Environmental protection inspection and air pollution in Hebei: Catch me if you can

Pan Chen

School of Economics, HUST

2018.3.27

Introduction

.

- China's economy has reached a high level
- However, at a huge cost of environment
- haze has become a regular customer of China, and half more of the land is capped in these particular matters.

- Air pollution has created a heavy burden to the country, and its negative effect has been studied by various papers (e.g., Word Bank, 2007; Matus et al., 2012; He et al., 2016).
- Note: this paper does not strictly distinct haze from air pollution since in China, haze conquers all other pollutants in air

- Stronger economic power wins China more opportunity in holding international meetings and games.
- leave good impression on foreigner participants, push a series of regulations aiming at reducing air pollution
- for example, the 2008 Beijing Olympic Games, APEC meeting
- effect cannot last for too long, but works immediately(Chen et al., 2013).

- Also during big domestic events, e.g., victory of China's resistance war
- curb the air pollution to make it a decent day
- 'Olympic blue', 'APEC blue' or 'military-parade blue'

- not just a matter of grey or blue sky, but concerning people's health.
- significant associations between air pollution exposure and increased mortality risks (Zhang et al., 2011; Shang et al., 2013; Zhang et al., 2014; He et al., 2016).
- Tanaka (2015), regulations on air pollution in China, infant mortality rate falls by 20 percent.

- CPC's principle, no permission
- the central government orders local governments to solve environmental problems, of which air pollution is at head.

- puzzle: environmental damage, economic growth and officers' promotion
- there will be a big cost: low-level economic growth for governments and unaffordable expenditure for factories (many are in small size)
- no incentives to follow the orders

- local governments' laziness has nowhere to hide
- how to meet the goals of pollution control and economic growth?
- manipulating statistics, Merliand and Raftery (2000); Ghanem and Zhang (2014)
- the crowd's calling for clean air grows and the air becomes worse, central government realize they must find a way out

- A new law took effect at the beginning of 2015, but no significant change
- environmental protection inspection project
- experiment in Hebei province, a (not one of) most serious haze-attacked region in China, in the first two months of 2016

- a panel data method initiated by Hsiao et al. (hereafter HCW)(2012) to evaluate the impact on air quality of 8 serious haze-capped cities of Hebei province.
- different from DID or synthetic control method
- (1) it allows for sample selection effect; (2) it allows for the impact of underlying factors to vary by cross-sectional units; (3) estimation of it is computationally much easier, and OLS is enough. (Bai et al., 2014)

- The inspection lowers 17 weeks' average AQI in Shijiazhuang (capital of Hebei province), Tangshan, Handan, Baoding, Langfang, Cangzhou, Hengshui, and Xingtai by 13.13 (8.93%),9.67 (7.6%), 9.6 (12.36%), 38 (16.25%), 39.73 (29.98%), 18.77 (16.2%), 4.73 (4.14%), and 17.65 (13.33%) respectively.
- longer post-treatment period, lower accuracy
- several robust checks and the conclusions do not change.

- The paper continues as follows:
- Section "Background"
- Section"The model and Lasso method"
- Section "The data"
- Section "Estimation of the treatment effect"
- Section "Robustness checks"
- Section "Conclusion"

Background

Hebei province and its haze

.

Figure 1: Location of Hebei

Figure 2: Constitutions of Hebei

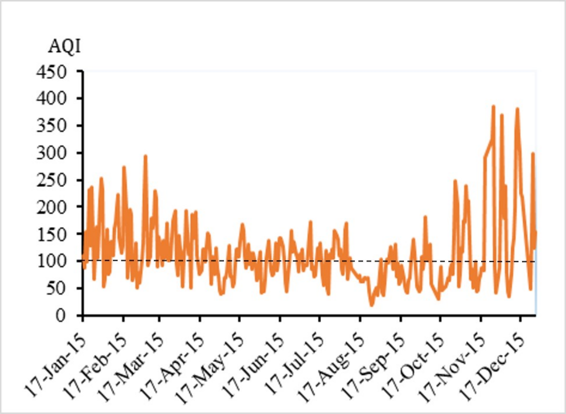

Figure 3: Daily AQI in Shijiazhuang city in 2015

- lightly polluted or worse (AQI >100) days is up to 173, 53% of the sample days.
- moderately polluted and worse (AQI>150) days, 21%
- all the 8 haze-plagued cities are in southern Hebei

- meteorological factors (Wei et al., 2010);
- approximately 65% of the PM2.5 in Shijiazhuang and Xingtai originated from the local southern area, and the rest from the northern area and the nearby Shanxi province.

Background

Environmental protection inspection

.

- CPC Central Committee and the State Council decide to constitute a group
- inspection items include air pollution, water pollution, noise and any other environment-related problems.
- equipped with a retired minister or equal level officer as leader, and the deputy leader should be a current vice minister of the Ministry of Environmental Protection.

- January 4th, 2016, the environmental inspection group formed up in Shijiazhuang to inspect all cities of Hebei province in the following two months.
- feedback meeting on May 2016, banned 200 factories, punished 123 people
- finishing the inspection at February 4th, submitted materials to central government then, and later at May 3rd gave feedbacks to Hebei government.
- on-site inspection before May 3rd and remote supervision after May 3rd.

The model and Lasso method

HCW approach

.

