre-inspection parallel trend tests on weather normalised pollutants

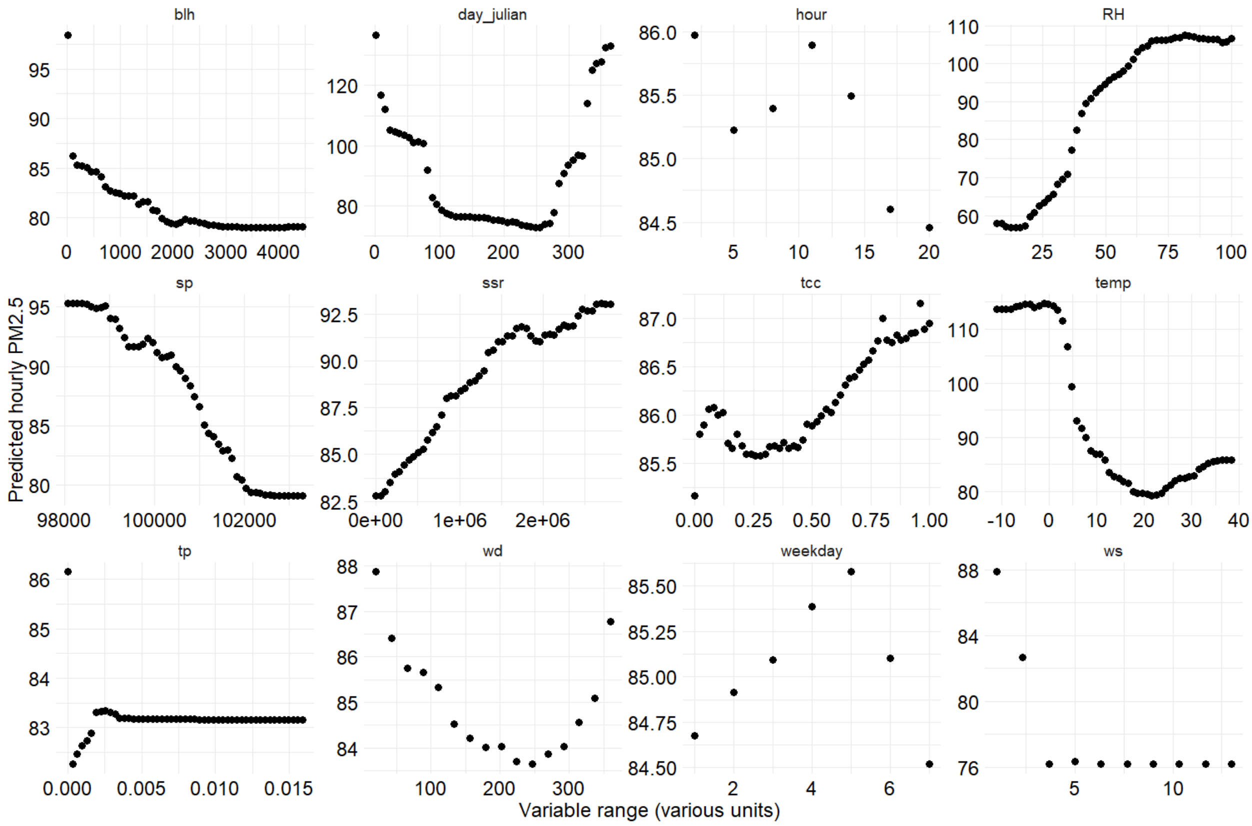


Figure C.5: The partial dependence of meteorological variable in $PM_{2.5}$ random forest model in Shijiazhuang

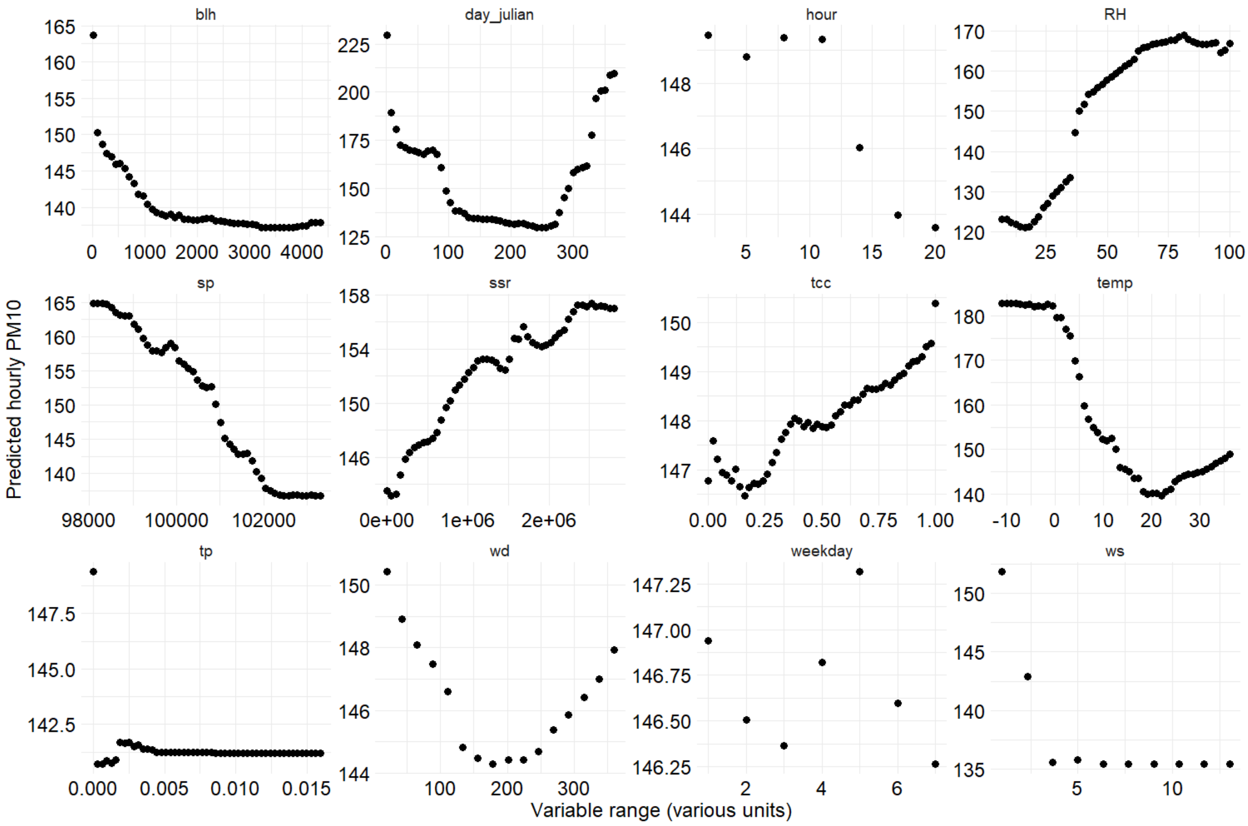


Figure C.6: The partial dependence of meteorological variable in PM10 random forest model in Shijiazhuang

