Empirical research on influential factors of air pollution in Chinese provincial capital cities

Chenwei Ma1 and Lijin Ding1*

1 School of Public Administration, Sichuan University, Chengdu, Sichuan, 610065, China
*Corresponding author’s e-mail: 2016141094163@stu.scu.edu.cn

Abstract. With the rapid development of China's economy, air pollution has gradually drawn experts and governors' attention. To investigate possible socioeconomic influential factors and distribution traits of air pollution in China, this paper adopts the static panel data regression and cluster analysis on the panel data of the air quality index of 30 provincial capital cities. The results of the research indicate that Chinese provincial capital cities can be divided into three categories based on the degree of contamination of the six typical pollutants. The research also find that industrial structure, green energy and the number of cars can significantly affect air pollution in capital cities, but the population density does not have a significant effect.

1. Introduction
According to the United Nations, more than half of the world’s population lives in cities, and this proportion is expected to rise to 60% by 2030. Urban areas account for about 70% of global carbon emissions and over 60% of resource use. The accompanying high concentration of production and living activities has brought tremendous pressure on the living environment and public health of urban residents. Taking air pollution as an example, as of 2016, about 90% of urban residents breathed air that did not meet safety standards, and 4.2 million people died from air pollution problem. In China, the problem of air pollution is also not optimistic. According to the "2018 China Eco-Environmental Status Bulletin" released by the Ministry of Ecology and Environment of China, among 169 key monitoring regional cities, 121 cities had less than 292 good air quality days, which is less than 80% of the year.

The 11th goal of the United Nations 17 sustainable development goals is to build sustainable cities and communities, which requires efficient urban planning and management methods to meet the challenges of rapid urbanization. One of the seven specific goals for building sustainable cities and communities is to reduce the negative impact of the urban environment on people by 2030, especially in terms of air pollution. Solving the air pollution problem is not only a livelihood project to improve people's sense of well-being and people's health, but also an important way to promote sustainable development. Under this goal, it is necessary to try to understand the socioeconomic factors that might affect air pollution in order to facilitate pollution control.

From what has been discussed above, this paper attempts to use the panel data from 30 provincial capital cities(also known as provincial capitals) in China as a research basis to explore whether there is spatial heterogeneity in air pollution in Chinese provincial capitals, and what socioeconomic factors might have led to air pollution in Chinese provincial capitals, so as for future study to put forward corresponding policy suggestions and solutions accordingly.
2. Literature reviews and research hypotheses

2.1. Environmental Kuznets curve
G M Grossman and A B Krueger used panel data in GEMS to prove that there is an inverted U-shaped relationship between per capita income and environmental pollution. After crossing the critical income level, economic growth would be conducive to the improvement of air quality and water quality. The relationship between environmental pollution and per capita income is called the Environmental Kuznets Curve (EKC)[1]. S Shao et al analysed the energy data of Shanghai from 1994 to 2009 and testified the relationship between industrial carbon dioxide emissions (ICE) and output per capita through the cubic term[2]. Their analysis results showed that coal consumption is the largest source of industrial CO2 emissions. Based on this, this article proposes the hypothesis 1:

H1. The increase in GDP per capita is negatively related to air pollution and there is an Environmental Kuznets Curve between GDP per capita and air pollution.

2.2. Socioeconomic factors
Scholars also found that various social and economic factors may cause environmental pollution. J M Dean used provincial data to find that the net effect of foreign direct investment on water pollution is positive[3]. Y Hao and Y M Liu studied the data of 73 cities in China in 2013 and found that car ownership and the secondary industry have a significant impact on PM2.5[4]. S Shao et al used provincial data to prove that smog pollution is mainly caused by backward industrial structure, unreasonable energy structure and high population density[5]. J Liu et al used prefecture-level cities data to prove that the existence of Environmental Kuznets Curve and the proportion of secondary industry, technical level, greening level and population density have a significant impact on atmospheric quality[6]. In summary, this paper proposes the following hypotheses:

H2. The proportion of secondary industry is positively related to air pollution, and the proportion of tertiary industry is negatively related to air pollution.
H3. Population density is positively related to air pollution.
H4. Foreign direct investment is negatively related to air pollution.
H5. Clean energy is negatively related to air pollution.
H6. The number of cars is positively related to air pollution.