- refer to correlations among cross-sectional units
- result from some common factors which affect all cross-sectional units, impacts are allowed to vary
- DID supposes each cross-unit of treatment and control groups bears the same influence, which can hardly apply to China (Bai et al., 2014)

- Follow HCW, we consider the general model as

$$y_{it}^{0} = a_i + b_i^{'} f_t + \mu_{it}, \tag{1}$$

where i=1,...,N; t=1,...T, a_i is the fixed individual-specific effects, b_i is a $K \times 1$ vector of (unobserved) factor loadings, f_t is a $K \times 1$ vector of (unobserved) factors, and μ_{it} is a weakly dependent and weakly stationary error term with zero mean. Interpretation of factors and factor loadings can be seen in Bai (2009).

consider a case that for $t = 1, ..., T_1$, no policy intervention for all i. Suddenly at T_{1+1} certain i receives a treatment and without loss of generality, we let this happen to the first unit. Those without treatment denoted by y_{jt}, $j = 2, ..., N$. Time line can be divided into two parts, the pretreatment period and the post-treatment period. For all units in the pretreatment period, we have

$$y_{it}^0 = a_i + b_i^{'} f_t + \mu_{it}, i = 1, ..., N; t = 1, ... T_1, \quad (2)$$

where y^0 denotes the (observed) outcome without any intervention.

Based on Eq.2, by stacking $i = 1, , N$ of pretreatment period we can obtain the following form

$$Y_t = Y_t^0 = a + Bf_t + \mu_t, t = 1, ..., T_1 \tag{3}$$

where $Y_t(y_{1t}, ..., y_{Nt})^{'}$ and $a = (a_1, ..., a_N)^{'}$ are both $N \times 1$ vectors, $B = (b_1, ..., b_2)^{'}$ is an $N \times K$ factor loading matrix, and $\mu_t = (\mu_{1t}, ..., \mu_{Nt})^{'}$ is an $N \times 1$ vector of error terms.

At time T_1+1, the first unit receives a policy intervention while the rest units keep unaffected. Hence, for the first unit from time T_1+1 on, y_{1t} takes the form

$$y_{1t} = y_{1t}^1 = a_1 + b_1^{'} f_t + \Delta_{1t} + \mu_{1t}, t = T_1, ..., T \tag{4}$$

where y_{1t}^1 is the observed post-treatment outcome of the first unit. Δ_{1t} is the treatment effect to the first unit at time $t = T_1 + 1$.

For units without any intervention at time T_1+1, we have

$$y_{it}^0 = a_i + b_i^{'} f_t \mu_{it}, i = 2, ...N; t = T_1, ..., T \quad (5)$$

Note that in this period ($t = T_1 + 1, ..., T$), all y_{it}^0 are observable for $i = 2, , N$. Ideally if y_{1t}^0 for $t \geqslant T_1 + 1$ can be observed, the treatment effect, $\Delta_{1t} = y_{1t}^1 - y_{1t}^0$, is of no difficulty to derive. Unluckily, y_{1t}^0 cannot directly be observed

- construct the counterfactual y_{1t}^{0} for $\geqslant t_1 + 1$.
  If N and T_1 are both large, the procedure developed by Bai and Ng (2002) can be employed to identify f_t and b_1 in Eq.1, $y_{it}^{0} = a_i + b_i^{'} f_t + \mu_{it}$, by the maximum likelihood approach, and then predict y_{1t}^{0}. However, when N or T_1 or both are not sufficiently large, may not be estimated accurately.(Bai and Ng, 2002). Hsiao et al. (2012) propose a novel way by using $\widetilde{Y}_t = (y_{2t}, ..., y_{NT})^{'}$ in lieu of f_t to predict y_{1t}^{0} for post-treatment period.

- Bai et al. (2014) and Li and Bell (2017) prove that when the series are non-stationary I(1) processes or trend-stationary, the OLS estimators are still consistent.

We define $\widetilde{Y}_t = (y_{2t}, ..., y_{NT})'$ so that $Y_t = (y_{1t}, \widetilde{Y}_t')'$, where $t \leqslant T_1$. Through analogy definitions we have $\widetilde{a}$ and $a = (a_1, \widetilde{a}')'$, $\widetilde{\mu}_t$ and $\mu_t = (\mu_{1t}, \widetilde{\mu}_t')'$ respectively. Going after Li and Bell (2017) we get

Assumption 1. (i) $||b_i|| \leqslant M < \infty$ for all $i = 1, ...N$, where M is a positive constant; (ii) μ_t is a weakly dependent process with $E(\mu_t) = 0$ and $E(\mu_t \mu_t') = V$, where V is an $N \times N$ diagonal matrix; (iii) $E(\mu_t f_t') = 0$ for all t; (iv) $E(\mu_{jt|d_{1t}}) = 0$ for $j \neq 1$, where d_{it}=1 if the ith unit is under treatment at time t and $d_{it} = 0$ otherwise.

Assumption 2. (i) Let $x_t = (1, \widetilde{Y}_t)$. Then $\{x_t\}_{t=1}^{T}$ is a weakly dependent and weakly stationary process, $\frac{1}{T_1} \sum_{t=1}^{T_1} x_t x_t' \xrightarrow{p} E(x_t x_t')$ as $T_1 \to \infty$, and $[E(x_t x_t')]$ is invertible. (ii)$Rank(\widetilde{B}) = K$, where definitions for $\widetilde{Y}_t$ and $\widetilde{a}$ apply for $\widetilde{B}$.