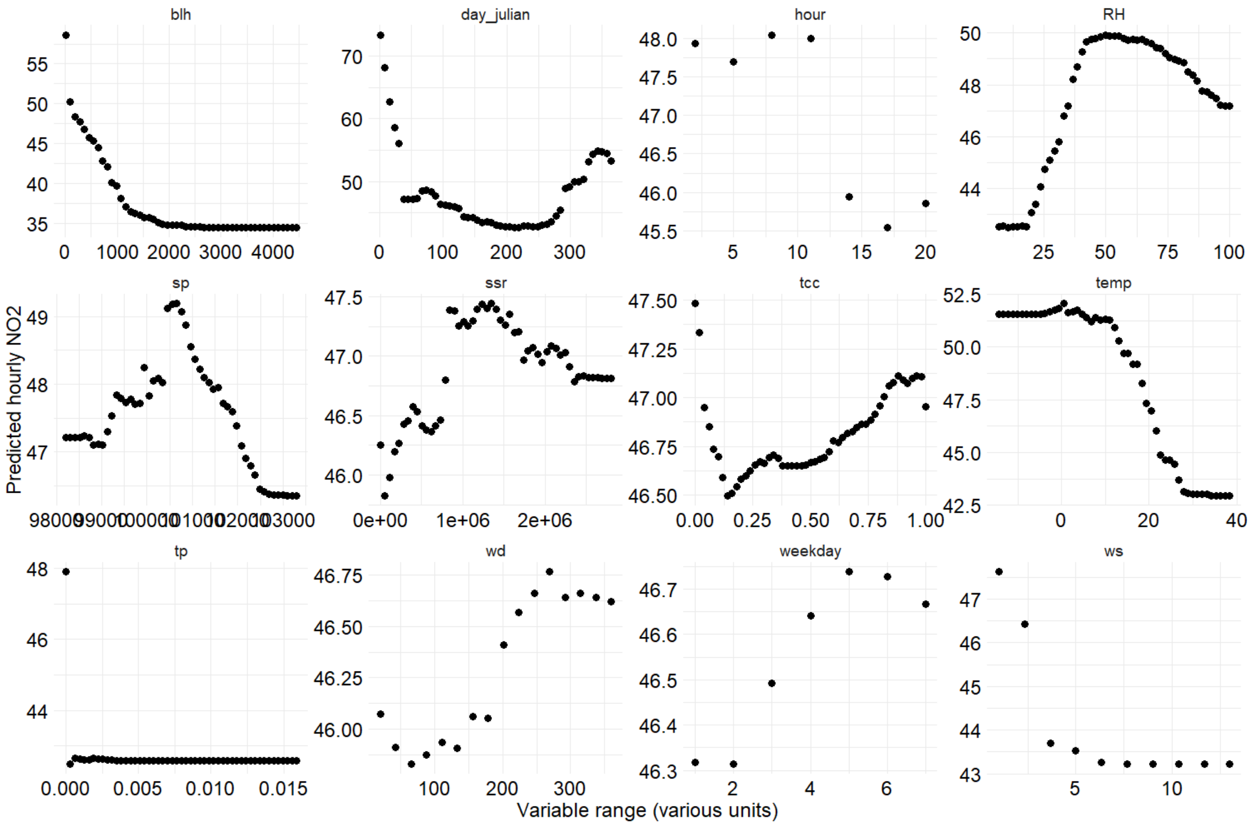


Figure C.7: The partial dependence of meteorological variable in NO_2 random forest model in Shijiazhuang

