



Association between air pollution and COVID-19 infection: evidence from data at national and municipal levels

Qiang Liu^{1,2} · Shengxia Xu¹ · Xiaoli Lu¹

Received: 13 October 2020 / Accepted: 3 March 2021 / Published online: 13 March 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The impact of high concentrations of air pollution on COVID-19 has been a major air quality and life safety issue in recent studies. This study aimed to assess the contribution of different air pollution indicators in different spaces on the newly confirmed cases of coronavirus. According to causality's results between air pollution (AP) and COVID-19 infection in 9 countries, first, we examined the non-linear relationship from AP to COVID-19 with PM_{2.5} as the rating variable (the cut point is 35 µg/m³) at the national level. It is concluded that the effects of PM_{2.5} and PM₁₀ on COVID-19 are more sensitive in Russia, England, Germany, and France, while O₃ and PM_{2.5} are more sensitive in America and Canada from 21 Jan to 20 May. Second, we examined the threshold effects from AP to COVID-19 with PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O₃ as the threshold variables, respectively, at the municipal level in China during the period 28 Jan to 31 May. It is concluded that except CO, the remaining 5 indicators are very sensitive to the increase of newly confirmed cases, and the spread of COVID-19 can be prevented and controlled by the determination of thresholds. In addition, the 9 countries and 27 provinces show that PM_{2.5} in high concentrations is the more sensitive pollutant on the spread of COVID-19 infection.

Keywords COVID-19 · Coronavirus · Air pollution · PM_{2.5} · Threshold effect · 9 countries

Introduction

The urgency and uncertainty specific to major public emergencies have brought great challenges to the development of the global economy, and since the unexpected pneumonia was

first reported on December 2019 in Wuhan, it spread rapidly to the rest of China, which has later become a global public health problem and was declared as a pandemic by the WHO (World Health Organization 2020). The coronavirus was named after crown spikes were found on its surface (Zaki

Highlights

- Unidirectional causal linkages from environmental pollution to COVID-19 infection were examined at the national level.
- Threshold associations of PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O₃ with COVID-19 newly confirmed cases were observed.
- PM_{2.5} in high concentrations is the more sensitive pollutant on the spread of COVID-19 infection regardless of whether it is in the national level or municipal level.
- Effects of PM_{2.5} and PM₁₀ on COVID-19 are more sensitive in Russia, England, Germany, and France, while O₃ and PM_{2.5} are more sensitive in America and Canada.

Responsible Editor: Lotfi Aleya

✉ Shengxia Xu
cuebxushengxia@163.com

Qiang Liu
cuebliuqiang@163.com

Xiaoli Lu
luxiaolicueb@163.com

¹ School of Statistics, Capital University of Economics and Business, No. 121, Zhangjia Road, Huaxiang, Fengtai District, Beijing 100070, China

² Beijing Key Laboratory of Megaregions Sustainable Development Modelling, Beijing 100070, China

et al. 2012), and the novel coronavirus (COVID-19) is an acute respiratory disease which may lead to pneumonia with symptoms such as fever, cough, and dyspnea (Jiang et al. 2020).

As of 18 June 2020, there have been > 8 million confirmed cases and over 450 thousand deaths were reported globally, with an approximate fatality rate greater than 2–3% as proposed by Rodriguez-Morales et al. (2020). Figure 1 presents the trend of newly confirmed cases of 9 countries of Asia, North America, and Europe during 21 Jan to 20 May in 2020. Asian countries are generally close to 0 after May, especially after experiencing a large outbreak in February in China; the newly confirmed cases fluctuated slightly within 100 from March. For the remaining 6 countries, April and May are the outbreak periods of COVID-19.