By definition, since $\widetilde{Y}_t, \widetilde{B}, \widetilde{a}, \widetilde{\mu}_t$ are vectors or matrix obtained by removing the first rows of Y_t, B, a and μ_t respectively, we can easily have $\widetilde{Y}_t = \widetilde{a} + \widetilde{B} f_t + \widetilde{\mu}_t$
, Multiplying it by $\widetilde{B}'$ on both sides and then finding solution for f_t, under Assumption 2(ii) we derive

$$f_t = (\widetilde{B}' \widetilde{B})^{-1} \widetilde{B}' (\widetilde{Y}_t - \widetilde{a} - \widetilde{\mu}_t) \tag{6}$$

Substituting Eq.6 into Eq.2 for $i = 1$, it produces

$$y_{1t}^0 = a_1 + b_1' (\widetilde{B}' \widetilde{B})^{-1} \widetilde{B}' (\widetilde{Y}_t - \widetilde{a} - \widetilde{\mu}_t) + \mu_{1t} \tag{7}$$

After some arrangement of terms we obtain

$$y_{1t}^{0} = \gamma_1 + \gamma_1^{'} \widetilde{Y}_t + \epsilon_{1t} \tag{8}$$

where $\gamma_1 = a_1 - b_1^{'}(\widetilde{B}^{'}\widetilde{B})^{-1}\widetilde{B}^{'}\widetilde{a}$, $\gamma_1^{'} = b_1^{'}(\widetilde{B}^{'}\widetilde{B})^{-1}\widetilde{B}^{'}$, $\epsilon_{1t} = \mu_{1t} - b_1^{'}(\widetilde{B}^{'}\widetilde{B})^{-1}\widetilde{B}^{'}\widetilde{\mu}_t$. Li and Bell (2017) then assume (Y_t, ϵ_{1t}) to be a weakly stationary process and define c_1 and c to be the minimizers of $\min_{c_1,c}[(\epsilon_{1t} - c_1 - c^{'}\widetilde{Y}_t)^2]$,where $c = (c_2, ..., c_N)^{'}$. Term $c_1 + c^{'}\widetilde{Y}_t$ is then called the linear projection of ϵ_{1t} onto $(1, \widetilde{Y}_t^{'})$ and denoted by $L(\epsilon_1 t|\widetilde{Y}_t)$. Therefore, ϵ_{1t} can be decomposed into $\epsilon_{1t} = L(\epsilon_{1t}|\widetilde{Y}_t) + \upsilon_{1t}$, where $\upsilon_{1t} = \epsilon_{1t} - L(\epsilon_{1t}|\widetilde{Y}_t)$. We can re-write Eq.8 as

$$
\begin{aligned}
y_{1t}^{0} &= \gamma_1 + \gamma_1^{'} \widetilde{Y}_t + L(\epsilon_{1t}|\widetilde{Y}_t) + \upsilon_{1t} \\
&= \gamma_1 + \gamma_1^{'} \widetilde{Y}_t + c_1 + c^{'} \widetilde{Y}_t^{2} + \upsilon_{1t} \\
&= \delta_1 + \delta^{'} \widetilde{Y}_t + \upsilon_{1t}
\end{aligned} \tag{9}
$$

Where $\delta_1 = \gamma_1 + c_1$ and $\delta = \gamma + c$. Because $L(\epsilon_{1t}|\widetilde{Y}_t) = 0$, which implies $E(\upsilon_{1t}) = 0$ and $E(\widetilde{Y}_t \upsilon_{1t}) = 0$, Hsiao et al. (2012) and Li and Bell (2017) prove that using pretreatment data, OLS regression of y_{it}^{0} (for $t \leqslant T_1$) on $(1, \widetilde{Y}_t^{'})$ gives consistent estimators for δ_1 and δ. Even if the series are non-stationary, Bai et al. (2014) still validate OLS estimators' consistency.

The model and Lasso method

Brief introduction of Lasso

.

- Consider a linear regression

$$y_{it}^{0} = x_{t}^{'}\beta + \upsilon_{1t}, t = 1, ..., T_1 \quad (10)$$

where definitions for y_{it}^{0}, $x_{t}^{'}$ inherit from the former part, $_beta = (\delta_1, \delta)^{'}$ is an N-vextor. Note that Lasso allows for the number of predictors N to be bigger than T_1, which is infeasible with OLS. For a given value of $\lambda \geqslant 0$, Lasso selects β to minimize

$$\sum_{t=1}^{T_1}[y_{1t}^{0} - x_{t}^{'}\beta]^2 + \lambda\sum_{j=2}^{N}|\delta_j| \quad (11)$$

Apart from the term with λ, the sum square term in Eq.11 is the same as that in ordinary linear regression.

- we may search for λ over a discrete set $\Lambda_m = \{0, \lambda_1, \lambda_2, ..., \lambda_m\}$, where λ_m is larger than all other elements.
- 5 fold cross-validation
- General process:
  1, calculates the largest value of λ (λ_m) that gives a non-null model
  2, searches for λ in Λ_m successively, saves the cross-validation error at the same time
  3, pin down the λ with minimum cross-validation error

The data

.