2.3. Spatial heterogeneity
X Y Li et al found that meteorological factors such as temperature and precipitation are significantly correlated with air pollution index through the study of nine typical cities in the north and dust storms are important factors affecting the spring air pollution of northern cities. When analysing the cities (Lanzhou and Urumqi) with the most serious air pollution, they also believed that the valley and basin topography also cause severe air pollution[7]. L Jiang et al found that there is a significant spatial heterogeneity in the effect of foreign direct investment on air pollution by geographic weighted regression[8]. Therefore, this paper reasoned out the hypothesis 7:

H7. There is spatial heterogeneity in air pollution in Chinese provincial capitals.

3. Data and model

3.1. Study area and indicator
This paper intends to focus on provincial capitals as they are the most representative cities of the social-economic development of China. Due to the lack of data, four cities are missed out, which included Hong Kong, Macau, Taipei and Lhasa.
Air quality index (AQI) is a dimensionless index developed by the Ministry of Environmental Protection and it is used to replace air pollution index (API) to measure air pollution. AQI reflects air pollution more accurate than API. Therefore, this paper chooses AQI as an indicator to measure air pollution.
3.2. Data source
The AQI and six pollutants data from 2013 to 2017 are extracted from the "China Environment Statistical Yearbook" from 2014 to 2018, and the AQI and six pollutants data for 2018 are extracted from the "China Statistical Yearbook" in 2019. The socioeconomic data from 2013 to 2017 are all extracted from the "China City Statistical Yearbook" from 2014 to 2018.
It should be noted that the statistical yearbook does not give the annual AQI numerical value, so this paper uses days of air quality equal to or above Grade II (i.e., when AQI ≤ 100) to replace AQI.

3.3. Econometric model
According to the previous research results and considering the availability and comparability of data, this paper uses days of air quality equal to or above Grade II (i.e., when AQI ≤ 100) as the explained variable (healthy days), which represents air pollution. And this paper selects the industrial structure, foreign direct investment, population density, number of cars, clean energy and economy as explanatory variables. The details of all variable are shown in Table 1.

| Variable abbreviation | Variable name       | Variable definition                                      |
|-----------------------|---------------------|----------------------------------------------------------|
| Hday                  | healthy days        | Days when AQI ≤ 100                                      |
| X2                    | secondary industry  | output value of secondary industry / GDP                 |
| X3                    | tertiary industry   | output value of tertiary industry / GDP                  |
| PGDP                  | economy             | per capita GDP                                           |
| FDI                   | foreign direct investment | amount of foreign capital actually utilized              |
| CE                    | clean energy        | total gas supply                                          |
| CAR                   | number of cars      | number of buses, trolley buses and taxis                 |
| PD                    | population density  | population density                                       |

In empirical researches, an additive regression model after taking the natural logarithm facilitates empirical estimation and hypothesis testing. And this paper uses panel data. Thus, the econometric model is as follows:

\[
\ln H_{day_{it}} = \alpha + \beta_1 \cdot \ln X_{2_{it}} + \beta_2 \cdot \ln X_{3_{it}} + \beta_3 \cdot \ln PGDP_{it} + \beta_4 \cdot \ln PGDP^2_{it} + \beta_5 \cdot \ln PD_{it} + \beta_6 \cdot \ln FDI_{it} + \beta_7 \cdot \ln CE_{it} + \beta_8 \cdot \ln CAR_{it} + \mu_i + \epsilon_{it}
\]

where subscript \( i \) represents cities, \( t \) represents the time in years, \( PGDP^2 \) represents the quadratic term of per capita GDP, \( \beta_1-\beta_8 \) are coefficients to be estimated, \( \mu_i \) represents the fixed effect, \( \alpha \) is the constant and \( \epsilon \) is the error term. Because of the cross-section dependence, Driscoll-Kraay standard error is adopted in this paper.

