Erasmus University Rotterdam
Erasmus School of Economics

Bachelor Thesis Economics and Business Economics

Beijing air quality program: the synthetic control approach

Student name: Nikita Nesterov
Student ID number: 422428
Supervisor: Prof. Dr. O.R. Marie
Date final version: 17.07.18

Abstract

This paper contributes to the empirical literature on China's air pollution and builds upon the ideas of Abadie and Gardeazabal (2003), and Abadie, Diamond, and Hainmueller (2010, 2015). It implements a synthetic control method to assess the effectiveness of Jingjinji region-specific air action plan in reducing the fine particulate matter concentrations in post-2013 Beijing. The results suggest that following the policy enactment, $PM_{2.5}$ levels in the Chinese capital have dropped by 24.11%, but at most one-fifth of this reduction is attributable to the regional program, as the pollution levels in Beijing decreased by 4.67 $\mu g/m^3$ relative to the corresponding synthetic control unit. Besides, the analysis of mechanisms responsible for the air quality improvement reveals fundamental changes to final energy structure and an average cutback in consumption of non-renewable energy products of 47.85%.

1 Introduction

China is the world's fastest-developing major economy with growth rates averaging 10% in the recent decades (IMF, 2018). This unprecedented economic expansion does, however, come at a cost. Ever since China has surpassed the United States in its carbon dioxide emissions in 2007, not only has it become the world's biggest polluter (World Bank, 2018), but also the deadliest country to reside in as far as the air pollution is concerned (WHO, 2016). In China, environmental concerns have always been a distant second priority to the pursuit of economic prosperity, but air pollution is undeniably a significant issue which carries both high public health burden, as well as substantial financial losses.

In the winter of 2012-2013 China has experienced a multifold of geographically widespread long-lasting extreme air pollution incidents which have affected up to a quarter of China's land area and an estimated 600 million people (Huang et al., 2014; Ji et al., 2014). During 2013 alone, 1.6 million Chinese have died due to exposure to ambient air pollution, resulting in an economic loss equivalent to 10.9% of GDP (World Bank & IHME, 2016). More than half of these premature deaths were attributable to fine particulate matter exposure, which made it country's fifth leading cause of death (GBD MAPS Working Group, 2016). Beijing has received the hardest hit by far, with $PM_{2.5}$ concentrations exceeding the record range of monitoring instruments on more than one occasion in a series of events referred to as the "airpocalypse" (Lim, 2013).

Not only have these accidents received extensive international media attention, but they have also marked a turning point for the governmental approach to environmental issues. Freshly inaugurated president Xi Jinping has promptly put pollution on top of the legislative agenda, and issued a National Action Plan on Prevention and Control of Air Pollution in September of 2013. This program has set a number of quantitative targets for various regions, mainly focusing on the ambient particulate matter concentrations. The strictest targeted plan was set for Beijing and required an overall 25% reduction in PM2.5 levels with an annual upper threshold of $60\,\mu g/m^3$. This plan specific to Jingjinji, also known as Beijing-Tianjin-Hebei region has proposed various to achieve the target set by the government, which included a more efficient emission control, industrial optimization and restructuring, as well as an increased supply of clean energy (Ministry of Ecology and Environment of China, 2013). On the 31st of January 2018, the government has officially announced that the action plan has achieved "better-than-expected results" (Youbin, 2018) with a 39.6% drop in average PM2.5 concentrations of 39.6%. The research outlined in this paper is intended to critically assess this statement and evaluate the effectiveness of the Jingjinji regional air program. Consequently, the central question this paper aims to answer is as follows:

To what extent is the air quality improvement in Beijing attributable to Jingjinji region-specific action plan?

Economic science is frequently concerned with gauging the effects of policy interventions and large-scale events through comparative case studies which compare the development of variables of interest in an affected unit with the development of the same variable in unaffected units. The synthetic control method is one of the ways to conduct these studies and postulates that a group of controls generally yields a closer resemblance to the treated unit, than any single unit in particular. Besides, the synthetic control method offers a number of other advantages compared to a difference-in-differences (DiD) estimation, as it provides a greater degree of transparency and makes the relative contribution of each control unit explicit.

The first study that made use of a synthetic control method was performed in 2003 by Alberto Abadie who investigated the effect of terroristic activity on economic development in Basque Country by constructing a control group using two Spanish regions (Abadie & Gardeazabal, 2003). Later, Abadie has assessed the effect of California Proposition 99 on tobacco consumption (Abadie, Diamond, & Hainmueller, 2010) and the impact of 1990 German reunification on the economic growth (Abadie, Diamond, & Hainmueller, 2015).

This research draws inspiration from the studies mentioned above and applies the synthetic control method to assess the effectiveness of Jingjinji region-specific air action plan. It appears that no single specific region can accurately estimate the values of $PM_{2.5}$ pollution predictors for Beijing in the pre-treatment period, but a weighted average of control regions does provide an adequate approximation. The results suggest that following the program enactment, air quality in Beijing in terms of fine particulate matter concentrations has improved relative to the control unit. Results of the synthetic control method indicate that the ambient $PM_{2.5}$ levels in Beijing were on average 4.67 $\mu g/m^3$ lower than what they would have been in the absence of the program, which equals a 6.50% reduction. The ensuing difference-in-differences analysis provides a slightly higher estimate of 5.38 $\mu g/m^3$, which is statistically significant. Finally, the study looks into the mechanisms behind this air quality improvement, and the results suggest that Beijing has significantly reduced its usage of numerous non-renewable energy products, including coal, coke and diesel oil. These findings endure a series of robustness checks and placebo studies.

The rest of the paper is organized as follows: section 2 reviews the related literature, 3 discusses the main ideas behind the synthetic control method and outlines the model. Section 4 describes the data, and section 5 applies the synthetic control method to estimate the policy effect and presents the findings, section 6 assesses their robustness. Section 7 extends the synthetic control framework with a difference-in-differences estimation and investigates the mechanisms behind the $PM_{2.5}$ development. Finally, section 8 concludes while also highlighting the research limitations and providing suggestions for future studies.

2 Literature

Consequences of the exposure to poor air quality have always attracted plenty of scientific attention. Numerous environmental studies have established an association between exposure to fine particulate air pollution and mortality (Samet, Dominici, Curriero, Coursac, & Zeger, 2000), various health issues (Pope & Dockery, 2006), and a reduction in life expectancy (Pope, Ezzati, & Dockery, 2009).

Air pollution is currently one of China's most significant concerns and has been researched elaborately in the recent years. One paper has investigated the impact of exposure to air pollution on life expectancy in the Huai River region, confirming with robust empirical evidence the existence of a negative association between air pollution and longevity (Chen, Ebenstein, Greenstone, & Li, 2013). Another study has highlighted an overall increase in mortality due to exposure to ambient air pollution prior to the 2013 action plan (M. Liu et al., 2017).

Air quality in the Chinese capital has likewise received elaborate attention both before the air program, as well as after its implementation. Some studies have found that due to strict control actions during the 2008 Summer Olympics, the city's air quality has seen a considerable improvement. Much to Beijing's residents' dismay, this improvement has been short-lived, and the pollution reverted to the conventional levels in the consecutive months (Streets et al., 2007). 2014 APEC Summit was another occurrence where Beijing has experienced a significant air quality improvement. Likewise, this improvement has proven to be temporary, and air quality has deteriorated shortly after the emission controls were lifted (Sheng et al., 2015).

Similarly, long-term developments of air quality in the Chinese capital were looked into and are believed to generally be improving since the all-time high of 2013 (Zhang et al., 2016). Many think that Beijing's air quality has a reasonable chance of achieving the target set by the action plan (Wang et al., 2018). Finally, one study has evaluated the results of Beijing-specific air policy and concluded that the program had provided an effective approach to alleviate the $PM_{2.5}$ pollution in Beijing and the surrounding area (Cai et al., 2017).