In order to control the spread of COVID-19, various studies have been conducted to explore important factors affecting the transmission of SARS-CoV-2 (Zhu et al. 2020). For example, Wang et al. (2020) have demonstrated human-to-human transmission of COVID-19 through direct contact or droplets; Kraemer et al. (2020) have examined a significant effect from population mobility to COVID-19 infection; Xie and Zhu (2020) have explored an association of ambient temperature with the infection of COVID-19. Previous studies have suggested that ambient air pollutants are risk factors for respiratory infection (Cai et al. 2007; Xu et al. 2016); therefore, it is interesting to investigate the effect of air pollution on COVID-19 infection, since COVID-19 is a respiratory disease (Zhu et al. 2020; Yang et al. 2020). Table 1 presents the mean value of six air pollution indicators of 9 countries during the same period with newly confirmed cases of COVID-19. Obviously, fine particulate matter (PM_{2.5}) is an air pollution index shared by all countries, so in the study at the national level, it must be the focus of our attention.

In addition, some studies have shown that the incidence of these diseases can also be caused by a long exposure to air

pollution, especially nitrogen dioxide (NO₂) which is one of its toxic components (Ogen 2020; Saeha et al. 2020); some studies have shown that ozone (O₃) in the troposphere plays a central role in the oxidation of chemically and climatically relevant trace gases and can regulate people's lifetime in the atmosphere (Monks et al. 2015; Wang et al. 2016); some studies have shown that chronic atmospheric pollution may favor coronavirus spreading, and air pollution (AQI as indicator) should be considered in epidemic prevention (Zhang et al. 2020; Fattorini and Regoli 2020), and so on.

The above-mentioned studies started from different perspectives of air pollution and conducted a series of researches; it can always be obtained that air pollution has a non-negligible effect on the spread of COVID-19. However, for research content, 6 air pollution indicators such as fine particulate matter (PM_{2.5}), inhalable particles (PM₁₀), sulfur dioxide (SO₂), carbon monoxide (CO), nitrogen dioxide (NO₂), and ozone (O₃) have not been found to study the impact of COVID-19 infection both from the overall and individual aspects, but these 6 indicators have shown a strong effect on COVID-19 from the different studies, so further research is needed. For the existing research, it has not been found that studies that consider both the impact of national air pollution on COVID-19 and the with-national differences, which are often an indispensable part of the COVID-19 infection study. Therefore, in this study, based on the results of dynamic interrelationships between air pollution and COVID-19 infection in 9 countries by focusing on accounting for structural changes in causal linkages, we examine the non-linear effects of different air pollution indicators on COVID-19 from the spatial and temporal dimensions both at the national and municipal levels. The objective of this work is to assess the contribution of different air pollution indicators in different spaces on the newly confirmed cases of coronavirus.