4. Data analyses and results

4.1. Air pollution status
The changing trend of the number of healthy days (i.e., when AQI ≤ 100) in 30 provincial capitals from 2013 to 2018 is shown in Figure 1 below. The gap between the two broken lines represents the improvement of air quality in each provincial capital.
As can be seen from Figure 1, 29 out of 30 provincial capitals have more healthy days in 2018 than in 2013. The only exception is Fuzhou, where the number of healthy days is reduced by 6 days, a decrease of about 2%. Although healthy days of 29 provincial capitals have increased to varying degrees, the overall air pollution situation is still not optimistic. Taking 2018 as an example, the accumulated number of healthy days in 30 provincial capitals was 8015 days, accounting for only 73.2% of the whole year.

4.2. Spatial heterogeneity of air pollution

First, this paper uses the average concentration of the six pollutants, which are PM2.5, PM10, NO2, CO, SO2 and O3 from 2013 to 2018 by cluster analysis to investigate if there is a spatial heterogeneity of the air pollution in provincial capitals[9]. The results of cluster analysis are shown in Table 2.

| Category | Nanjing | Hangzhou | Chongqing | Shanghai | Guangzhou | Wuhan |
|----------|---------|----------|-----------|----------|-----------|-------|
| I        | Nanjing | Xining   | Hohhot    | Kunming  | Nanchang  | Haikou |
| II       | Taiyuan | Yinchuan | Guiyang   | Shenyang | Changchun |       |
| III      | Lanzhou | Urumqi   | Xi’an     | Hefei    | Zhengzhou | Shijiazhuang |

According to the result of one-way ANOVA, there is a significant difference in the degree of air pollution among the three categories of cities (F = 35.84, P < 0.00). This shows that the H7 has been testified, that is, there is spatial heterogeneity in air pollution of provincial capitals. According to the standard of healthy days, the air pollution level of category II cities is the least severe, followed by category I and category III cities. From the perspective of the region of the three categories of cities, category II cities with the least air pollution are mostly located in the southwest and northeast regions. Category I cities with higher air pollution are mostly located in the southeast coastal areas, while category III cities with the most severe air pollution are mostly located in the surrounding areas of Beijing-Tianjin-Hebei and the northwest regions. Obviously, the economic situation of the three categories of cities is also different. The result of one-way ANOVA shows that there is a significant difference in per capital GDP between the three categories of cities (F = 6.48, P < 0.00).
4.3. Regression results and robustness analysis

According to the regression (1) in Table 3, the research hypotheses H2, H4, H5, H6 have been testified, the research hypothesis H3 failed testify, and the research hypothesis H1 has been partially testified. The specific analysis results are as follows. The regression coefficient of quadratic term of GDP per capita is significantly positive, which indicates the existence of an U-shaped EKC. But the regression coefficient of one-time item of GDP per capita is not significant. H1 has been partially testified. The increase in the proportion of the secondary industry would increase air pollution, while the increase in the proportion of the tertiary industry can reduce air pollution. H2 has been testified. Foreign direct investment and clean energy have a negative effect on air pollution. H4 and H5 have been testified. The number of cars has a positive impact on air pollution. H6 has been testified. Population density has no significant effect on air pollution, thus H3 failed testify.

|               | (1)          | (2)          |
|---------------|--------------|--------------|
| Hday          | -0.309**     | 0.454**      |
| lnX2          | (0.0676)     | (0.117)      |
| lnX3          | 0.708***     | -2.211***    |
|               | (0.0290)     | (0.146)      |
| lnPGDP        | -0.103       | 1.449***     |
|               | (0.158)      | (0.233)      |
| lnPGDP2       | 0.0784**     | -0.456***    |
|               | (0.0224)     | (0.0499)     |
| lnPD          | 0.0208       | -0.0732      |
|               | (0.0236)     | (0.0484)     |
| lnFDI         | 0.0341***    | -0.0116      |
|               | (0.00469)    | (0.0236)     |
| lnCE          | 0.193**      | -0.353**     |
|               | (0.0622)     | (0.0907)     |
| lnCAR         | -0.0975**    | 0.179**      |
|               | (0.0245)     | (0.0487)     |
| cons          | 4.423**      | -0.341       |
|               | (1.035)      | (1.718)      |
| N             | 150          | 150          |
| \(R^2\)       | 0.483        | 0.544        |

Table 3. Results of fixed effect model.