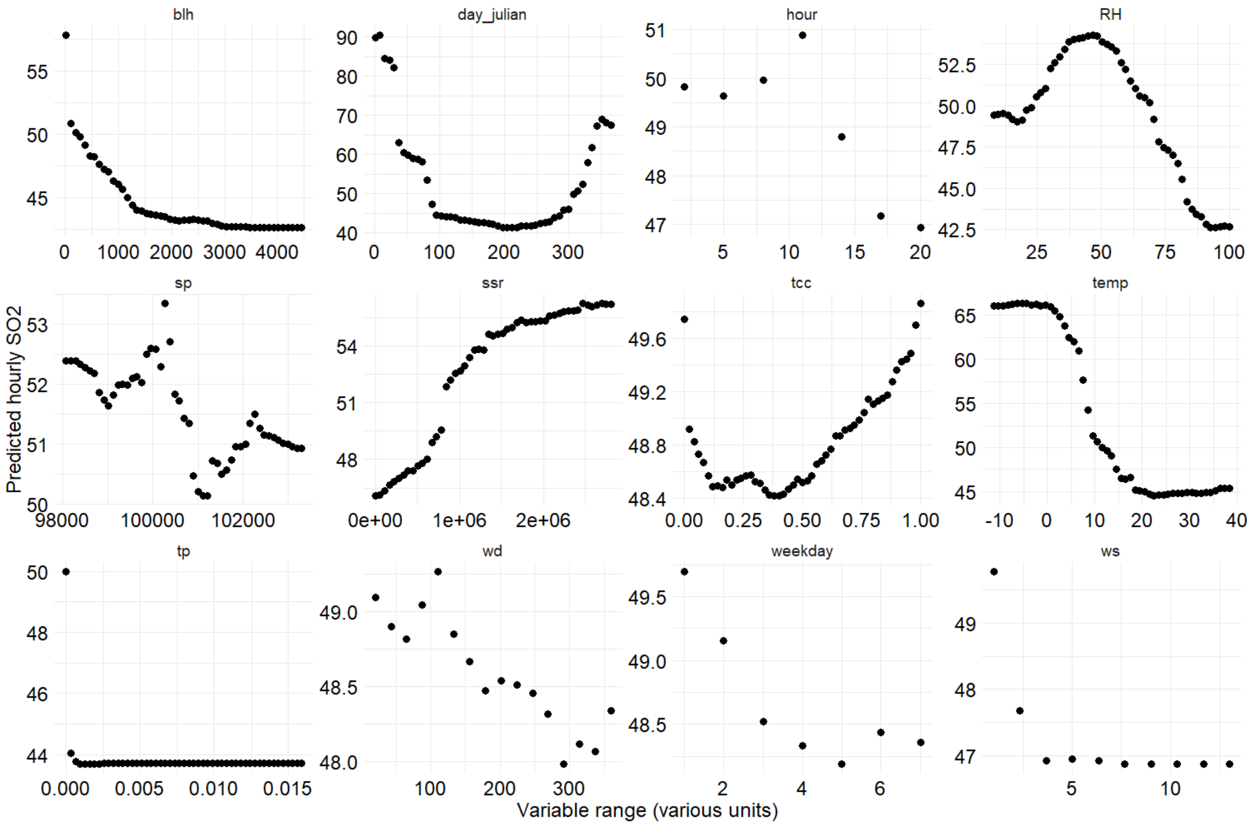
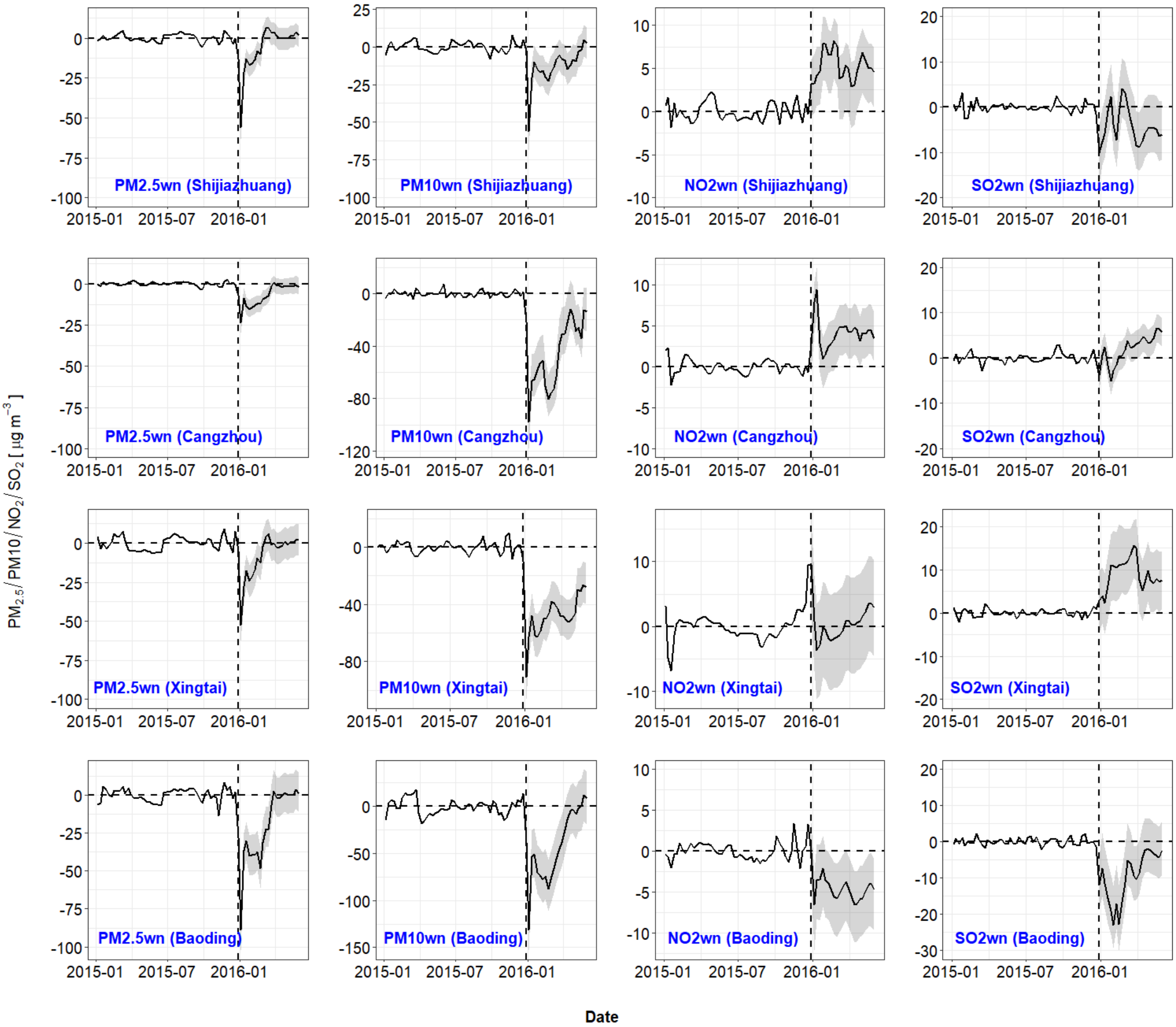
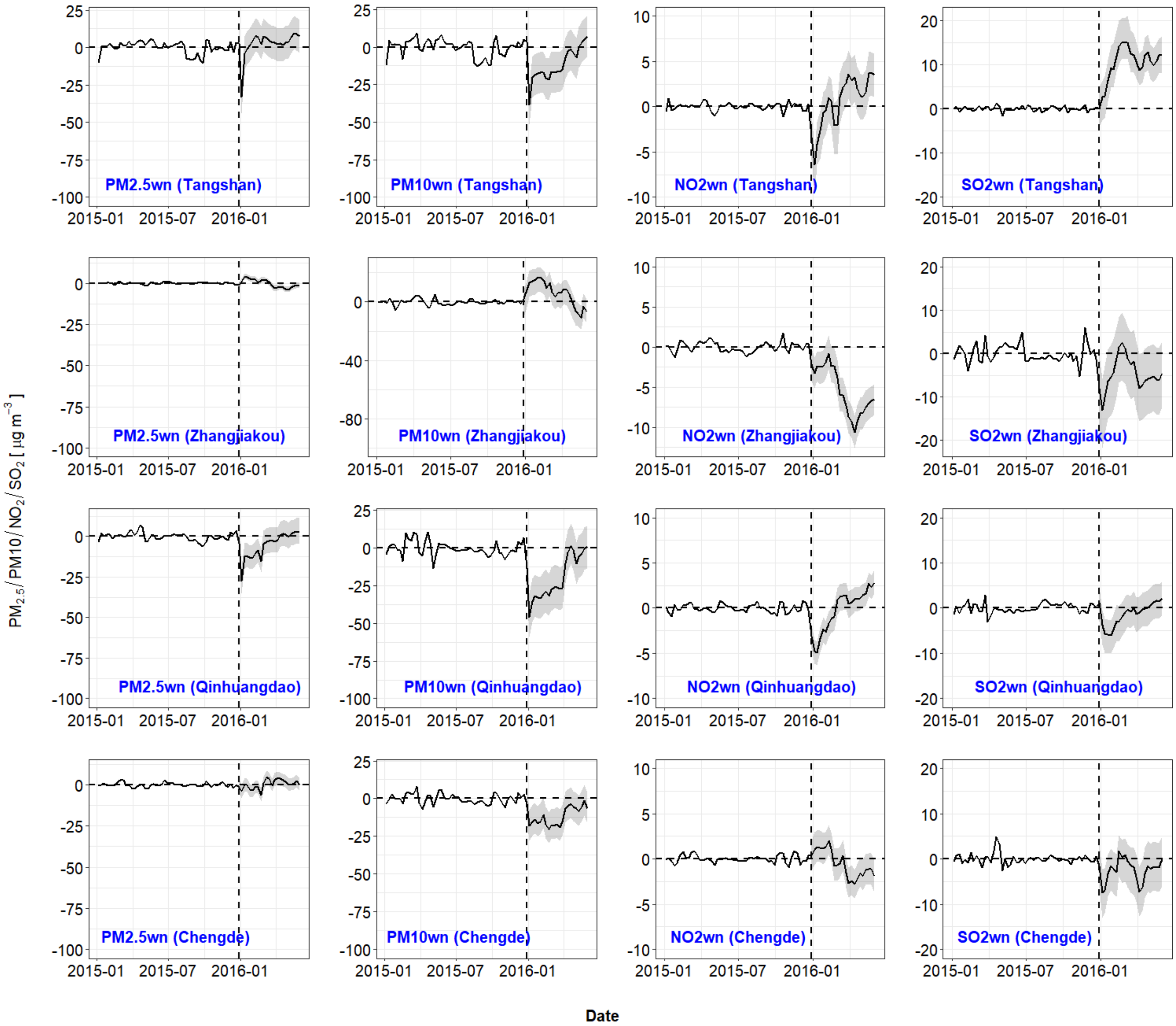


Figure C.8: The partial dependence of meteorological variable in SO_2 random forest model in Shijiazhuang



Note: all figures show the point-wise 95% confidence interval using Jackknife+ procedure.

Figure C.9: The heterogeneity effects of inspection on weather normalised pollutants in Shijiazhuang, Cangzhou, Xingtai and Baoding



Note: all figures show the point-wise 95% confidence interval using Jackknife+ procedure.