- The Air Quality Index (AQI): refers to six kind of pollutants, including PM2.5, PM10, SO_2, NO_2, O_3, and CO.
- data before 2014 just contains a few cities. not until May 15th of the year that AQI started to be reported regularly. Even in this way, data of many cities are still absent.
- we start the collection of the hourly data from the first day of 2015, end it at the last day of December 2016.
- Data after December 2016 is available

- problem with hourly data, how about daily data?
- monthly versus weekly
- date set finally contains 357 cities, 346=357−11

- Note again that HCW still works if the series is non-stationary I(1) process.
- augmented Dicky-Fuller tests in log first differences (with trend) on 357 cities' AQI, Mackinnon approximated *p*-values are all significant at the 10% level or above

Estimating the treatment effect

.

- normally the pretreatment period should be 46 weeks. However, we cut off the last two weeks, so the formal pretreatment period contains 44 weeks.
- We take the 44 weeks' AQI data into the following model

$$q_{1t}^{0} = \gamma_1 + \gamma^{'} \widetilde{Q}_t + \epsilon_{1t} \tag{12}$$

where q_{1t}^{0} is, for instance, the AQI in Shijiazhuang city in pretreatment period. $\widetilde{Q}_t$ are AQI in control cities.

Results of Shijiazhuang and Tangshan

.

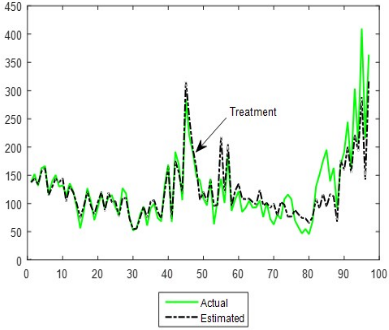

Figure 5: Treatment effect for Shijiazhuang

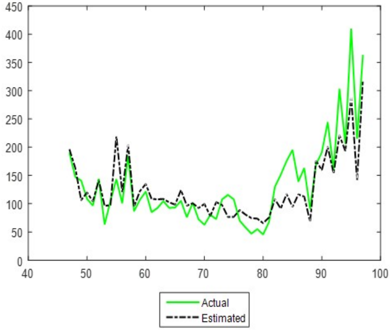

Figure 6: Signified treatment effect for Shijiazhuang

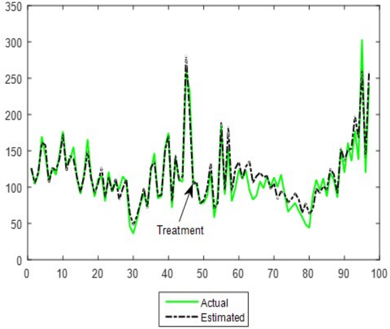

Figure 7: Treatment effect for Tangshan

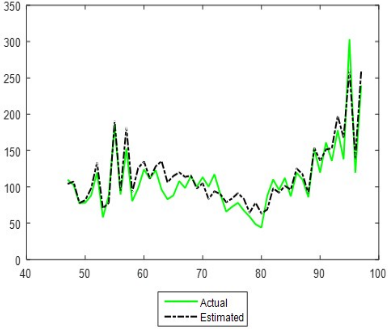

Figure 8: Signified treatment effect for Tangshan

- treatment effect equals $q^1_{1t} - \widehat{q}^0_{1t}$, ATE equals $q^1_{1t} - \widehat{q}^0_{1t} / T_2$

Table 1: The Average treatment effects (ATEs) for Shijiazhuang and Tangshan

	Shijiazhuang	Tangshan
17 weeks	-13.13^{**}	-9.67^{***}
	(−2.28)	(−2.76)
34 weeks	-11.41^{***}	-8.82^{**}
	(−2.59)	(−2.22)
51 weeks	6.82	-6.97^{**}
	(0.74)	(−2.32)

Parentheses present the T-statistic calculated by means of Theorem 3.2 of Li and Bell (2017); *,**, and *** imply significance at 10%, 5%, and 1% respectively. T-statistics in this paper is based on Lasso method, while Li and Bell (2017) derive it with least squares estimator.

$$T.S. \stackrel{def}{=} \sqrt{T_2}\frac{\widehat{\Delta}_1}{\sqrt{\widehat{\Sigma}}} = \frac{\widehat{\Delta}_1}{\sqrt{\widehat{\Sigma}/T_2}} \xrightarrow{H_0} N(0,1)$$

where $\widehat{\Sigma} = \widehat{\Sigma}_1 + \widehat{\Sigma}_2$
$\widehat{\Sigma}_1/T_2 = (T_2/T_1)\widehat{E}(x_t)'(V'/T_2)\widehat{E}(x_t)$, V' is a consistent estimator of $Var(\sqrt{T_1}\widehat{\beta})$;
$\widehat{\Sigma}_2 = \frac{1}{T_2}\sum_{t=T_1+1}^{T}\sum_{s=T_1+1,|t-s|\leqslant l}^{T}[\widehat{\Delta}_{1t} - \widehat{\Delta}_1][\widehat{\Delta}_{1s} - \widehat{\Delta}_1]$
where $\widehat{\Delta}_1 = T_2^{-1}\sum_{s=T_1+1}^{T}\Delta_{1s}$, $l \longrightarrow \infty$ and $l/T_2 \longrightarrow 0$ as $T_2 \longrightarrow \infty$. For example, one may choose $l = O(T_2^{1/4})$ (Newwy and West, 1987a,b; White, 2000, 2001) or use a faster rate for l (Andrews, 1991).