3 Methodology

Following Abadie and Gardeazabal (2003), and Abadie, Diamond, and Hainmueller (2010, 2015), this paper evaluates the impact of the Jingjinji region-specific air program on Beijing ambient $PM_{2.5}$ concentrations by implementing a synthetic control method. This approach aims to approximate characteristics of Beijing during the pre-policy period using a combination of unaffected Chinese regions by constructing a synthetic control unit. Consequently, the post-policy counterfactual outcomes are used to estimate the results that would have been observed in Beijing in the absence of the policy.

3.1 Econometric model

To set the notation, suppose that $J + 1$ Chinese regions are observed and, without loss of generality, only Beijing is affected by the air program. This leaves J regions to form a pool of controls indicated by $i = 1, \ldots, J + 1$. Let $t = 1, \ldots, T$ denote the time frame with $t = 1, \ldots, T_0$ standing for the period prior to passage of the program and $t = T_0 + 1, \ldots, T$ the period after its implementation.

Let Y_{it}^N be the ambient $PM_{2.5}$ levels that would be observed in region i at time t in the absence of the intervention, and Y_{it}^I the outcome that would be observed if the region was exposed to intervention in periods $t = T_0 + 1, \ldots, T$. Additionally, assume that the policy has no effect prior to the period from when on it is implemented: $Y_{it}^I = Y_{it}^N$ for $t = 1, \ldots, T_0$. The final assumption is that of no inference between units which implies that the pollution levels in control regions are not affected by the Beijing air program. This inference is addressed later in the analysis.

The model aims to estimate the effect of the air program on Beijing ambient fine particulate matter concentrations: $\alpha_{it} = Y_{it}^I - Y_{it}^N$. Let D_{it} be the indicator that takes on the value of one if region i is exposed to the intervention at time t, and zero otherwise. Under the assumption that only the first region is exposed and only after the period T_0, $D_{it} = 1$ if and only if $i = 1$ and $t > T_0$. Hence, the indicator takes on this value exclusively for Beijing during the post-intervention period. Then, the observed outcome can be written as: $Y_{it} = Y_{it}^N + \alpha_{it} D_{it}$. The objective is to estimate the program effect: $\alpha_{1t} = Y_{1t}^I - Y_{1t}^N$, where Y_{1t}^I is observed, but Y_{1t}^N is not. The synthetic control method constructs a counterfactual that yields an estimate for this unobserved variable by means of the following model:

$$Y_{1t}^N = \delta_t + \theta_t \zeta_i + \lambda_t \eta_i + \varepsilon_{it} \tag{1}$$

here, δ_t denotes the unknown common factor, θ_t is a $(1 \times r)$ vector of unknown parameters, ζ_i is a $(r \times 1)$ vector of observed outcomes not affected by the intervention, λ_t is a $(1 \times f)$ vector of unobserved common factors, and η_i is a $(f \times 1)$ vector of unknown factor loadings. Finally, ε_{it} denotes the unobserved unit-level transitory shocks with the mean of zero.

3.2 Weights estimation

This subsection outlines the way in which weights of the control regions are estimated to form a suitable control unit. Consider a $(J \times 1)$ vector of weights W with $W = (w_2, \ldots, w_{J+1})$ such that $w_j \geq 0$ and $w_2 + \cdots + w_{J+1} = 1$. This vector is meant to minimize the difference in the pre-policy characteristics between Beijing and the weighted average of other Chinese regions. Borrowing from equation (1), the value of the outcome variable is given by

$$\sum_{j=2}^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_{j=2}^{J+1} w_j \zeta_j + \lambda_t \sum_{j=2}^{J+1} w_j \eta_j + \sum_{j=2}^{J+1} w_j \varepsilon_{jt} \tag{2}$$

Assuming that there exist optimal weights w^* such that

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \sum_{j=2}^{J+1} w_j^* Y_{j2} = Y_{12}, \ldots, \sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0}; \sum_{j=2}^{J+1} w_j^* Z_j = Z_1 \tag{3}$$

and given a sufficient number of pre-policy periods, α_{1t} can be estimated as $\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$. In actuality, equation (2) does not strictly hold, but to an extent determined by how well synthetic control matches Beijing's true characteristics. Essentially, equation (1) generalizes the commonly used difference-in-differences model, which can still be obtained within this framework if λ_t is kept constant for all t. While the difference-in-differences approach allows for unobserved confounders, it restricts their effect to be constant over time, so they can be differenced out. Synthetic control method, on the other hand, does allow the unobserved characteristics to vary with time, but as a result, they can't be eliminated by taking the time differences. In spite of that, a synthetic control which satisfies

$$\sum_{j=2}^{J+1} w_j^* \zeta_j = \zeta_1; \sum_{j=2}^{J+1} w_j^* \eta_j = \eta_1 \tag{4}$$

will provide an unbiased estimator of Y_{it}^N. However, since $\eta_1, \ldots, \eta_{J+1}$ are unobserved, choosing a synthetic control in this way is not feasible. Still, the factor model $Y_{it} = Y_{it}^N + \alpha_{it} D_{it}$ implies that a synthetic control can fit ζ_1 and a set of pre-intervention outcomes $Y_{11}, \ldots, Y_{1T_0}$, as long as it fits both ζ_1 and η_1, hence equation (3) does approximately hold. The methodology outlined in this section is put into practice using *R* and specifically, *synth* extension package (Heinmueller, 2014). Appendix 1 provides information on the implementation.

3.3 Difference-in-differences estimation

One of the main limiting factors of the synthetic control method is the inability to accurately gauge the statistical significance of policy effect. A difference-in-differences estimation following Card and Krueger (1993) does allow for this inference. Besides, DiD enables this research to address the previously specified no inference assumption. There is a reason to believe that the improvement in Beijing air quality could have positively influenced the air quality in the neighboring Shenyang, thereby violating the assumption. If that is the case, the estimated effect would be lower, as the particulate matter concentrations in the control region would be artificially reduced. Additionally, the DiD framework provides a way to assess the development of mechanisms responsible for the changes in $PM_{2.5}$ levels. This estimation technique is less sophisticated compared to the synthetic control method and is outlined in Appendix 2.

4 Data

4.1 Background

Studies commonly turn to official government data in assessing the effectiveness of policies, but there are a few issues associated with it. Firstly, China's Ministry of Ecology and Environment is severely underfunded and as a result, does not have the resources to quantify the extent of pollution and provide credible oversight. More importantly, Chinese politicians are highly sensitive when it comes to publicizing environmental data. The 2007 China's Ministry of Land study illustrates this conjecture best, as the results were ready for publication within eighteen months, but were not disclosed until the end of 2013 and even then, to an extent. The full scope of the findings would have created a major liability for the central government as according to the study, 16.1% of China's soil and about 19.4% of farmland appeared to be contaminated (Guangwei, 2014).

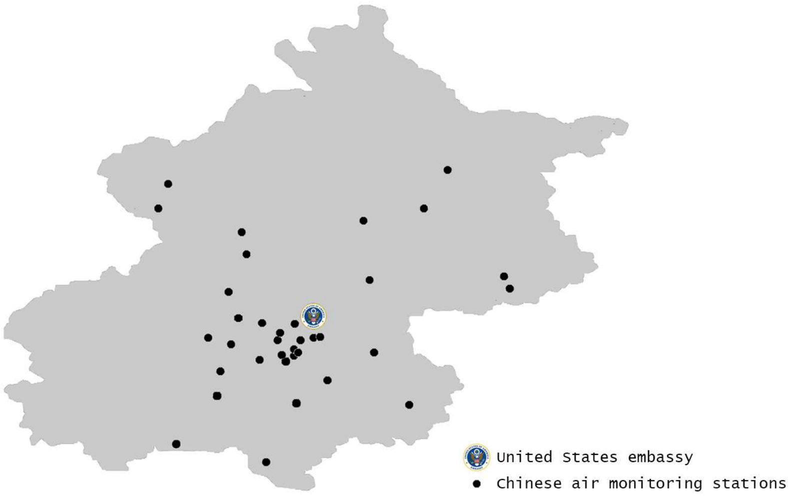

Figure 1: Locations of the air monitoring stations

4.2 Variable of interest

On the grounds that the official data can neither be considered comprehensive nor genuinely reliable, this paper considers data which originates from the United States embassy and consulates equipped with the air quality monitors for measuring fine particulate matter concentrations. $PM_{2.5}$ particulates are less than 2.5 micrometers in diameter and are believed to pose the largest health risks, as they are small enough to directly enter the lungs and the bloodstream (Fann et al., 2012). Besides, $PM_{2.5}$ is a globally recognized air quality standard which allows for international comparison.