Figure C.10: The heterogeneity effects of inspection on weather normalised pollutants in Tangshan, Zhangjiakou, Qinhuangdao and Chengde

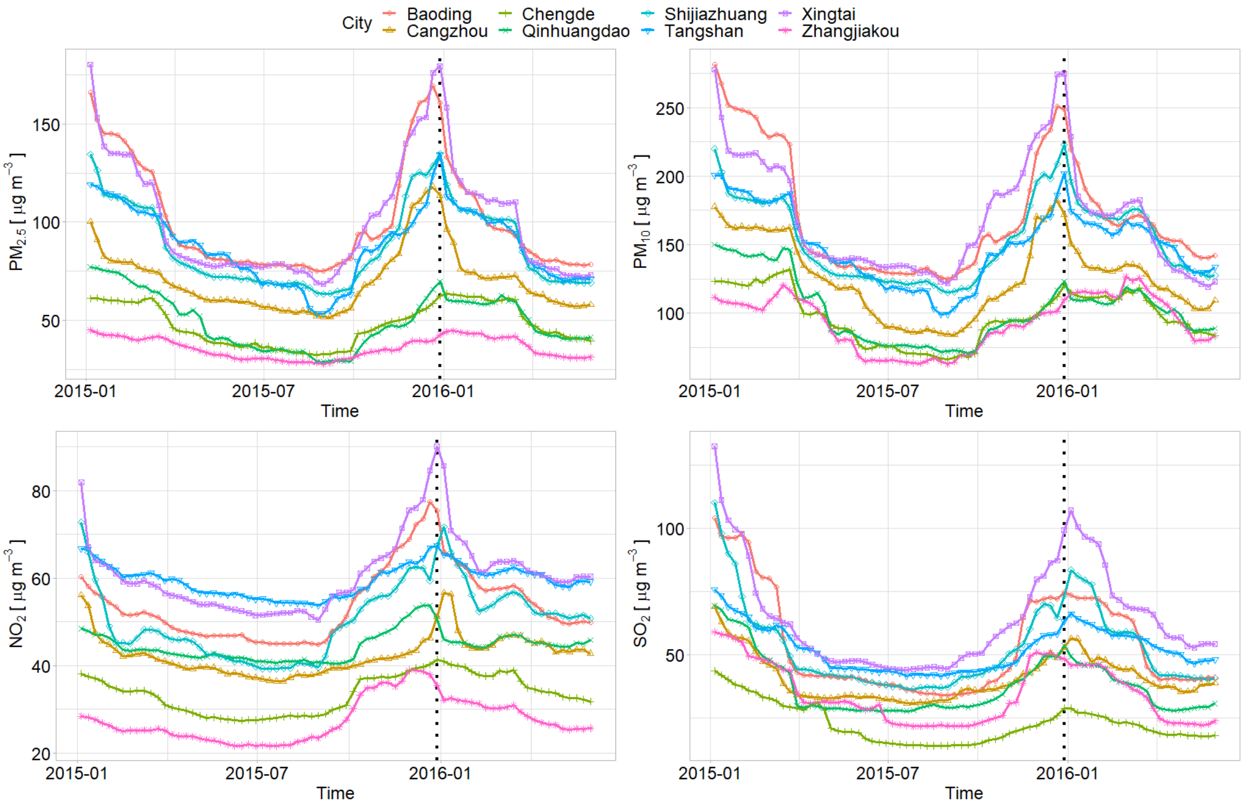


Figure C.11: The plot of weekly weather normalised pollutants in 8 Hebei cities

Table C.3: Control groups used in the main analysis and sensitivity analysis

SFull	Full sample, use the 27 cities as the control group	Changchun, Changsha, Chengdu, Chongqing, Fuzhou, Guangzhou, Guiyang, Hangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Lhasa, Nanchang, Nanjing, Nanning, Shanghai, Shenyang, Taiyuan, Urumqi, Wuhan, Xian, Xining, Yinchuan, Zhengzhou
SCG1	Excluding cities from "Northeastern three provinces (Heilongjiang, Jilin and Liaoning)"	Exclude Changchun, Shenyang and Harbin from the full control sample (SFull)
SCG2	Excluding Southern Chinese cities (cities locate to the south of the "Qin-mountain and Huai-river line")	Exclude Chengdu, Chongqing, Fuzhou, Guangzhou, Guiyang, Hangzhou, Nanchang, Nanjing, Nanning and Shanghai from the full control sample (SFull) (we keep Changsha, Wuhan, Hefei and Kunming in the "SCG2" group since these four cities have partially centralised winter heating programs.)
SCG3	Excluding cities with cleaner air quality (excluding cities in Yunan, Guizhou, Guangxi and Fujian provicnes)	Exclude kunming, Guiyang, Nanning and Fuzhou from the full control sample (SFull)

Appendix Four

Chapter 5 Appendices

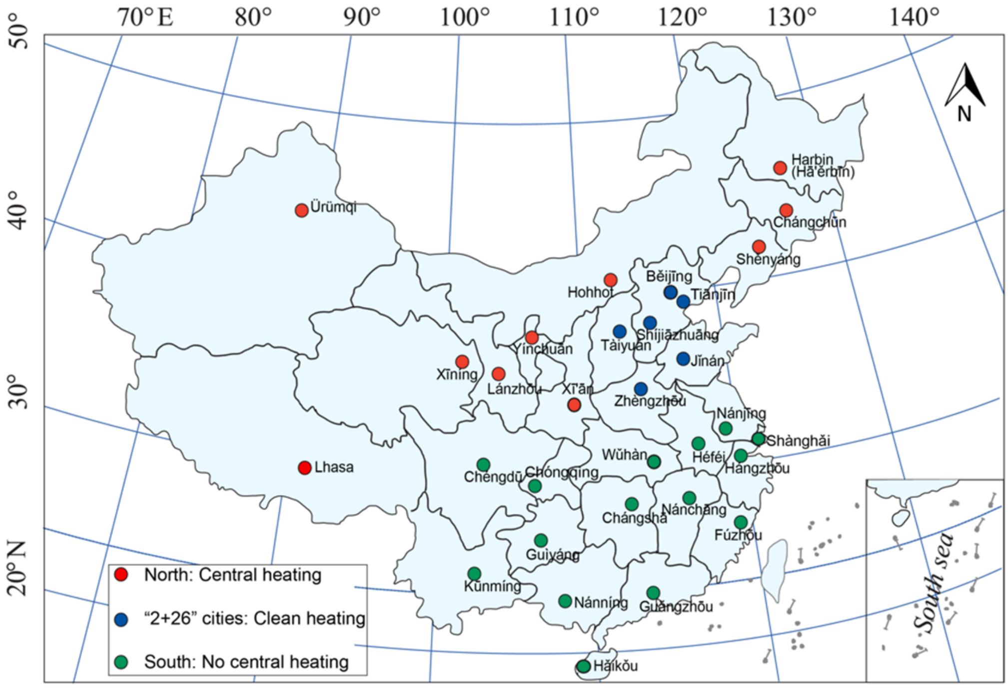


Figure D.1: The location of "2+26", "North" and "South" cities in China.

Note: the northern and southern cities are divided by the "Huai-River" line, Lhasa belongs to northern city but did not have winter heating until winter season in 2020. The blue and red dots refer to "2+26" and "North" cities (2 treatment groups), respectively, and the green dots refer to southern cities that are not exposed to winter heating policy (control group).