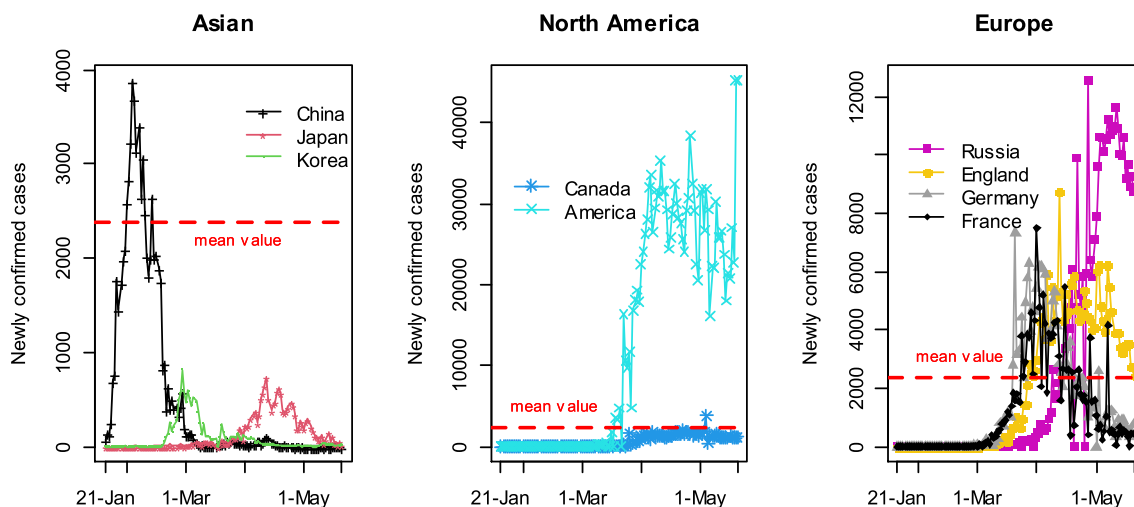


Fig. 1 Newly confirmed cases of COVID-19 infection from the national level

Table 1 Mean value of air pollution from the national level

AP	China	Japan	Korea	Canada	America	Russia	England	Germany	France
PM _{2.5}	108.19	39.17	79.36	26.20	41.26	42.55	38.09	34.67	53.96
PM ₁₀	48.50	11.65	42.98	N	N	28.53	17.05	20.64	28.13
SO ₂	N	N	2.70	N	N	N	N	N	N
CO	5.83	2.55	5.38	1.48	N	N	N	N	N
NO ₂	13.61	12.40	28.55	6.60	N	18.36	11.97	15.29	26.80
O ₃	41.61	31.22	30.26	28.53	29.06	17.85	N	N	29.78

Note: Letter “N” indicates that this indicator is not included in air pollution

Materials and methods

Database of air pollution and COVID-19 infection

The data of PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O₃ are used to characterize the indicator of air pollution has been collected from real-time air quality index of Air Pollution in the World database.¹ And the municipal levels’ data was collected from the National Health Commission of the People’s Republic of China. However, due to the huge workload of collecting air pollution data from countries around the world, we only selected data from 9 countries in Asia, North America, and Europe for studying. Moreover, not all 6 indicators representing air pollution in selected countries are included; therefore, we show the details in Table 2.

The data concerning the number of newly confirmed cases (NCC) was collected and organized from daily report published on the official website of the WHO, and each country on a regional/administrative level. For some regions/countries where only cumulative confirmed cases data is released, we calculate them based on date data. The use of this method is intended to highlight the spatial variation of the epidemic which exists not only between the different countries but also inside each country. In terms of the difference within a country, we take 27 municipalities in China, excluding 7 regions including Hong Kong, Macau, Taiwan, Xinjiang, Tibet, Hainan, and Inner Mongolia, as an example for analysis in this paper.

Moreover, if highly new confirmed cases are observed in two different countries, we need to identify their common factors which may explain why. Therefore, data was collected from 9 countries, China, Japan, Korea, Canada, America, Russia, England, Germany, and France. In addition, information about new confirmed cases was also taken from the Ministry of Health, the national agency of public health, or the state health offices of each country to check with the international data and make adjustments.

Design of causality test

This study examines the association between air pollution and COVID-19 infection from the different levels. Before deciding on the final role-relationship model, we need to test the causal linkage between air pollution and COVID-19 infection. A simple approach to account for structural breaks (both abrupt and gradual) in causality analysis proposed by Nazlioglu et al. (2016) and developed by Durusu-Ciftci et al. (2020) after augmenting the Toda-Yamamoto framework with a Fourier approximation. They relax the assumption that the intercept terms are constant over time, and the gradual structural shifts with an unknown time, number, and forms of breaks are captured by the Fourier approximation. In line with that, we define the causal linkage between air pollution (AP) and COVID-19 infection as

$$AP_t = c_{10} + \sum_{k=1}^K a_{1k} \sin\left(\frac{2\pi tk}{T}\right) + \sum_{k=1}^K a_{2k} \cos\left(\frac{2\pi tk}{T}\right) + \sum_{j=1}^{p+d} \beta_{1j} \text{COVID}_{t-j} + \varepsilon_{1t} \tag{1}$$

$$\text{COVID}_t = c_{20} + \sum_{k=1}^K a_{1k} \sin\left(\frac{2\pi tk}{T}\right) + \sum_{k=1}^K a_{2k} \cos\left(\frac{2\pi tk}{T}\right) + \sum_{j=1}^{p+d} \beta_{2j} \text{COVID}_{t-j} + \varepsilon_{2t} \tag{2}$$