No data is perfect in every respect and the United States Beijing observations emerge from a single monitoring station, whereas the government possesses 34, as illustrated in Figure 1 (Beijing Government, 2018). The U.S. provides hourly observations for five major Chinese cities for a period between 2012 and 2016, which are publicly accessible. In February 2014 this caused major civil unrest as the United States Beijing Air Quality Index indicated that the city's air quality was at or above the "unhealthy for sensitive groups" level for 70% and "hazardous" or beyond for 25% of the time (United States Embassy Beijing, 2014).

The timeframe of this study spans from 2012 up to 2016. While it is desirable to have an extended pre-treatment period, the sample begins in early 2012, as only then, $PM_{2.5}$ concentration readings are available for all control regions. The policy was passed in September of 2013 resulting in respectively 20 and 28 months of pre- and post-intervention data.[1]

[1] For the purpose of this research, hourly data is aggregated to a monthly level.

4.3 STL decomposition

Pollution readings exhibit a high degree of seasonality which needs to be addressed before carrying out the research. There are numerous ways in which observations can be stripped of its seasonal component. Most commonly used smoothing technique is the simple and exponential moving averages. This paper takes a different approach in arriving to seasonally-adjusted data, namely the Seasonal Trend decomposition by Loess (STL). This method developed by Cleveland, Cleveland, McRae, and Terpenning (1990) splits the time series into three components: trend, seasonal and remainder. STL has a number of advantages over the moving averages decomposition, as it not only can handle any type of seasonality but also allows the seasonal components to vary over time, is robust to outliers, and doesn't sacrifice any of the first and last observations.

Figure 2 illustrates the time series decomposition for Beijing observations. The raw data is shown in the top panel, the second and third panels graph the seasonal and trend component, respectively. The remainder, displayed in the fourth panel is the residual variation. The trend indicates a decline in the ambient $PM_{2.5}$ levels in late 2014, which is partially attributable to the air program. In addition, Table 1 provides descriptive statistics for the seasonally-adjusted particulate matter concentrations, the main variable of interest. Statistics suggest that Shenyang had the highest average $PM_{2.5}$levels, closely followed by Beijing. These alarmingly high estimates are negative externalities of the heavy industry located in Northeast China.

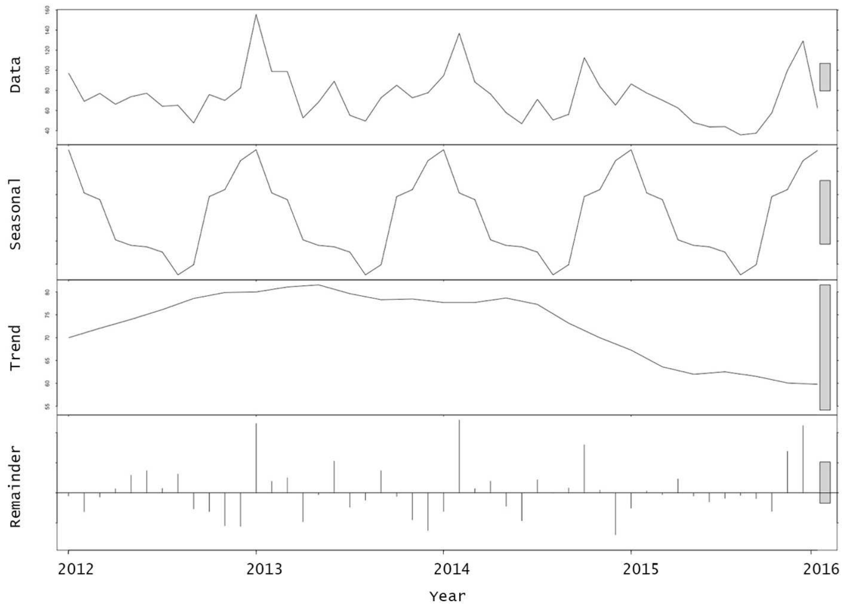

Figure 2: STL decomposition of Beijing $PM_{2.5}$ observations

Table 1: Descriptive statistics of the seasonally-adjusted $PM_{2.5}$ concentrations

	Average	Median	Std. Dev.	Min	Max
Beijing	82.14	74.90	9.40	59.54	86.63
Chengdu	75.11	76.40	8.95	56.11	89.09
Guangzhou	46.53	48.39	9.91	33.24	60.69
Shanghai	49.97	51.30	4.99	39.07	56.27
Shenyang	84.20	82.77	7.95	54.84	92.91

4.4 Predictor variables

The availability of data constraints the pre-treatment period for the depended variable to 20 months preceding the intervention. Statistics on the predictor variables are compiled from other sources allowing for an additional 11 years of data. The primary goal of synthetic control is to approximate characteristics of the treated unit during the pre-treatment period, hence extending the timeframe should positively influence the precision.

4.4.1 Population density

Population density is first on the list of predictor variables for the $PM_{2.5}$ pollution, as it is proven to be positively associated with particulate matter emissions (He et al., 2001). The data on this variable is derived from the Chinese national population surveys and considers metropolitan statistical area population densities.

4.4.2 GDP per capita

Capita-adjusted GDP is the second predictor variable and is a commonly used measure of financial development. Data on this variable originates from the China statistical yearbooks and is measured in 2013 RMB. Economic advancement is generally considered one of the primary driving forces behind the $PM_{2.5}$ emissions (Guan et al., 2014) with a positive correlation between GDP values and pollution readings (Selden & Song, 1994).

4.4.3 Meteorological indicators

Meteorological conditions greatly affect the fine particulate matter concentrations implying that different weather conditions can induce different pollution readings even if the underlying emissions are the same (Liang et al., 2015). Commonly, non-parametric models and Kernel statistical learning are used to account for meteorological factors. These are a powerful tool for weather-adjusting the data but are beyond the scope of this paper. As pollution levels need to ideally be compared under similar weather conditions, air temperature, wind speed, and relative humidity are added to the list of predictor variables. Data on these variables is compiled from the Weather Underground weather service website. Air temperature is argued to be positively correlated with the $PM_{2.5}$ levels as the higher temperature can lead to slightly higher pollution readings (Y. Liu, Franklin, Kahn, & Koutrakis, 2007). Wind speed is also proven to be influential in studies on particulate matter pollution as wind transferred dust greatly contributes to the $PM_{2.5}$ concentrations (Hueglin et al., 2005). Therefore, higher average wind speed can lead to an increase in particulate matter pollution. Relative humidity is defined as the amount of ambient water vapor as a percentage of volume needed for saturation. An increase in precipitation with a consequent rise in relative humidity is proven to cause a decrease in fine particulate matter concentrations through a process known as "scavenging" (Liao, Chen, & Seinfeld, 2006).

4.4.4 Air pollution indicators

This paper primarily looks into the developments of fine particulate matter concentrations, but since ambient pollutants are typically intercorrelated, particulate matter (PM_{10}), sulfur dioxide (SO_2) and nitrogen dioxide (NO_2) are taken into account. Annual statistics on these variables are derived from the China statistical yearbooks and unlike $PM_{2.5}$pollution, cover the entire pre-treatment period. Particulate matter (PM_{10}) generally exhibits a strong positive correlation with fine particulate matter (Airborne Particles Expert Group, 1999). Sulfur dioxide and nitrogen dioxide also show a positive correlation of 0.45 (Venners et al., 2003) and 0.70 (Beckerman et al., 2008), respectively.

4.4.5 Human impact indicators

Human activity is the leading cause of $PM_{2.5}$ pollution. Hence, this paper considers a number of consumption indicators of various energy sources, including coal, coke, crude oil, gasoline, kerosene, diesel oil, natural gas, and electricity. Data on this last group of variables is compiled from Chinese national bureau of statistics. Natural gas is measured in cubic meters, electricity in kilowatt hours, and the remaining predictors in kilograms.