Table D.1: Winter heating start and end time in 29 cities

City	2015	2016	2017	2018	2019	2020
Beijing	/11.15	3.15/11.13	3.15/11.15	3.20/11.15	3.15/11.15	3.15/11.15
Changchun	4.10/10.25	4.10/10.25	4.10/10.25	4.10/10.25	4.10/10.25	4.10/10.25
Changsha	NA	NA	NA	NA	NA	NA
Chengdu	NA	NA	NA	NA	NA	NA
Chongqing	NA	NA	NA	NA	NA	NA
Fuzhou	NA	NA	NA	NA	NA	NA
Guangzhou	NA	NA	NA	NA	NA	NA
Guiyang	NA	NA	NA	NA	NA	NA
Haikou	NA	NA	NA	NA	NA	NA
Hangzhou	NA	NA	NA	NA	NA	NA
Harbin	4.10/10.20	4.15/10.20	4.20/10.20	4.20/10.20	4.20/10.20	4.20/10.20
Hefei	NA	NA	NA	NA	NA	NA
Hohhot	/10.15	/10.15	/10.15	/10.15	/10.15	/10.15
Jinan	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15
Kunming	NA	NA	NA	NA	NA	NA
Lanzhou	3.31/11.01	3.31/11.01	3.31/11.01	3.31/11.01	3.31/11.01	3.31/11.01
Nanchang	NA	NA	NA	NA	NA	NA
Nanjing	NA	NA	NA	NA	NA	NA
Nanning	NA	NA	NA	NA	NA	NA
Shanghai	NA	NA	NA	NA	NA	NA
Shenyang	3.31/11.01	3.15/11.15	3.15/11.10	3.31/11.10	3.31/11.10	3.31/11.10
Shijiazhuang	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15
Taiyuan	4.01/11.01	4.01/11.01	4.01/11.01	4.01/11.01	4.01/11.01	4.01/11.01
Tianjin	3.31/11.15	3.31/11.05	3.31/11.01	3.31/11.01	3.31/11.15	3.31/11.15
Urumqi	4.10/10.10	4.10/10.10	4.10/10.10	4.10/10.10	4.10/10.10	4.10/10.10
Wuhan	NA	NA	NA	NA	NA	NA
Xian	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15	3.15/10.15
Xining	/10.15	/10.15	/10.15	/10.15	/10.15	/10.15
Yinchuan	/11.01	/11.01	/11.01	/11.01	/11.01	/11.01
Zhengzhou	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15	3.15/11.15

Source: Collected from government and news websites. NA refers to no winter heating.

References

Abadie, Alberto (2019). “Using synthetic controls: Feasibility, data requirements, and methodological aspects”. **in***Journal of Economic Literature*.

Abadie, Alberto and Matias D Cattaneo (2018). “Econometric methods for program evaluation”. **in***Annual Review of Economics*: 10, **pages** 465–503.

Abadie, Alberto, Alexis Diamond and Jens Hainmueller (2010). “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program”. **in***Journal of the American statistical Association*: 105.490, **pages** 493–505.

— (2015). “Comparative politics and the synthetic control method”. **in***American Journal of Political Science*: 59.2, **pages** 495–510.

Abadie, Alberto and Javier Gardeazabal (2003). “The economic costs of conflict: A case study of the Basque Country”. **in***American economic review*: 93.1, **pages** 113–132.

Almond, Douglas, Yuyu Chen, Michael Greenstone et al. (2009). “Winter heating or clean air? Unintended impacts of China’s Huai river policy”. **in***American Economic Review*: 99.2, **pages** 184–90.

Altschul, Stephen F, Thomas L Madden, Alejandro A Schäffer et al. (1997). “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs”. **in***Nucleic acids research*: 25.17, **pages** 3389–3402.

Andrews, Steven Q (2008). “Inconsistencies in air quality metrics:‘Blue Sky’days and PM10 concentrations in Beijing”. **in***Environmental Research Letters*: 3.3, **page** 034009.

Anh, Vo, Hiep Duc and Merched Azzi (1997). "Modeling anthropogenic trends in air quality data". **in***Journal of the Air & Waste Management Association*: 47.1, **pages** 66–71.

Athey, Susan and Guido W Imbens (2017). "The state of applied econometrics: Causality and policy evaluation". **in***Journal of Economic Perspectives*: 31.2, **pages** 3–32.

— (2019). "Machine learning methods that economists should know about". **in***Annual Review of Economics*: 11, **pages** 685–725.

Auffhammer, Maximilian, Antonio M Bento and Scott E Lowe (2009). "Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of a spatially disaggregated analysis". **in***Journal of Environmental Economics and Management*: 58.1, **pages** 15–26.

— (2011). "The city-level effects of the 1990 Clean Air Act amendments". **in***Land Economics*: 87.1, **pages** 1–18.

Baker, Stephen, Jonathan Hardy, Kenneth E Sanderson et al. (2007). "A novel linear plasmid mediates flagellar variation in Salmonella Typhi". **in***PLoS Pathog*: 3.5, e59.

Barwick, Panle Jia, Shanjun Li, Liguo Lin et al. (2019). *From fog to smog: The value of pollution information.* techreport. National Bureau of Economic Research.

Ben-Michael, Eli, Avi Feller and Jesse Rothstein (2020). *The Augmented Synthetic Control Method.* techreport. arXiv. org.

— (2021). "The augmented synthetic control method". **in***Journal of the American Statistical Association*: just-accepted, **pages** 1–34.

Bento, Antonio, Matthew Freedman and Corey Lang (2015). "Who benefits from environmental regulation? Evidence from the Clean Air Act Amendments". **in***Review of Economics and Statistics*: 97.3, **pages** 610–622.

Breiman, Leo (2001). "Random forests". **in***Machine learning*: 45.1, **pages** 5–32.

Breiman, Leo, Jerome Friedman, Charles J Stone et al. (1984). *Classification and regression trees.* CRC press.

Broner, Fernando, Paula Bustos and Vasco M Carvalho (2012). *Sources of comparative advantage in polluting industries.* techreport. National Bureau of Economic Research.

Buchard, V, AM Da Silva, CA Randles et al. (2016). "Evaluation of the surface PM2. 5 in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States". **in***Atmospheric Environment*: 125, **pages** 100–111.

Cai, Hongbin, Yuyu Chen and Qing Gong (2016). "Polluting thy neighbor: Unintended consequences of China s pollution reduction mandates". **in***Journal of Environmental Economics and Management*: 76, **pages** 86–104.

Cao, Jing, Richard Garbaccio and Mun S Ho (2009). *China's 11th five-year plan and the environment: reducing SO2 emissions.*

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy et al. (2013). "Catastrophic natural disasters and economic growth". **in***Review of Economics and Statistics*: 95.5, **pages** 1549–1561.

Chay, Kenneth, Carlos Dobkin and Michael Greenstone (2003). "The Clean Air Act of 1970 and adult mortality". **in***Journal of risk and uncertainty*: 27.3, **pages** 279–300.

Chay, Kenneth Y and Michael Greenstone (2003a). *Air quality, infant mortality, and the Clean Air Act of 1970.* techreport. National Bureau of Economic Research.

— (2003b). "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession". **in***The quarterly journal of economics*: 118.3, **pages** 1121–1167.

— (2005). "Does air quality matter? Evidence from the housing market". **in***Journal of political Economy*: 113.2, **pages** 376–424.

Chen, Ye, Hongbin Li and Li-An Zhou (2005). "Relative performance evaluation and the turnover of provincial leaders in China". **in***Economics Letters*: 88.3, **pages** 421–425.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone et al. (2013). "Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy". **in***Proceedings of the National Academy of Sciences*: 110.32, **pages** 12936–12941.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar et al. (2012). "Gaming in air pollution data? Lessons from China". **in** *The BE Journal of Economic Analysis & Policy*: 13.3.

— (2013). "The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality". **in** *Journal of Environmental Economics and Management*: 66.3, **pages** 424–443.

Chen, Yvonne Jie, Pei Li and Yi Lu (2018). "Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China". **in** *Journal of Development Economics*: 133, **pages** 84–101.