OLS estimate: $Var(\widehat{\beta}) = \delta^2(X'X)^{-1}$, $\widehat{\delta}^2 = \frac{e^{'}e}{n-k-1}$

Lasso estimate is a non-linear and non-dfferentiable function of response values even for a fixed value of λ.The covariance matrix of the estimates may then be approximated by:

$(X^{'}X + \lambda W^{-})^{-1}X^{'}X(X^{'}X + \lambda W^{-})^{-1}\widehat{\delta}^2$ where W is a diagonal matrix with diagnoal elements $|\widetilde{\beta}_j|$ and W^{-} denotes the generalized inverse W (Tibshirani, 1996).

Results of Handan and Baoding

.

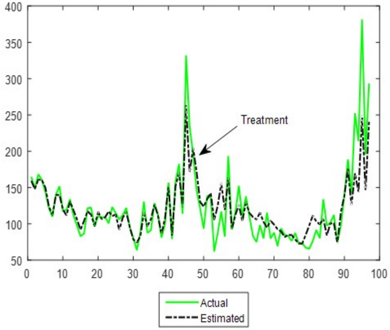

Figure 9: Treatment effect for Handan

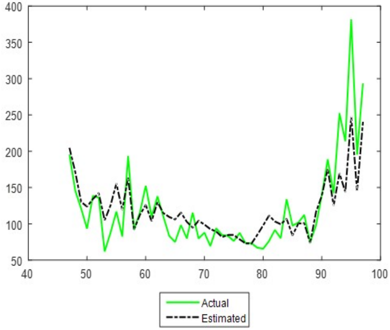

Figure 10: Signified treatment effect for Handan

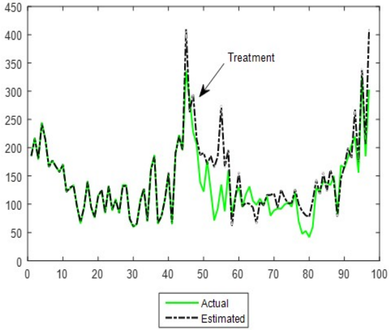

Figure 11: Treatment effect for Baoding

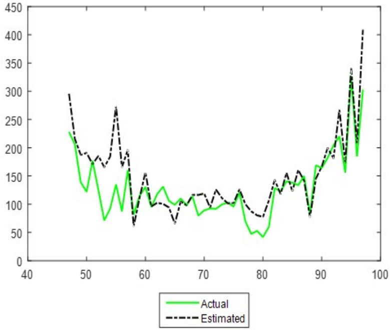

Figure 12: Signified treatment effect for Baoding

Table 2: ATEs for Handan and Baoding

	Handan	Baoding
17 weeks	−9.60	−38.00**
	(−1.21)	(−2.08)
34 weeks	−10.04**	−25.18**
	(−2.16)	(−2.16)
51 weeks	0.96	−21.66***
	(0.11)	(−2.54)

Results of Langfang and Cangzhou

.

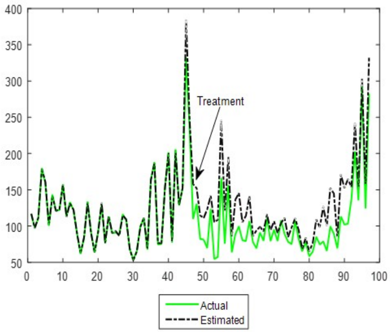

Figure 13: Treatment effect for Langfang

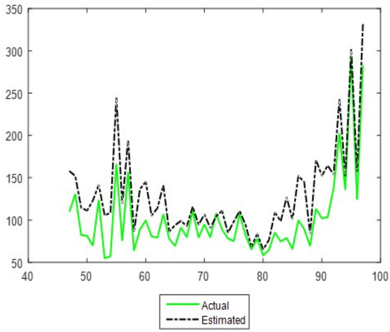

Figure 14: Signified treatment effect for Langfang

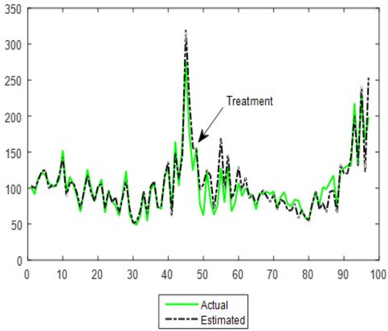

Figure 15: Treatment effect for Cangzhou

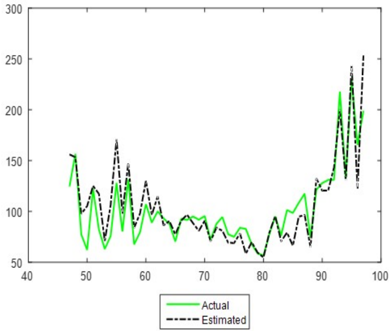

Figure 16: Signified treatment effect for Cangzhou

Table 3: ATEs for Langfang and Cangzhou

	Langfang	Cangzhou
17 weeks	-39.73^{***}	-18.77^{***}
	(-9.99)	(-6.29)
34 weeks	-25.09^{***}	-7.29
	(-4.31)	(-1.48)
51 weeks	-28.45^{***}	-3.05
	(-6.27)	(-0.73)

Results of Hengshui and Xingtai

.