Production and consumption of coal and crude oil are proven to be positively correlated with fine particulate matter levels and contribute greatly to its composition (Ito et al., 2011). Traffic exhaust fumes are also a significant source of $PM_{2.5}$ pollution (De Kok, Driece, Hogervorst, & Briedé, 2006), as well as electricity (Davidson, Phalen, & Solomon, 2005) due to its origin, which in China lies primarily in non-renewable energy.

4.5 Descriptive statistics

Table 2 summarizes the descriptive statistics of the predictor variables. The regions are vastly different in their characteristics and comparing Beijing to a single control, or a simple weighted average is unlikely to yield accurate results. Synthetic control estimation offers a solution by fine-tuning weights of untreated

Table 2: Descriptive statistics of the predictor variables

Variable	Average	Median	Std. Dev.	Min	Max
Population density	1548.20	1708.00	646.19	630.00	2177.00
GDP per capita	17872.80	19718.00	5274.01	11204.00	24230.00
Temperature	15.31	17.77	6.59	5.48	22.80
Relative humidity	63.54	61.02	6.41	58.29	73.11
Wind power	9.20	8.87	3.01	5.09	13.45
Particulate matter	100.16	110.00	17.17	80.00	116.00
Sulfur dioxide	42.20	41.30	3.49	39.90	48.40
Nitrogen dioxide	53.56	53.60	3.33	48.80	57.10
Coal	1940.58	1755.26	989.09	390.07	4151.06
Coke	248.71	217.11	206.34	15.60	784.25
Crude oil	631.95	519.95	468.59	4.67	1611.92
Gasoline	116.88	106.82	62.53	17.25	263.57
Kerosene	70.10	25.07	63.26	4.33	273.48
Diesel oil	129.96	121.01	65.35	19.40	279.22
Natural gas	160.53	128.56	144.51	1.41	746.94
Electricity	3446.11	3555.04	1406.27	625.80	6140.58

regions to construct a fitting comparison unit. The statistics display a high degree of variation for the majority of the predictor variables, yet the most striking inequalities are observed in the consumption of energy products with standard deviations reaching values as high as 90% of the average. The air pollution indicators, on the other hand, are moderately stable across regions.

5 Results

Figure 3 plots the development of fine particulate matter concentrations for Beijing and the average of unaffected regions. During the entire pre-intervention period, Beijing shows considerably higher $PM_{2.5}$ concentrations compared to the controls. Specifically, at the time of policy enactment, Chinese capital has over 50% higher pollution level. Air quality begins to improve a year after policy implementation as fine particulate matter levels in Beijing experience a noticeably sharper decline compared to untreated regions. This development serves as a preliminary indication of the policy effect. However, to accurately assess the effectiveness of the program, a synthetic control which plots the theoretical $PM_{2.5}$concentrations in Beijing in the absence of intervention needs to be constructed. A fitting synthetic control is capable of clearly illustrating any discrepancies caused by the policy.

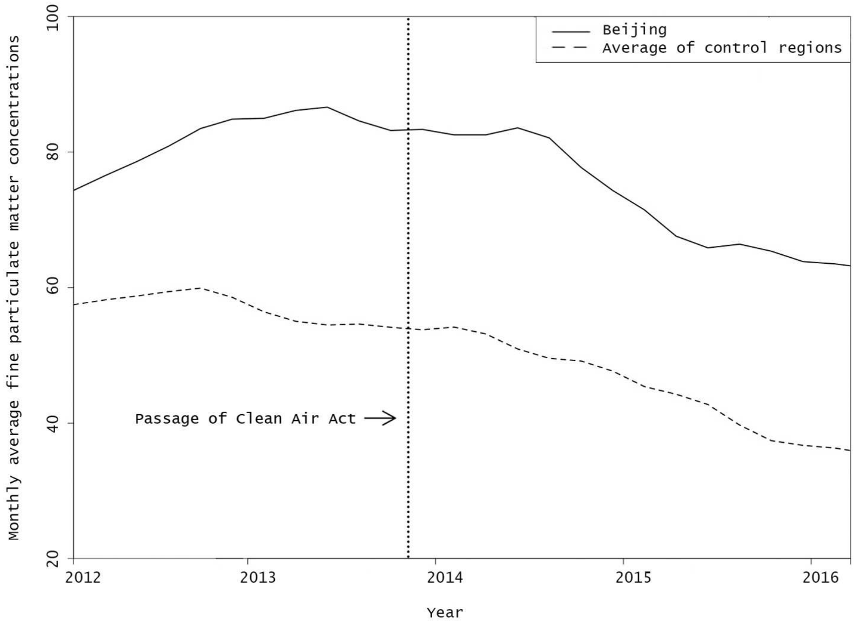

Figure 3: Trends in $PM_{2.5}$ concentrations for Beijing and the average of control group

Table 3: Control region weights in Synthetic Beijing

Region	Weight
Chengdu	0.40
Guangzhou	0.01
Shanghai	0.15
Shenyang	0.44

Table 3 provides the weights assigned to control regions which make up a synthetic control unit. By design, these weights sum up to one. Beijing carries the highest resemblance with the neighboring city of Shenyang, and Chengdu follows in a close second place. Shanghai and Guangzhou appear to be vastly different in their pre-intervention characteristics and as a result, end up in a distant third and fourth place, respectively.

Table 4 compares the actual pre-treatment characteristics of Beijing with its synthetic counterpart and the average of control regions. The estimates suggest that the mean values of predictor variables in the unaffected areas don't provide a sufficiently accurate comparison. On average, the values of the control group deviate from Beijing by 29.40%. The synthetic control unit offers an almost three times more precise approximation with an average deviation of 9.96%.

Table 4: Mean values of the predictor variables during pre-intervention period

Variable	Beijing	Synthetic Beijing	Average of controls
Population density	1167.00	1474.00	1643.00
GDP per capita	20407.00	13651.76	17239.25
Temperature	12.53	12.50	16.01
Relative humidity	56.29	60.85	64.86
Wind power	8.72	8.56	9.36
Particulate matter	116.00	107.53	96.75
Sulfur dioxide	39.90	40.24	43.03
Nitrogen dioxide	56.20	54.49	52.90
Coal	1439.92	1476.42	2065.74
Coke	199.31	207.45	258.15
Crude oil	519.39	549.87	660.09
Gasoline	168.12	157.35	104.07
Kerosene	170.89	121.77	44.90
Diesel oil	96.51	106.56	138.33
Natural gas	337.68	286.50	114.90
Electricity	3819.93	3710.03	3352.66
Average deviation		0.096	0.294

Beijing is nearly identical to its synthetic counterpart in terms of air temperature, wind power, and sulfur dioxide concentrations. In general, every predictor variable for the per capita GDP is more accurately approximated by the synthetic control as compared to a simple average. The high difference in the GDP values is directly induced by the composition of the counterfactual. The relatively less wealthy Shenyang and Chengdu make up for 84% of the total resulting in lower GDP value compared to when regions are weighted equally.

It is worth noting, that pioneering studies which made use of the synthetic control method have generally accomplished a near perfect fit with less than 1% discrepancy in (Abadie et al., 2010), and 6% in (Abadie et al., 2015).[2] It is challenging to estimate a precise control unit due to a limited size of the donor pool, and this is the best obtainable approximation. It is far from perfect, but still sufficiently accurate to allow for plausible inferences with regard to the policy effect, as illustrated in Figure 4 which plots the development of ambient $PM_{2.5}$ pollution for Beijing and its synthetic counterpart.

[2] Author's calculations based on the mean values of predictor variables in the quoted studies.