Cheng, Chu-Chuan and Yu-Bong Lai (2012). "Does a stricter enforcement policy protect the environment? A political economy perspective". **in** *Resource and Energy Economics*: 34.4, **pages** 431–441.

Cheung, Chun Wai, Guojun He and Yuhang Pan (2020). "Mitigating the air pollution effect? The remarkable decline in the pollution-mortality relationship in Hong Kong". **in** *Journal of environmental economics and management*: 101, **page** 102316.

Cole, Matthew A, Robert JR Elliott and Bowen Liu (2020). "The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach". **in** *Environmental and Resource Economics*: 76.4, **pages** 553–580.

Currie, Janet, Eric A Hanushek, E Megan Kahn et al. (2009). "Does pollution increase school absences?" **in** *The Review of Economics and Statistics*: 91.4, **pages** 682–694.

Currie, Janet and Matthew Neidell (2005). "Air pollution and infant health: what can we learn from California's recent experience?" **in** *The Quarterly Journal of Economics*: 120.3, **pages** 1003–1030.

Currie, Janet, Matthew Neidell and Johannes F Schmieder (2009). "Air pollution and infant health: Lessons from New Jersey". **in** *Journal of health economics*: 28.3, **pages** 688–703.

Currie, Janet and Reed Walker (2011). "Traffic congestion and infant health: Evidence from E-ZPass". **in** *American Economic Journal: Applied Economics*: 3.1, **pages** 65–90.

Dai, Qili, Linlu Hou, Bowen Liu et al. (2021). "Spring Festival and COVID-19 lockdown: disentangling PM sources in major Chinese cities". **in***Geophysical Research Letters*: e2021GL093403.

Davis, Lucas W (2008). "The effect of driving restrictions on air quality in Mexico City". **in***Journal of Political Economy*: 116.1, **pages** 38–81.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller et al. (2019). "The mortality and medical costs of air pollution: Evidence from changes in wind direction". **in***American Economic Review*: 109.12, **pages** 4178–4219.

Donghua, Chen, Zhang Tiesheng and Li Xiang (2008). "Law environment, government regulation and implicit contract: Empirical evidence from the scandals of China's listed companies". **in***Frontiers of Economics in China*: 3.4, **pages** 560–584.

Du, Bin, Zhigang Li and Jia Yuan (2014). "Visibility has more to say about the pollution–income link". **in***Ecological Economics*: 101, **pages** 81–89.

Duflo, Esther, Michael Greenstone, Rohini Pande et al. (2013). "Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India". **in***The Quarterly Journal of Economics*: 128.4, **pages** 1499–1545.

— (2018). "The value of regulatory discretion: Estimates from environmental inspections in India". **in***Econometrica*: 86.6, **pages** 2123–2160.

Dustmann, Christian, Uta Schönberg and Jan Stuhler (2017). "Labor supply shocks, native wages, and the adjustment of local employment". **in***The Quarterly Journal of Economics*: 132.1, **pages** 435–483.

Earnhart, Dietrich H and Robert L Glicksman (2015). "Coercive vs. cooperative enforcement: Effect of enforcement approach on environmental management". **in***International Review of Law and Economics*: 42, **pages** 135–146.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, Peng Yin et al. (2015). "Growth, pollution, and life expectancy: China from 1991-2012". **in***American Economic Review*: 105.5, **pages** 226–31.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He and Maigeng Zhou (2017). “New evidence on the impact of sustained exposure to air pollution on life expectancy from China’s Huai River Policy”. **in** *Proceedings of the National Academy of Sciences*: 114.39, **pages** 10384–10389.

Eckert, Heather (2004). “Inspections, warnings, and compliance: the case of petroleum storage regulation”. **in** *Journal of Environmental Economics and Management*: 47.2, **pages** 232–259.

Fan, Maoyong, Guojun He and Maigeng Zhou (2020). “The winter choke: coal-fired heating, air pollution, and mortality in China”. **in** *Journal of health economics*: 71, **page** 102316.

Fang, Hanming, Long Wang and Yang Yang (2020). “Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in china”. **in** *Journal of Public Economics*: 191, **page** 104272.

Faustini, Annunziata, Regula Rapp and Francesco Forastiere (2014). “Nitrogen dioxide and mortality: review and meta-analysis of long-term studies”. **in** *European Respiratory Journal*: 44.3, **pages** 744–753.

Foster, Andrew, Emilio Gutierrez and Naresh Kumar (2009). “Voluntary compliance, pollution levels, and infant mortality in Mexico”. **in** *American Economic Review*: 99.2, **pages** 191–97.

Fu, Shihe and Yizhen Gu (2017). “Highway toll and air pollution: Evidence from Chinese cities”. **in** *Journal of environmental Economics and Management*: 83, **pages** 32–49.

Gao, Yun (2016). “China’s response to climate change issues after Paris Climate Change Conference”. **in** *Advances in Climate Change Research*: 7.4, **pages** 235–240.

Grange, Stuart K and David C Carslaw (2019). “Using meteorological normalisation to detect interventions in air quality time series”. **in** *Science of The Total Environment*: 653, **pages** 578–588.

Grange, Stuart K, David C Carslaw, Alastair C Lewis et al. (2018). “Random forest meteorological normalisation models for Swiss PM 10 trend analysis”. **in** *Atmospheric Chemistry and Physics*: 18.9, **pages** 6223–6239.

Gray, Wayne B and Jay P Shimshack (2011). “The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence”. **in***Review of Environmental Economics and Policy*: 5.1, **pages** 3–24.

Greenstone, Michael (2003). “Estimating regulation-induced substitution: The effect of the Clean Air Act on water and ground pollution”. **in***American Economic Review*: 93.2, **pages** 442–448.

— (2004). “Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations?” **in***Journal of environmental economics and management*: 47.3, **pages** 585–611.

Greenstone, Michael and Ted Gayer (2009). “Quasi-experimental and experimental approaches to environmental economics”. **in***Journal of Environmental Economics and Management*: 57.1, **pages** 21–44.

Greenstone, Michael and Rema Hanna (2014). “Environmental regulations, air and water pollution, and infant mortality in India”. **in***American Economic Review*: 104.10, **pages** 3038–72.

Greenstone, Michael, Guojun He, Shanjun Li et al. (2021). *China’s War on Pollution: Evidence from the First Five Years*. techreport. National Bureau of Economic Research.

Greenstone, Michael, John A List and Chad Syverson (2012). *The effects of environmental regulation on the competitiveness of US manufacturing*. techreport. National Bureau of Economic Research.

Gupta, Pawan, Sundar A Christopher, Jun Wang et al. (2006). “Satellite remote sensing of particulate matter and air quality assessment over global cities”. **in***Atmospheric Environment*: 40.30, **pages** 5880–5892.

Hanna, Rema (2010). “US environmental regulation and FDI: evidence from a panel of US-based multinational firms”. **in***American Economic Journal: Applied Economics*: 2.3, **pages** 158–89.

Hanna, Rema Nadeem and Paulina Oliva (2010). “The impact of inspections on plant-level air emissions”. **in***The BE Journal of Economic Analysis & Policy*: 10.1.

Hao, Jiming, Shuxiao Wang, Bingjiang Liu et al. (2001). "Plotting of acid rain and sulfur dioxide pollution control zones and integrated control planning in China". **in** *Water, Air, and Soil Pollution*: 130.1, **pages** 259–264.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction.* Springer Science & Business Media.