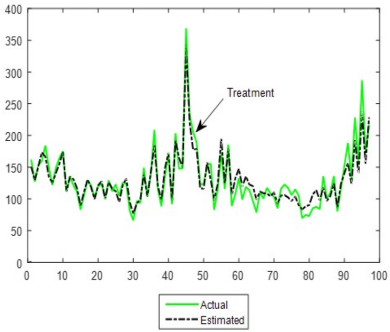

Figure 17: Treatment effect for Hengshui

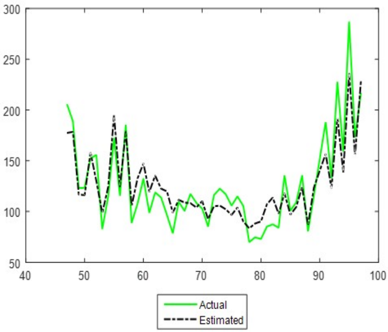

Figure 18: Signified treatment effect for Hengshui

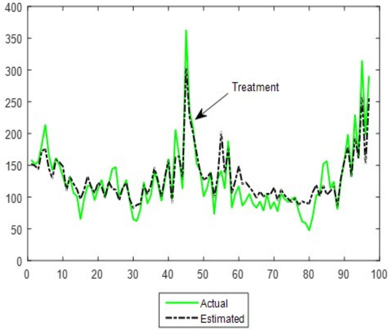

Figure 19: Treatment effect for Xingtai

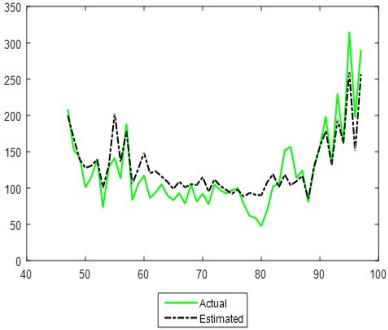

Figure 20: Signified treatment effect for Xingtai

Table 4: ATEs for Hengshui and Xingtai

	Hengshui	Xingtai
17 weeks	−4.73	−17.65***
	(−0.81)	(−4.09)
34 weeks	−2.99	−16.58**
	(−0.71)	(−4.21)
51 weeks	0.47	−6.49
	(0.11)	(−1.03)

Summary

.

Table 5: Classification of ATEs for cities

	High	Moderate	Slight or None
17 weeks	Baoding Langfang	Shijiazhuang Cangzhou Xingtai	Handan Tangshan Hengshui
34 weeks	Baoding Langfang	Shijiazhuang Xingtai	Tangshan Handan Cangzhou Hengshui
51 weeks	Baoding Langfang		Shijiazhuang Tangshan Handan Cangzhou Hengshui Xingtai

- the following two points in these cases deserve a discussion:
  First, generally declining trend in treatment effect
  Second, Shijiazhuang and Handan stand out as a consequence of a huge gap between actual and estimated AQI at the end of post-treatment period.

Robustness Check

Out-of-sample prediction

.

- Theoretically, one can improve the accuracy of in-sample prediction to 100%.
- cut 10% of the net pretreatment period data