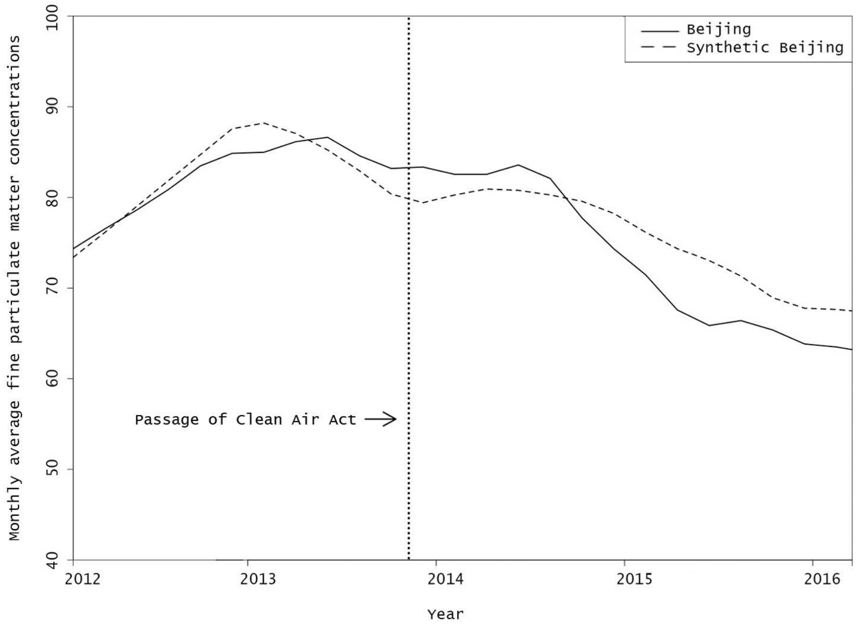

Figure 4: Trends in $PM_{2.5}$ concentrations for Beijing and Synthetic Beijing

In contrast to equally weighted controls, synthetic Beijing tracks the trajectory of real Beijing far more closely during the intervention period and provides a reasonable approximation for the $PM_{2.5}$ concentrations that would have been observed in Beijing post-2013 in the absence of the air program. Due to the developments in the Chinese environmental legislation, particulate matter pollution trends are mostly downward sloping. Following the policy enactment in September of 2013, Beijing $PM_{2.5}$levels have remained virtually unchanged for nearly a year. It isn't uncommon for environmental programs to exhibit a delayed impact, as established by among others, Popp (2003), and Divan and Rosencranz (2001) who addressed the effectiveness of 1970 U.S. clean air act and 1974 act on pollution prevention and control in India, respectively. The lag is caused by the fact that the pollution-reducing measures, such as transmission to sustainable energy and industry restructuring are time-consuming and provide gradual rather than instant results. Later part of the analysis highlights the developments in the consumption of non-renewable energy in detail. One thing is clear: in the absence of the region-specific policy, Beijing would have still experienced an air quality improvement, albeit of a lesser magnitude.

On average, the policy has reduced the $PM_{2.5}$ pollution by an additional 4.67 $\mu g/m^3$per month, or 6.50% compared to the baseline level. Fine particulate matter concentrations have dropped by a total of 24.11% since the policy effect was first observed, but at most one-fifth of this reduction is directly attributable to the region-specific program. The regional plan appears insignificant on its

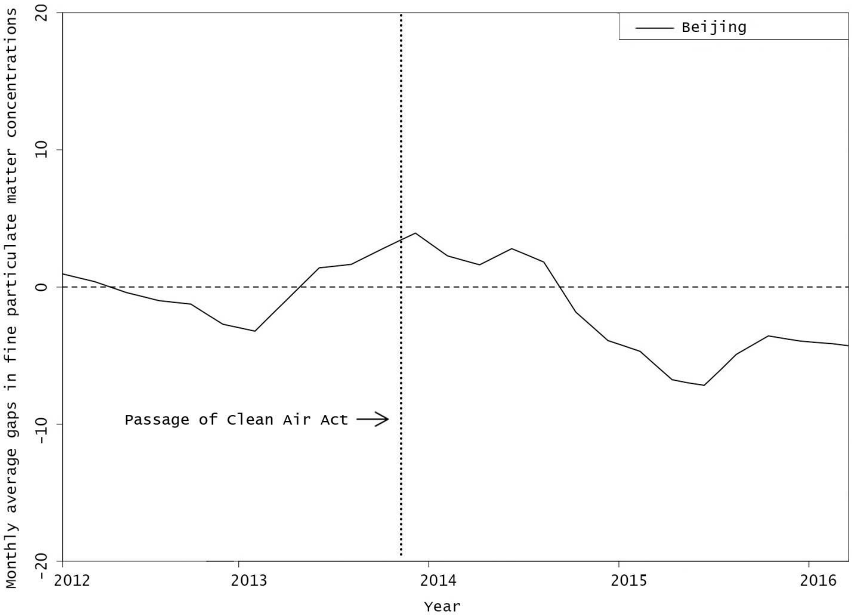

Figure 5: Gaps in $PM_{2.5}$concentrations between Beijing and Synthetic Beijing

own, but in combination with the ongoing air quality improvement, the Jingjinji program has undoubtedly contributed to Beijing's pursuit of 60 µg/m^{3}annual threshold. The timeframe of this study doesn't extend into 2018 as the respective data is not publically available yet. However, a proclaimed 39.6% reduction certainly seems attainable based on the observed developments.

Lastly, Figure 5 looks at the policy effect from an alternative angel by plotting the differences in $PM_{2.5}$ levels between Beijing and its counterfactual. It is possible that the observed differences are created artificially by the lack of fit and are not indicative of policy effect. This concern is addressed in the following section which compares Beijing gaps with those of the untreated regions.

6 Placebo studies and robustness checks

Placebo studies assess the robustness of the estimates by trying to answer the question whether the results of comparable magnitude and direction could be driven by chance. In this effort, the synthetic control method is applied to the untreated regions, and if gaps similar to the one estimated for treated unit one are observed, then analysis fails to provide sufficient evidence for the presence of the policy effect. On the other hand, if the differences in Beijing observations appear to be unusually large relative to the placebo gaps, then the program is believed to have been effective in reducing the pollution.

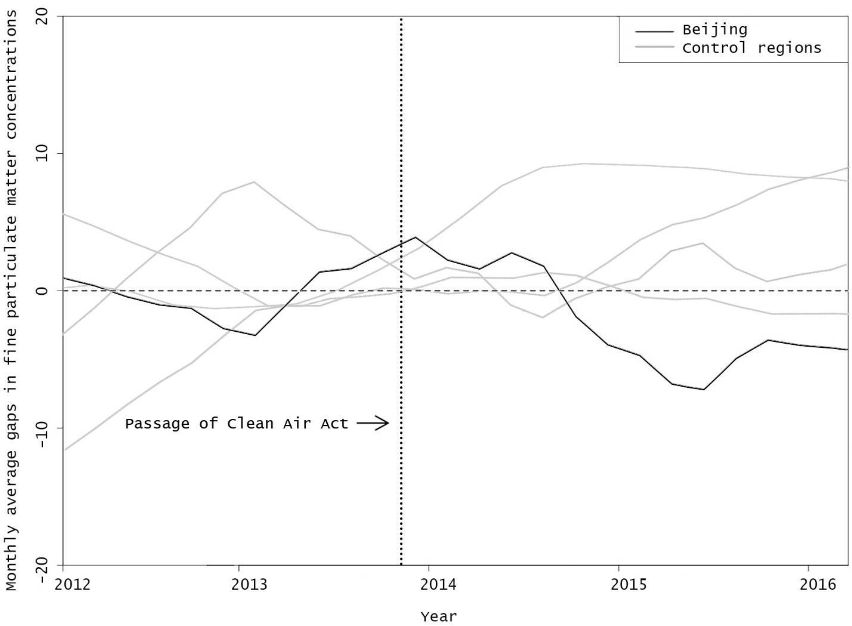

Figure 6: Gaps in $PM_{2.5}$ concentrations in Beijing and placebo gaps in control regions

Table 5: Ratios of post- and pre-intervention mean squared prediction errors

Region	Ratio
Beijing	3.98
Chengdu	0.85
Guangzhou	2.37
Shanghai	1.12
Shenyang	2.86

Figure 6 shows the results of the placebo studies. Gray lines represent gaps associated with the control regions measured as differences in $PM_{2.5}$ levels between each unit and its synthetic counterpart. The black line represents Beijing gaps borrowed directly from Figure 5. The latter appears to reach the highest negative value in the post-intervention period with a $7.26\,\mu g/m^3$ difference in particulate matter pollution concentrations, which is over 2.5 times greater than the value of the largest placebo gap. However, the precision of the placebo tests leaves much to be desired. Hence, the unusually large positive gaps are most likely caused by the lack of fit, as opposed to an intervention effect. A way to assign goodness of fit values to these observations is the mean squared prediction error (MSPE) which measures the magnitude of the gap between each region and its synthetic counterpart. Note that a large post-intervention MSPE is not indicative of a significant policy effect if the pre-intervention MSPE is also large. Intuitively, a high post-intervention prediction error is not indicative of policy effect if the synthetic control doesn't closely reproduce the outcome of interest during the pre-treatment period.