In some cases, AQI cannot reflect the comprehensive impact of the six pollutants [10]. Therefore, this paper uses exploratory factor analysis to obtain the air pollution score “F” as a new dependent variable for robust analysis. The results of robust analysis are shown as the regression (2) in Table 3. Except for foreign direct investment and GDP per capita, the significance test results of other variables have not changed. This shows that the regression results are relatively robust.
5. Discussion and conclusions
From the time dimension, air pollution in Chinese provincial capitals has been reduced to varying degrees, but it is still not optimistic. From the spatial dimension, there is obvious spatial heterogeneity in the air pollution of provincial capitals. The air pollution level in provincial capitals of the southwest, northeast and coastal regions is relatively light. The air pollution level in the provincial capitals of the Beijing-Tianjin-Hebei region and the northwest region is relatively heavy.
The tertiary industry and clean energy can reduce air pollution, while the number of cars and the secondary industries might aggravate air pollution level. Population density has no significant impact on air pollution in provincial capitals. The significance test results of GDP per capita in the two regressions are different, but regression coefficients of quadratic term of GDP per capita are significant. This may be because GDP per capita does not represent the economic development level comprehensively.

6. Prospect
Although AQI may reflect air pollution more accurately than API, there is a series of comprehensive impact of pollutant sources has not been included in some cases in this research. Therefore, we intend to consider and gather more data to testify in future research.

Acknowledgments
This paper is supported by the foundation of The National Social Science Fund of China, Project ID: 18CGL040.

References
[1] Grossman, G.M., Krueger, A.B. (1995) Economic Growth and the Environment. The Quarterly Journal of Economics., 110:353-377.
[2] Shao S, Yang L.L., Yu M.B., Yu M.L. (2011) Estimation, characteristics, and determinants of energy-related industrial CO2 emissions in Shanghai (China), 1994-2009. Energy Policy., 39:6476-6494.
[3] Dean J.M. (2002) Does trade liberalization harm the environment? A new test. Canadian Journal of Economics/Revue canadienne d’Economique, 35:819-842.
[4] Hao Y, Liu Y.M., (2016) The influential factors of urban PM2.5 concentrations in China: a spatial econometric analysis. Journal of Cleaner Production., 112: 1443-1453.
[5] Shao S, Li X, Cao J.H., Yang L.L. (2016) China's Economic Policy Choices for Governing Smog Pollution Based on Spatial Spillover Effects. Economic Research Journal., 51:73-88.
[6] Liu J, Wang H.W., Yang J. (2017) Research into the Influential Factors of Air Pollution in China: An Analysis of Dynamic Spatial Panel Model of Chinese Cities. Journal of Hohai University (Philosophy and Social Sciences), 19:61-67+91-92
[7] Li X.Y., Ding X.M. et al. (2011) Characteristics of air pollution index in typical cities of North China. Journal of Arid Land Resources and Environment, 25:96-101.
[8] Jiang L, Zhou H.F., Bai L. (2018) Spatial Heterogeneity Analysis of Impacts of Foreign Direct Investment on Air Pollution: Empirical Evidence from 150 Cities in China Based on AQI. Scientia Geographica Sinica., 38:351-360.
[9] Wang B, Gao H.W. (2008) Characteristics of air pollution index in coastal cities of China. Ecology and Environmental Sciences, 02:542-548.
[10] Hu, J.L., Ying Q, Wang Y.G., Zhang H.L. (2015) Characterizing multi-pollutant air pollution in China: Comparison of three air quality indices. Environment International, 84:17-25.