He, Guojun, Maoyong Fan and Maigeng Zhou (2016). "The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games". **in** *Journal of Environmental Economics and Management*: 79, **pages** 18–39.

He, Guojun, Yuhang Pan and Takanao Tanaka (2020). "The short-term impacts of COVID-19 lockdown on urban air pollution in China". **in** *Nature Sustainability*: 3.12, **pages** 1005–1011.

He, Guojun and Shaoda Wang (2017). "Do college graduates serving as village officials help rural China?" **in** *American Economic Journal: Applied Economics*: 9.4, **pages** 186–215.

He, Kebin, Hong Huo and Qiang Zhang (2002). "Urban air pollution in China: current status, characteristics, and progress". **in** *Annual review of energy and the environment*: 27.1, **pages** 397–431.

Hering, Laura and Sandra Poncet (2014). "Environmental policy and exports: Evidence from Chinese cities". **in** *Journal of Environmental Economics and Management*: 68.2, **pages** 296–318.

Hu, Jianlin, Qi Ying, Yungang Wang et al. (2015). "Characterizing multi-pollutant air pollution in China: Comparison of three air quality indices". **in** *Environment international*: 84, **pages** 17–25.

Huang, Wei, Jianguo Tan, Haidong Kan et al. (2009). "Visibility, air quality and daily mortality in Shanghai, China". **in** *Science of the Total Environment*: 407.10, **pages** 3295–3300.

Imbens, Guido W and Donald B Rubin (2015). *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

Isaksen, Elisabeth Thuestad (2020). "Have international pollution protocols made a difference?" **in** *Journal of Environmental Economics and Management*: 103, **page** 102358.

Jia, Kai and Shaowei Chen (2019). "Could campaign-style enforcement improve environmental performance? Evidence from China's central environmental protection inspection". **in***Journal of environmental management*: 245, **pages** 282–290.

Jin, Yana, Henrik Andersson and Shiqiu Zhang (2016). "Air pollution control policies in China: a retrospective and prospects". **in***International Journal of Environmental Research and Public Health*: 13.12, **page** 1219.

Johnston, Andrew C and Alexandre Mas (2018). "Potential unemployment insurance duration and labor supply: The individual and market-level response to a benefit cut". **in***Journal of Political Economy*: 126.6, **pages** 2480–2522.

Jones, Alan M, Roy M Harrison and Jacob Baker (2010). "The wind speed dependence of the concentrations of airborne particulate matter and NOx". **in***Atmospheric Environment*: 44.13, **pages** 1682–1690.

Kagan, Robert A, Neil Gunningham and Dorothy Thornton (2003). "Explaining corporate environmental performance: how does regulation matter?" **in***Law & Society Review*: 37.1, **pages** 51–90.

Kahn, Matthew E, Pei Li and Daxuan Zhao (2015). "Water pollution progress at borders: the role of changes in China's political promotion incentives". **in***American Economic Journal: Economic Policy*: 7.4, **pages** 223–42.

Kaufman, Yoram J, Nadine Gobron, Bernard Pinty et al. (2002). "Relationship between surface reflectance in the visible and mid-IR used in MODIS aerosol algorithm-theory". **in***Geophysical Research Letters*: 29.23, **pages** 31–1.

Kleven, Henrik Jacobsen, Camille Landais and Emmanuel Saez (2013). "Taxation and international migration of superstars: Evidence from the European football market". **in***American economic review*: 103.5, **pages** 1892–1924.

Kostka, Genia and Yu Xiaofan (2015). "Career backgrounds of municipal party secretaries in China: why do so few municipal party secretaries rise from the county level?" **in***Modern China*: 41.5, **pages** 467–505.

Kreif, Noémi, Richard Grieve, Dominik Hangartner et al. (2016). "Examination of the synthetic control method for evaluating health policies with multiple treated units". **in***Health economics*: 25.12, **pages** 1514–1528.

Kumar, Naresh, Allen D Chu, Andrew D Foster et al. (2011). "Satellite remote sensing for developing time and space resolved estimates of ambient particulate in Cleveland, OH". **in***Aerosol Science and Technology*: 45.9, **pages** 1090–1108.

Kumar, Naresh and Andrew D Foster (2009). "Air quality interventions and spatial dynamics of air pollution in Delhi and its surroundings". **in***International journal of environment and waste management*: 4.1-2, **pages** 85–111.

Lagravinese, Raffaele, Francesco Moscone, Elisa Tosetti et al. (2014). "The impact of air pollution on hospital admissions: evidence from Italy". **in***Regional Science and Urban Economics*: 49, **pages** 278–285.

Laplante, Benoıt and Paul Rilstone (1996). "Environmental inspections and emissions of the pulp and paper industry in Quebec". **in***Journal of Environmental Economics and management*: 31.1, **pages** 19–36.

Li, Hongbin and Li-An Zhou (2005). "Political turnover and economic performance: the incentive role of personnel control in China". **in***Journal of public economics*: 89.9-10, **pages** 1743–1762.

Li, Xiao, Yuanbo Qiao, Junming Zhu et al. (2017). "The "APEC blue" endeavor: Causal effects of air pollution regulation on air quality in China". **in***Journal of cleaner production*: 168, **pages** 1381–1388.

List, John A, Daniel L Millimet, Per G Fredriksson et al. (2003). "Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator". **in***Review of Economics and Statistics*: 85.4, **pages** 944–952.

Liu, Haoran, Cheng Liu, Zhouqing Xie et al. (2016). "A paradox for air pollution controlling in China revealed by "APEC Blue" and "Parade Blue"". **in***Scientific reports*: 6.1, **pages** 1–13.

Maddison, David (2005). "Air pollution and hospital admissions: an ARMAX modelling approach". **in***Journal of Environmental Economics and Management*: 49.1, **pages** 116–131.

Mohan, Preeya (2017). "The economic impact of hurricanes on bananas: a case study of Dominica using synthetic control methods". **in***Food policy*: 68, **pages** 21–30.

Mullainathan, Sendhil and Jann Spiess (2017). "Machine learning: an applied econometric approach". **in***Journal of Economic Perspectives*: 31.2, **pages** 87–106.

Qin, Yu and Hongjia Zhu (2018). "Run away? Air pollution and emigration interests in China". **in***Journal of Population Economics*: 31.1, **pages** 235–266.

Rohde, Robert A and Richard A Muller (2015). "Air pollution in China: mapping of concentrations and sources". **in***PloS one*: 10.8, e0135749.

Rost, Jutta, Thomas Holst, Elke Sahn et al. (2009). "Variability of PM10 concentrations dependent on meteorological conditions". **in***International Journal of Environment and Pollution*: 36.1-3, **pages** 3–18.

Shang, Yu, Zhiwei Sun, Junji Cao et al. (2013). "Systematic review of Chinese studies of short-term exposure to air pollution and daily mortality". **in***Environment international*: 54, **pages** 100–111.

Shi, Zongbo, Congbo Song, Bowen Liu et al. (2021). "Abrupt but smaller than expected changes in surface air quality attributable to COVID-19 lockdowns". **in***Science Advances*: 7.3, eabd6696.

Shimshack, Jay P and Michael B Ward (2008). "Enforcement and over-compliance". **in***Journal of Environmental Economics and Management*: 55.1, **pages** 90–105.

Sun, Cong, Siqi Zheng and Rui Wang (2014). "Restricting driving for better traffic and clearer skies: did it work in Beijing?" **in***Transport Policy*: 32, **pages** 34–41.

Tanaka, Shinsuke (2015). "Environmental regulations on air pollution in China and their impact on infant mortality". **in***Journal of health economics*: 42, **pages** 90–103.