Table 5 reports the ratios of post- and pre-intervention prediction errors, and Beijing stands out with the post-intervention gap nearly four times as large as the pre-intervention gap. To put it into perspective, Abadie et al. (2015) have estimated a ratio of over 15, and Abadie et al. (2010) an impressive 130. Consequently, this analysis can hardly be considered indicative on its own and it remains challenging to accurately establish the significance of the policy effect. While the synthetic control method is a more sophisticated way to gauge the effectiveness of a policy, the difference-in-differences analysis can numerically measure the statistical significance of an estimate.

7 Extensions

7.1 Significance of the estimates

In addition to outlining the general intuition behind the difference-in-differences estimation, Appendix 2 specifies the parallel trend assumption necessary for the unbiasedness of the DiD estimator. Figure 3 at the beginning of the results section demonstrates its validity for the Beijing observations, as the time series follow similar paths in the pre-intervention period. The difference-in-differences estimation outlined in this section weights control regions both conventionally, as well as according to synthetic control estimates.

Table 6 shows the regression results. The first thing to note is the significance of policy effect coefficients in both weighting alternations. Secondly, the coefficient of synthetically weighted regression which indicates a $PM_{2.5}$ reduction of 5.38 $\mu g/m^3$ is similar to the estimate of a synthetic control method. Other estimates also fall in line with the earlier observations. The time trend coefficient is negative and statistically significant. Moreover, the regional difference

Table 6: Results of the difference-in-differences analysis for $PM_{2.5}$concentrations

Beijing $PM_{2.5}$ levels	Equal weights	Synthetic weights	Placebo time
Air action plan effect	-4.515**	-5.380***	-2.167
	(1.938)	(1.976)	(1.540)
Time trend	-10.160***	-9.010***	-13.075***
	(1.181)	(1.291)	(1.314)
Regional difference	14.921***	6.359***	11.140***
	(1.125)	(1.170)	(1.176)
R squared	0.497	0.783	0.364
Sample size	132	132	132

*p-value < 0.1, **p-value < 0.05, ***p-value < 0.01

coefficient indicates that the $PM_{2.5}$levels in Beijing are significantly higher compared to control regions, especially when they are weighted equally. Besides, the policy coefficient of a placebo policy of one year earlier is both smaller in value and not statistically significant, indicating that there was no significant change in $PM_{2.5}$concentrations a year prior to the policy enactment. In conclusion, the research presents sufficient empirical evidence to reject the null hypothesis that the policy effect is not significantly different from zero.

7.2 Spatial displacement

Geographical displacement of a policy effect can violate the previously made assumption of no inference between units. As Figure 8 illustrates, the most likely candidate to benefit from the Beijing air program is the neighboring Shenyang. Placebo tests assign treatment status to each of the control regions to assess whether spatial diffusion has indeed occurred. Table 7 summarizes the results. The coefficients were mostly to be expected: statistically significant negative time trend and large variations in $PM_{2.5}$concentrations across regions. Most importantly, the policy effect coefficient for Beijing stands out as it not only is statistically significant but also takes on the highest negative value. Guangzhou also appears to have a highly significant negative coefficient of policy effect, but a considerable distance between two regions makes it highly unlikely that Guangzhou could have benefited from Beijing air program. Seemingly, neither did Shenyang benefit from the pollution controls, as the coefficient is not only relatively low in value, but also not statistically significant. Hence, the null hypothesis concerning the policy effect on the $PM_{2.5}$ levels in Shenyang cannot be rejected as the associated coefficient appears not to be significantly different from zero.

Figure 7: Geographical locations of the regions

Table 7: Results of the placebo difference-in-differences analysis

$PM_{2.5}$ levels	Beijing	Shanghai	Guangzhou	Chengdu	Shenyang
Policy effect	-5.380***	3.493*	-1.494***	5.282	1.012
	(1.976)	(2.018)	(0.868)	(2.050)	(0.302)
Time trend	-9.010***	-11.819***	-8.210**	-8.650**	-11.671***
	(1.291)	(1.383)	(1.183)	(1.044)	(1.188)
Regional difference	6.359***	-19.253***	-14.456***	9.530**	28.176***
	(1.170)	(1.409)	(0.972)	(1.588)	(1.066)
R squared	0.783	0.747	0.832	0.639	0.578
Sample size	132	132	132	132	132

*p-value < 0.1, **p-value < 0.05, ***p-value < 0.01

7.3 Mechanisms behind the reduction

The study has successfully established the significance of the effect of Beijing air program on ambient particulate matter pollution. The final step is to determine the factors which have contributed to the observed changes in $PM_{2.5}$ levels. The last part of the analysis takes a close look at the consumption of some of the non-renewable energy products, considered earlier as predictor variables. Table 8 shows the results of difference-in-difference estimation for each energy products. Firstly, the air policy enactment has seemingly influenced consumption of the majority of considered variables, with only gasoline consumption staying unaffected. That is likely to due it being highly inelastic in short- and mid-run, as it is used in cars. The time trend coefficients show an increase in consumption across the board with only the coefficient of diesel oil consumption not being statistically significant. Lastly, Beijing differs significantly from other regions in consumption of all but one energy product, namely the electricity.

While regression estimates provide a way to make inferences with regard to the significance of the effects, it doesn't offer a suitable framework to compare consumption changes. Table 9 shows the percentage differences for the considered variables in Beijing, a control region and the difference in differences which indicates the program effect. Following the passage of the air program, Beijing has significantly reduced its consumption of heavy industrial products, namely coal, coke, crude oil, and diesel oil. The utilization of gasoline, kerosene and natural gas, on the other hand, has increased. If all products are considered to contribute to $PM_{2.5}$ pollution evenly, then Beijing has reduced the average consumption by 47.85%. Of course, not every form of energy is equally polluting, but a precise analysis of the contribution of each is beyond the scope of this paper. The main takeaway is that following the policy implementation, Beijing has adjusted its energy structure by reducing consumption of coal and coke in favor of low-carbon natural gas, the result which lies in line with the predictions of previous studies Feng, Chen, and Zhang (2013).

Table 8: Changes in the consumption of non-renewable energy products

Variable	Coal	Coke	Crude oil	Gasoline	Kerosene	Diesel oil	Natural gas	Electricity
Air action plan effect	-1398.867***	-286.107**	-263.504***	-12.161	81.784**	-68.361***	277.689***	-785.806***
	(195.748)	(40.751)	(53.115)	(17.493)	(17.026)	(15.233)	(67.494)	(313.277)
Time trend	352.133***	101.757***	165.205**	64.785***	18.358***	54.286	102.304**	1586.478**
	(130.867)	(28.004)	(43.171)	(12.352)	(3.773)	(13.622)	(14.821)	(264.406)
Regional difference	-689.642***	-87.237**	-293.984***	62.912***	125.691**	-47.981***	205.170***	348.576
	(158.903)	(40.208)	(39.215)	(15.995)	(13.895)	(14.798)	(37.232)	(294.312)
R squared	0.701	0.592	0.796	0.521	0.849	0.603	0.779	0.552
Sample size	136	136	136	136	136	136	136	136

*p-value $<$ 0.1, **p-value $<$ 0.05, ***p-value $<$ 0.01

Table 9: Percentage changes in the consumption of non-renewable energy products

Variable	Coal	Coke	Crude oil	Gasoline	Kerosene	Diesel oil	Natural gas	Electricity
Beijing	-0.644	-0.962	-0.183	0.331	0.654	-0.142	1.404	0.218
Control	0.152	0.364	0.199	0.675	0.667	0.369	1.563	0.476
Difference-in-differences	-0.796	-1.326	-0.382	-0.344	-0.013	-0.512	-0.159	-0.259

8 Discussion and conclusion

8.1 Main findings

This paper has utilized a synthetic control method to evaluate the impact of 2013 Beijing region-specific air program on the PM2.5 pollution levels by comparing the pre-program concentrations to the post-program concentrations, while also accounting for the overall trend. Synthetic control approach has allowed this study to utilize a more similar control region for Beijing than a difference-in-differences analysis would, whereas the latter has provided numerical estimates for the significance of the policy effect.