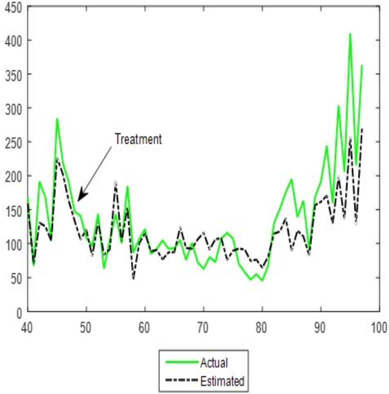

Figure 21: Out-of-sample fit: Shijiazhuang

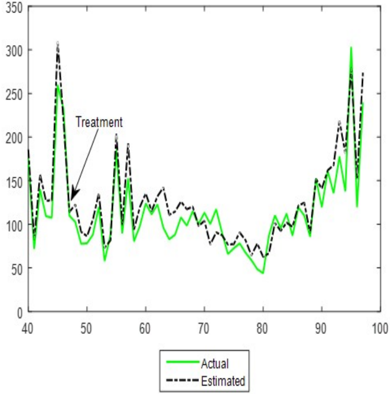

Figure 22: Out-of-sample fit: Tangshan

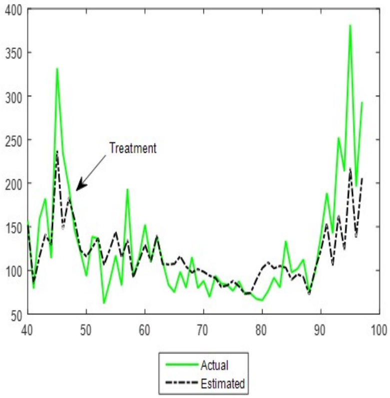

Figure 23: Out-of-sample fit: Handan

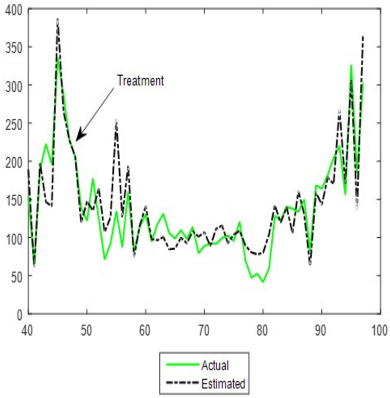

Figure 24: Out-of-sample fit: Baoding

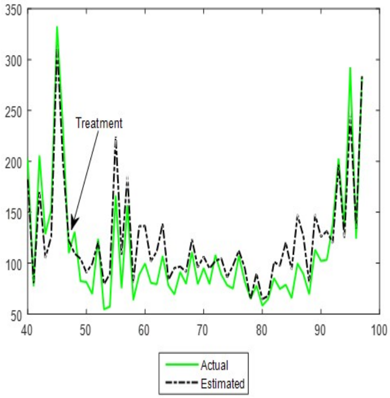

Figure 25: Out-of-sample fit: Langfang

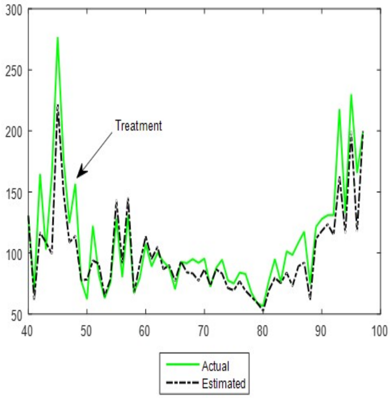

Figure 26: Out-of-sample fit: Cangzhou

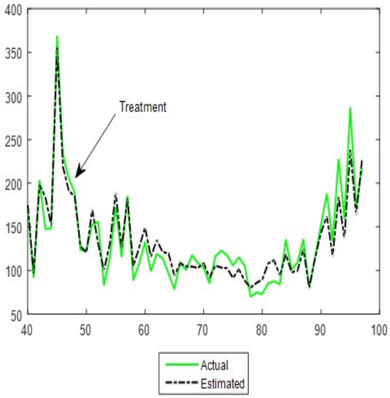

Figure 27: Out-of-sample fit: Hengshui

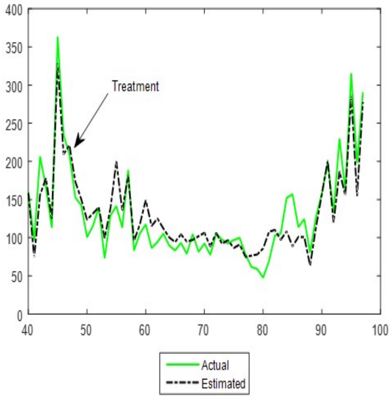

Figure 28: Out-of-sample fit: Xingtai

Robustness Check
Replacing AQI with PM2.5

.

- Obtaining AQI relates to several steps, and an intermediate step requires calculating an index of each pollutant
- index of PM2.5 prevails over other pollutants in most regions of China