Tanaka, Shinsuke, Wesley Yin and Gary H Jefferson (2014). "Environmental regulation and industrial performance: Evidence from China". **in** *Work. Pap., Tufts Univ., Medford, MA*.

Telle, Kjetil (2013). "Monitoring and enforcement of environmental regulations: Lessons from a natural field experiment in Norway". **in** *Journal of Public Economics*: 99, **pages** 24–34.

Van Donkelaar, Aaron, Randall V Martin, Michael Brauer et al. (2010). "Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application". **in** *Environmental health perspectives*: 118.6, **pages** 847–855.

Varian, Hal R (2014). "Big data: New tricks for econometrics". **in** *Journal of Economic Perspectives*: 28.2, **pages** 3–28.

Venigalla, Mohan and Michael Krimmer (2006). "Impact of electronic toll collection and electronic screening on heavy-duty vehicle emissions". **in** *Transportation research record*: 1987.1, **pages** 11–20.

Viard, V Brian and Shihe Fu (2015). "The effect of Beijing's driving restrictions on pollution and economic activity". **in** *Journal of Public Economics*: 125, **pages** 98–115.

Vu, Tuan V, Zongbo Shi, Jing Cheng et al. (2019). "Assessing the impact of clean air action on air quality trends in Beijing using a machine learning technique". **in** *Atmospheric Chemistry and Physics*: 19.17, **pages** 11303–11314.

Wager, Stefan and Susan Athey (2018). "Estimation and inference of heterogeneous treatment effects using random forests". **in** *Journal of the American Statistical Association*: 113.523, **pages** 1228–1242.

Wang, Alex L (2013). "The search for sustainable legitimacy: environmental law and bureaucracy in China". **in** *Harv. Envtl. L. Rev.*: 37, **page** 365.

Wang, Hongru, Chengming Fan and Sicheng Chen (2021). "The impact of campaign-style enforcement on corporate environmental Action: Evidence from China's central environmental protection inspection". **in** *Journal of Cleaner Production*: 290, **page** 125881.

Wang, Siwen, Hang Su, Chuchu Chen et al. (2020). "Natural gas shortages during the "coal-to-gas" transition in China have caused a large redistribution of air pollution in winter 2017". **in** *Proceedings of the National Academy of Sciences*: 117.49, **pages** 31018–31025.

Wang, Wentao, Toby Primbs, Shu Tao et al. (2009). "Atmospheric particulate matter pollution during the 2008 Beijing Olympics". **in** *Environmental science & technology*: 43.14, **pages** 5314–5320.

Wang, Wenwen, Xinran Sun and Ming Zhang (2021). "Does the central environmental inspection effectively improve air pollution?-An empirical study of 290 prefecture-level cities in China". **in** *Journal of Environmental Management*: 286, **page** 112274.

Wang, Zhanshan, Yunting Li, Tian Chen et al. (2015). "Changes in atmospheric composition during the 2014 APEC conference in Beijing". **in** *Journal of Geophysical Research: Atmospheres*: 120.24, **pages** 12695–12707.

Wark, Kenneth, Cecil F Warner, Davis Wayne T et al. (1998). *Air pollution: its origin and control.* Addison-Wesley.

Wetterstrand, K A (2016). *DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP).* URL: www.genome.gov/sequencingcosts.

Wise, Erika K and Andrew C Comrie (2005). "Extending the Kolmogorov–Zurbenko filter: application to ozone, particulate matter, and meteorological trends". **in** *Journal of the Air & Waste Management Association*: 55.8, **pages** 1208–1216.

Wong, Christine and Valerie J Karplus (2017). "China's war on air pollution: can existing governance structures support new ambitions?" **in** *The China Quarterly*: 231, **pages** 662–684.

Wu, Huangjian, Xiao Tang, Zifa Wang et al. (2018). "Probabilistic automatic outlier detection for surface air quality measurements from the China national environmental monitoring network". **in** *Advances in Atmospheric Sciences*: 35.12, **pages** 1522–1532.

Wu, Jing, Yongheng Deng, Jun Huang et al. (2013). *Incentives and outcomes: China's environmental policy.* techreport. National Bureau of Economic Research.

Wu, Ruxin and Piao Hu (2019). "Does the "miracle drug" of environmental governance really improve air quality? Evidence from China's system of central environmental protection inspections". **in** *International journal of environmental research and public health*: 16.5, **page** 850.

Wu, Xiao, Rachel C Nethery, Benjamin M Sabath et al. (2020). "Exposure to air pollution and COVID-19 mortality in the United States". **in** *MedRxiv*.

Xiang, C and T van Gevelt (2020). "Central inspection teams and the enforcement of environmental regulations in China". **in** *Environmental Science & Policy*: 112, **pages** 431–439.

Xu, Xuchang, Changhe Chen, Haiyin Qi et al. (2005). "Power-sector energy consumption and pollution control in China". **in** *Urbanization, Energy, and Air Pollution in China: The Challenges Ahead: Proceedings of a Symposium*: National Academies Press, **page** 217.

Xu, Yiqing (2017). "Generalized synthetic control method: Causal inference with interactive fixed effects models". **in** *Political Analysis*: 25.1, **pages** 57–76.

Yin, Peng, Michael Brauer, Aaron J Cohen et al. (2020). "The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990–2017: an analysis for the Global Burden of Disease Study 2017". **in** *The Lancet Planetary Health*: 4.9, e386–e398.

Young, Oran R, Dan Guttman, Ye Qi et al. (2015). "Institutionalized governance processes: Comparing environmental problem solving in China and the United States". **in** *Global Environmental Change*: 31, **pages** 163–173.

Zhang, Wei, C-Y Cynthia Lin Lawell and Victoria I Umanskaya (2017). "The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence". **in***Journal of Environmental Economics and Management*: 82, **pages** 181–220.

Zhang, Yali, Wenqi Li and Feng Wu (2020). "Does energy transition improve air quality? Evidence derived from China's Winter Clean Heating Pilot (WCHP) project". **in***Energy*: 206, **page** 118130.

Zhang, Yangmei, Tuan Van Vu, Junying Sun et al. (2019). "Significant changes in chemistry of fine particles in wintertime Beijing from 2007 to 2017: Impact of clean air actions". **in***Environmental science & technology*: 54.3, **pages** 1344–1352.

Zhao, Bin, SX Wang, Huan Liu et al. (2013). "NO x emissions in China: historical trends and future perspectives". **in***Atmospheric Chemistry and Physics*: 13.19, **pages** 9869–9897.

Zhao, Huiyu and Robert Percival (2017). "Comparative environmental federalism: Subsidiarity and central regulation in the United States and China". **in***TEL*: 6, **page** 531.

Zheng, Siqi and Matthew E Kahn (2013). "Understanding China's urban pollution dynamics". **in***Journal of Economic Literature*: 51.3, **pages** 731–72.

— (2017). "A new era of pollution progress in urban China?" **in***Journal of Economic Perspectives*: 31.1, **pages** 71–92.

Zheng, Siqi, Matthew E Kahn, Weizeng Sun et al. (2012). "Urban Political Accountability in China". **in**.

— (2014). "Incentives for China's urban mayors to mitigate pollution externalities: the role of the central government and public environmentalism". **in***Regional Science and Urban Economics*: 47, **pages** 61–71.

Ziegler, Andreas and Inke R König (2014). "Mining data with random forests: current options for real-world applications". **in***Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*: 4.1, **pages** 55–63.