Empirical estimates suggest that following the implementation of the air program, fine particulate matter pollution has experienced a noticeable decrease in Beijing, relative to the control regions, as Beijing had on average $4.67\,\mu g/m^3$ lower $PM_{2.5}$ levels. Overall, the pollution levels have dropped by 24.11%, but only one-fifth of this decrease can be owed to the regional policy. The difference-in-difference analysis provides a similar estimation, in addition to confirming the statistical significance of the policy effect. Generally, when the ongoing downward trend and the effect of a policy are jointly considered, the $PM_{2.5}$ levels could have surely dropped by proclaimed 39.6% in 2018, in addition to falling below the $60\,\mu g/m^3$ threshold. A look into the consumption of non-renewable energy products has revealed that Beijing underwent a transaction in its final energy composition by switching from carbon-rich coal and coke to less polluting natural gas.

This paper contributes to the ever-expanding archive of literature on Chinese air pollution, but the first one to utilize the synthetic control method to evaluate air program effectiveness. The main conclusion of this study is that the Jingjinji regional program did, in fact, significantly contribute to the improvement of Beijing air quality by effectively reducing the fine particulate matter concentrations.

8.2 Limitations and further research

The main limiting factor encountered by this research is the limited availability of data. Under ideal circumstances, the timeframe of the study should be expanded to include a greater number of observations as the length of the pre-intervention period positively impacts the integrity of the synthetic control estimation. Likewise, a restricted number of control regions has provided only a moderate synthetic unit approximation, which can be enhanced by a larger donor pool. Besides, an economic evaluation of the program which weights implementation and other various costs up against the monetary value of human lives saved is worth considering. In addition, assessment of policy effect based on the official government data could provide curious insights regarding various omitted variables. Government corruption is a major problem in a modern-day China and could be an influential factor.

9 Acknowledgements

I would like to express my sincere gratitude to my supervisor , Prof. Dr. Olivier Marie for his invaluable guidance and support throughout this study.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495–510.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Airborne Particles Expert Group. (1999). *Source apportionment of airborne particulate matter in the united kingdom*. Department of the Environment, Transport and the Regions.

Beckerman, B., Jerrett, M., Brook, J. R., Verma, D. K., Arain, M. A., & Finkelstein, M. M. (2008). Correlation of nitrogen dioxide with other traffic pollutants near a major expressway. *Atmospheric Environment*, *42*(2), 275–290.

Beijing Government. (2018). Index of air quality: Monitoring stations. Retrieved May 5, 2018, from http://zx.bjmemc.com.cn/getAqiList.shtml?timestamp=1525540153340

Cai, S., Wang, Y., Zhao, B., Wang, S., Chang, X., & Hao, J. (2017). The impact of the air pollution prevention and control action plan on pm2. 5 concentrations in jing-jin-ji region during 2012–2020. *Science of the Total Environment*, *580*, 197–209.

Card, D., & Krueger, A. B. (1993). *Minimum wages and employment: A case study of the fast food industry in new jersey and pennsylvania*. National Bureau of Economic Research.

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

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). Stl: A seasonal-trend decomposition. *Journal of Official Statistics*, *6*(1), 3–73.

Davidson, C. I., Phalen, R. F., & Solomon, P. A. (2005). Airborne particulate matter and human health: A review. *Aerosol Science and Technology*, *39*(8), 737–749.

De Kok, T. M., Driece, H. A., Hogervorst, J. G., & Briedé, J. J. (2006). Toxicological assessment of ambient and traffic-related particulate matter: A review of recent studies. *Mutation Research/Reviews in Mutation Research*, *613*(2), 103–122.

Divan, S., & Rosencranz, A. (2001). *Environmental law and policy in india: Cases, materials, and statutes*. Oxford University Press New Delhi.

Fann, N., Lamson, A. D., Anenberg, S. C., Wesson, K., Risley, D., & Hubbell, B. J. (2012). Estimating the national public health burden associated with exposure to ambient pm2. 5 and ozone. *Risk analysis*, *32*(1), 81–95.

Feng, Y., Chen, S., & Zhang, L. (2013). System dynamics modeling for urban energy consumption and co2 emissions: A case study of beijing, china. *Ecological Modelling*, *252*, 44–52.

GBD MAPS Working Group. (2016). Burden of disease attributable to coal-burning and other air pollution sources in China. Retrieved May 5, 2018, from https://www.healtheffects.org/publication/burden-disease-attributable-coal-burning-and-other-air-pollution-sources-china

Guan, D., Su, X., Zhang, Q., Peters, G. P., Liu, Z., Lei, Y., & He, K. (2014). The socioeconomic drivers of china primary pm2. 5 emissions. *Environmental Research Letters*, *9*(2), 024010.

Guangwei, H. (2014). The victims of China soil pollution crisis. Retrieved May 5, 2018, from https://www.chinadialogue.net/article/show/single/en/7073-Special-report-The-victims-of-China-s-soil-pollution-crisis

He, K., Yang, F., Ma, Y., Zhang, Q., Yao, X., Chan, C. K., ... Mulawa, P. (2001). The characteristics of pm2. 5 in beijing, china. *Atmospheric Environment*, *35*(29), 4959–4970.

Heinmueller, J. (2014). Package "synth". Retrieved July 5, 2018, from https://cran.r-project.org/web/packages/Synth/Synth.pdf

Huang, R. J., Zhang, Y. [Yanlin], Bozzetti, C., Ho, K.-F., Cao, J.-J., Han, Y., ... Canonaco, F., et al. (2014). High secondary aerosol contribution to particulate pollution during haze events in china. *Nature*, *514*(7521), 218.

Hueglin, C., Gehrig, R., Baltensperger, U., Gysel, M., Monn, C., & Vonmont, H. (2005). Chemical characterisation of pm2. 5, pm10 and coarse particles at urban, near-city and rural sites in switzerland. *Atmospheric Environment*, *39*(4), 637–651.

IMF. (2018). Macroeconomic and Financial Data. Retrieved May 5, 2018, from http://data.imf.org

Ito, K., Mathes, R., Ross, Z., Nádas, A., Thurston, G., & Matte, T. (2011). Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in new york city. *Environmental health perspectives*, *119*(4), 467.

Ji, D., Li, L., Wang, Y., Zhang, J., Cheng, M., Sun, Y., ... Hu, B., et al. (2014). The heaviest particulate air-pollution episodes occurred in northern china in january, 2013: Insights gained from observation. *Atmospheric Environment*, *92*, 546–556.

Liang, X., Zou, T., Guo, B., Li, S., Zhang, H., Zhang, S., ... Chen, S. X. (2015). Assessing beijing's pm2. 5 pollution: Severity, weather impact, apec and winter heating. *Proc. R. Soc. A*, *471*(2182), 20150257.

Liao, H., Chen, W.-T., & Seinfeld, J. H. (2006). Role of climate change in global predictions of future tropospheric ozone and aerosols. *Journal of Geophysical Research: Atmospheres*, *111*(D12).