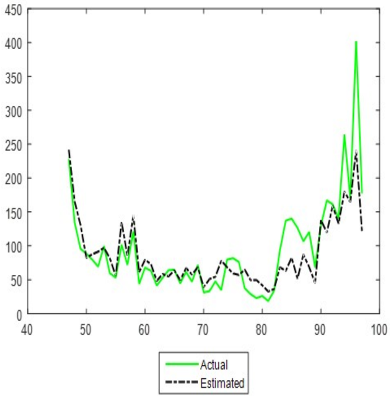

Figure 29: Estimation using PM2.5: Shijiazhuang

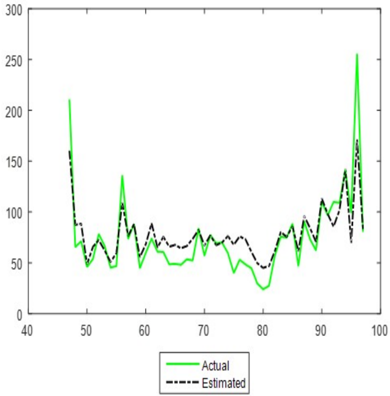

Figure 30: Estimation using PM2.5: Tangshan

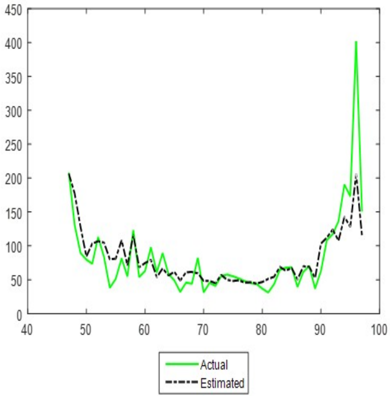

Figure 31: Estimation using PM2.5: Handan

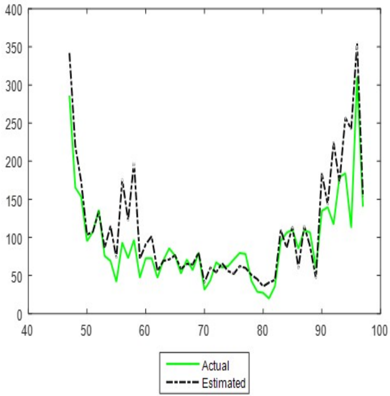

Figure 32: Estimation using PM2.5: Baoding

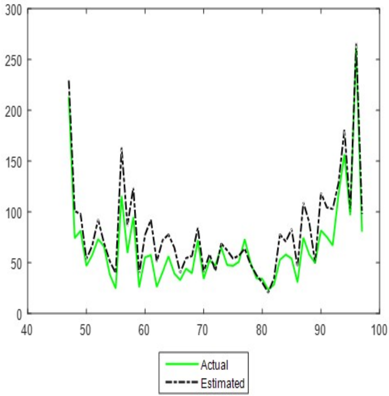

Figure 33: Estimation using PM2.5: Langfang

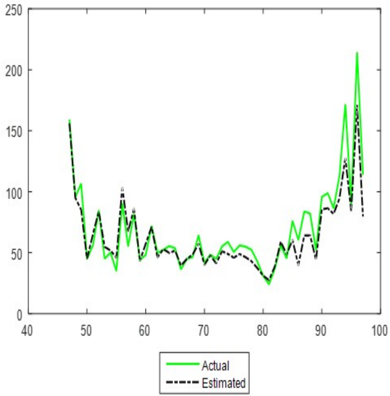

Figure 34: Estimation using PM2.5: Cangzhou

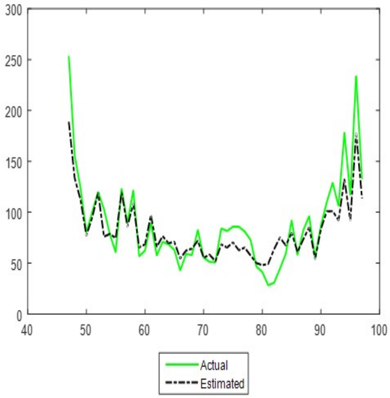

Figure 35: Estimation using PM2.5: Hengshui

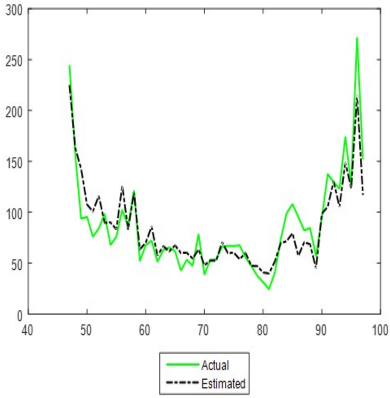

Figure 36: Estimation using PM2.5: Xingtai

The exogeneity of control units

.

- HCW methodology depends on the exogeneity of control variables
- our study faces a distinct situation.
- hardly possible that polluting sources operate with several branches and have made full preparation for a cat-mouse game
- possibility of successfully tracing the air pollutant in other province is far lower than tracing water and residue pollution
- we conservatively pay attention to a relatively short period amounting to 17 weeks so as to prevent the treatment from spillover

- Lastly but most importantly, serious haze pollution has a feature in location
- Then the location or geography can be regarded as a common factor to some extent
- Actually, control variables selected by Lasso (see Table 6) stand by our side.

Table 6: Number of selected units in neighborhood

	Number of selected control units	Number of selected units in neighborhood
Shijiazhuang	24	11
Tangshan	17	8
Handan	1 6	9
Baoding	42	11
Langfang	34	13
Cangzhou	17	11
Hengshui	6	3
Xingtai	5	5

- Several months later, shortening the duration from the previous 2 months to 1 month, the first round formal inspection from central government boots up
- run the Lasso regression with pretreatment period amounting to 44 weeks once more

Table 7: ATEs for the 8 cities after kicking-out

	17 weeks	34 weeks	51 weeks
Shijiazhuang	−15.71***	−12.88***	2.18
	(−3.34)	(−3.10)	(0.28)
Tangshan	−10.67***	−10.57***	−8.19***
	(−3.58)	(−2.83)	(−2.76)
Handan	−13.00	−15.76***	−3.91
	(−1.57)	(−3.39)	(−0.43)
Baoding	−30.85*	−18.04	−12.15
	(−1.68)	(−1.53)	(−1.40)

	17 weeks	34 weeks	51 weeks
Langfang	-30.26^{*}	-17.01^{***}	-16.48^{***}
	(-8.79)	(-3.08)	(-3.72)
Cangzhou	-12.05^{***}	-6.07^{*}	-0.69
	(-3.01)	(-1.63)	(-0.18)
Hengshui	-4.06	-2.97	0.51
	(-0.69)	(-0.71)	(0.12)
Xingtai	-15.18^{***}	-16.52^{***}	-5.22
	(-3.39)	(-4.05)	(-0.77)

Conclusion

.

- at 17 weeks' average, the inspection decreases AQI in Shijiazhuang (capital of Hebei province), Tangshan, Handan, Baoding, Langfang, Cangzhou, Hengshui, and Xingtai by 13.13 (8.93%), 9.67 (7.6%), 9.6 (12.36%), 38 (16.25%), 39.73 (29.98%), 18.77 (16.2%), 4.73 (4.14%), and 17.65 (13.33%) respectively.
- The three weeks-average treatment effects indicate the effect of environmental inspection generally declines in most treated cities. What's more, average treatment effects in few cities even vanish.
- Mixed with economy and politics, air pollution in China is almost impossible to be wiped out by a single try
- Further study, Elastic-net, when two units are correlated, Lasso tends to select one of them randomly, while Elastic-net will pick up both.

Thank You!

Email: lantocp@163.com