Lim, L. (2013). Beijing airpocalypse spurs pollution controls, public pressure. Retrieved May 5, 2018, from https://www.npr.org/2013/01/14/169305324/beijings-air-quality-reaches-hazardous-levels

Liu, M., Huang, Y., Ma, Z., Jin, Z., Liu, X., Wang, H., ... Bi, J., et al. (2017). Spatial and temporal trends in the mortality burden of air pollution in china: 2004–2012. *Environment international*, *98*, 75–81.

Liu, Y., Franklin, M., Kahn, R., & Koutrakis, P. (2007). Using aerosol optical thickness to predict ground-level pm2. 5 concentrations in the st. louis area: A comparison between misr and modis. *Remote sensing of Environment*, *107*(1-2), 33–44.

Ministry of Ecology and Environment of China. (2013). Action Plan on Prevention and Control of Air Pollution introducing ten measures to improve air quality. Retrieved May 5, 2018, from http://english.mep.gov.cn/News_service/infocus/201309/t20130924_260707.htm

Pope, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the air & waste management association*, *56*(6), 709–742.

Pope, C. A., Ezzati, M., & Dockery, D. W. (2009). Fine-particulate air pollution and life expectancy in the united states. *New England Journal of Medicine*, *360*(4), 376–386.

Popp, D. (2003). Pollution control innovations and the clean air act of 1990. *Journal of Policy Analysis and Management*, *22*(4), 641–660.

Samet, J. M., Dominici, F., Curriero, F. C., Coursac, I., & Zeger, S. L. (2000). Fine particulate air pollution and mortality in 20 us cities, 1987–1994. *New England journal of medicine*, *343*(24), 1742–1749.

Selden, T. M., & Song, D. (1994). Environmental quality and development: Is there a kuznets curve for air pollution emissions? *Journal of Environmental Economics and management*, *27*(2), 147–162.

Sheng, L., Lu, K., Ma, X., Hu, J.-k., Song, Z.-x., Huang, S.-x., & Zhang, J.-p. (2015). The air quality of Beijing-Tianjin-Hebei regions around the Asia-Pacific economic cooperation (APEC) meetings. *Atmospheric Pollution Research*, *6*(6), 1066–1072.

Streets, D. G., Fu, J. S., Jang, C. J., Hao, J., He, K., Tang, X., ... Zhang, Q., et al. (2007). Air quality during the 2008 Beijing Olympic Games. *Atmospheric environment*, *41*(3), 480–492.

United States Embassy Beijing. (2014). Beijing air. Retrieved May 5, 2018, from https://twitter.com/BeijingAir

Venners, S. A., Wang, B., Xu, Z., Schlatter, Y., Wang, L., & Xu, X. (2003). Particulate matter, sulfur dioxide, and daily mortality in chongqing, china. *Environmental health perspectives*, *111*(4), 562.

Wang, L. [Li], Zhang, F., Pilot, E., Yu, J., Nie, C., Holdaway, J., ... Vardoulakis, S., et al. (2018). Taking action on air pollution control in the beijing-tianjin-hebei (bth) region: Progress, challenges and opportunities. *International journal of environmental research and public health*, *15*(2), 306.

WHO. (2016). Monitoring health for the SDGs. Retrieved May 5, 2018, from http://www.who.int/gho/publications/world_health_statistics/2016/en/

World Bank. (2018). Carbon dioxide emissions per country. Retrieved May 5, 2018, from http://databank.worldbank.org/data/reports.aspx?source=2&series=EN.ATM.CO2E.PC&country

World Bank & IHME. (2016). The cost of air pollution: Strengthening the economic case for action. Retrieved May 5, 2018, from http://documents.worldbank.org/curated/en/781521473177013155/The-cost-of-air-pollution-strengthening-the-economic-case-for-action

Youbin, L. (2018). China achieves desired results in Clean Air Action Plan: official. Retrieved May 5, 2018, from http://english.sepa.gov.cn/News_service/media_news/201802/t20180201_430711.shtml

Zhang, Wang, S., Hao, J., Wang, X., Wang, S., Chai, F., & Li, M. (2016). Air pollution and control action in Beijing. *Journal of Cleaner Production*, *112*, 1519–1527.

Appendix 1: Synthetic control method implementation

This appendix elaborates upon the implementation of synthetic control method using *synth* package in *R*. First, let U_i be a $(r \times 1)$ vector of observed covariates for each unit and let $K = (k_1, \ldots, k_{T_0})'$ be a $(T_0 \times 1)$ vector that denotes a linear combination of pre-intervention outcomes $\overline{Y}_i^K = \sum_{s=1}^{T_0} k_s Y_{is}$, which is used to to control for unobserved common factors with time-varying effects. To implement the synthetic control method numerically, difference between the treated unit and its synthetic counterpart is defined by combining characteristics of the exposed unit in a $(k \times 1)$ vector $X_1 = \left(U_1', \overline{Y}_1^{K_1}, \ldots, \overline{Y}_1^{K_M}\right)'$ and the absolute distance between X_1 and X_0W subject to a weight constraint. Specifically, *synth* solves for an optimal vector of weights W^* that minimizes

$$||X_1 - X_0W||v = \sqrt{(X_1 - X_0W)'\, V\, (X_1 - X_0W)}$$

where V is a $(k \times k)$ symmetric semidefinite matrix which allows for different weights of the variables in X_0 and X_1 depending on their power of predicting the outcome. While *synth* permits manual choice of weights for the predictor variables, this paper employs a data-driven procedure, similarly to Abadie and Gardeazabal (2003), which generates V^* such that the mean squared prediction error (MSPE) of the synthetic control estimator is minimized over the pre-intervention period.

Appendix 2: Difference-in-differences estimation

This appendix outlines the intutition behind the difference-in-differences estimation. To set the notation, let $I = 0, 1$ indicated the treatment status, where 0 refers to the control group of Chinese regions unaffected by the intervention, and 1 refers to Beijing. Particulate matter concentrations are assumed to be observed during two periods $t = 0, 1$ where 0 indicates the time period prior to the implementation of the air program, and 1 the period after it has been implemented. Observed $PM_{2.5}$values are indexed by $i = 1, \ldots, N$. Lastly, Beijing pre- and post-intervention average particulate matter concentrations are denoted as Y_0^I and Y_1^I, respectively. Similarly, Y_0^C and Y_1^C refer to pre- and post-intervention averages of the control regions. These outcomes are given by the following equation:

$$Y_i = \alpha + \beta I_i + \gamma t_i + \delta(I_i t_i) + \varepsilon_i$$

Here, α is the constant term, β is the effect specific to the treated unit, γ is the common time trend, δ is the true effect of the intervention, and ε is the error term. For δ to provide an unbiased estimation, ε needs to have an average expected value of zero and to be uncorrelated with the other variables in the equation. The latter is known as the parallel-trend assumption and is crucial for the DiD estimation. Consider $cov(\varepsilon_i, I_i t_t) = E(cov(\varepsilon_i, I_i t_t)) = \varphi$. If Y_i follows a different trend for the treatment and the control group, respectively $\gamma^I = \gamma + \varphi$, and $\gamma^C = \gamma$, the estimated intervention effect will be biased:

$$E(\hat{\delta}_{DiD}) = (\gamma^I + \delta) - \gamma^C = \gamma + \varphi + \beta - \gamma = \delta + \varphi$$

A DiD estimator takes the difference between a difference in average outcome in the treatment and controls groups prior- and post-intervention:

$$\hat{\delta}_{DiD} = (Y_1^I - Y_0^I) - (Y_1^C - Y_0^C)$$

Consequently, $\hat{\delta}_{DiD}$ can be rewritten in terms of expected values:

$$\begin{aligned}
\hat{\delta}_{DiD} &= (E(Y_1^I) - E(Y_0^I)) - (E(Y_1^C) - E(Y_0^C)) \\
\hat{\delta}_{DiD} &= (\alpha + \beta + \gamma + \delta) - (\alpha + \beta) - (\alpha + \gamma) - \gamma \\
\hat{\delta}_{DiD} &= (\gamma + \delta) - \gamma \\
\hat{\delta}_{DiD} &= \delta
\end{aligned}$$