where $t(t = 1, 2, \dots, T)$ is the t th period, c_{10} and c_{20} are the vector of intercept terms, K is the number of frequencies, p is the lag order, d is the maximum integration order of variables, β_{1j} and β_{2j} are coefficient matrices, and ε_{1t} and ε_{2t} are the white-noise residuals. In addition, a_{1k} and a_{2k} can calculate the amplitude and displacement of the frequency, separately. In addition, COVID_t represents the new confirmed cases of COVID-19 at t time and AP_t represents the air pollution. The assumption here is that air pollution and COVID-19 infection variables will change as the structure shifts.

¹ Data sources: <http://aqicn.org>

Table 2 Indicators of air pollution from different countries

AP	Asia			North America		Europe			
	China	Japan	Korea	Canada	America	Russia	England	Germany	France
PM _{2.5}	Y	Y	Y	Y	Y	Y	Y	Y	Y
PM ₁₀	Y	Y	Y	N	N	Y	Y	Y	Y
SO ₂	N	N	Y	N	N	N	N	N	N
CO	Y	Y	Y	Y	N	N	N	N	N
NO ₂	Y	Y	Y	Y	N	Y	Y	Y	Y
O ₃	Y	Y	Y	Y	Y	Y	N	N	Y

Note: Letter “Y” indicates that this indicator is included in air pollution, and letter “N” indicates not included

The most important problem in Eq. (1) and Eq. (2) is that it requires determining the number of cumulative Fourier frequency K and lag lengths p . In order to determine the optimal number of lags in a causality analysis in common is to benefit from Akaike information criterion (AIC). As the recent works in the Granger causality literature shown (Buhlmann and Meneil 2002; Durusu-Ciftci et al. 2020), AIC can also be used for determining the number of Fourier frequency and lag lengths, wherefore, we also follow the above convention, using AIC to determine Fourier frequency K and lag lengths p in this paper.

Statistical analysis

The regression discontinuity design (RDD) can alleviate the endogenous problem of parameter estimation (Thistlethwaite and Campbell 1960); meanwhile, in some situations, RDD can be embedded in a panel context (Lee and Card 2008), whereby period by period, the treatment variable is determined according to the realization of the assignment variable, which has recently been used in more studies (Tang et al. 2019; Federico and BugniaIvan 2020). Considering the unity (both for time and individuals) and difference (only for individuals) of international comparison, we use RDD to analyze the relationship between air pollution and COVID-19 infection at the national level, and the model was defined as follows:

$$COVID_{it} = \alpha_1 + \beta_1(AP_{it}-c)^k + \tau D_{it} + \beta_2(AP_{it}-c)^k D_{it} \quad (3)$$

where $i(i = 1, 2, \dots, N)$ and $t(t = 1, 2, \dots, T)$ denote the individuals and time, respectively. $COVID_{it}$ represents the new confirmed cases of COVID-19 in country i at time t ; α_i stand for the individual effect; EP_{it} represents the air pollution in country i at time t portrayed by $PM_{2.5}$; c is a constant term senting by the cut point; D_{it} is the treatment variable and can be set as a dummy variable:

$$D_{it} = \begin{cases} 1, AP_{it} > c \\ 0, AP_{it} \leq c \end{cases} \quad (4)$$

k is the power of the polynomial, usually replaced by 1 to 4, which is determined by comparing AIC; β_1 and β_2 are the parameters to be estimated, and τ is the marginal treatment efficiency of air pollution intervention at the cut point. It needs to be emphasized that we introduce control variables (other air pollution indicators other than $PM_{2.5}$) on the basis of Eq. (3) to analyze the differences between countries:

$$COVID_{it} = a_i + \beta_1(PM_{2.5it}-c)^k + \tau D_{it} + \beta_2(PM_{2.5it}-c)^k D_{it} + \gamma_{1i}PM_{10it} + \gamma_{2i}O_{3it} + \gamma_{3i}NO_{2it} + \gamma_{4i}SO_{2it} + \gamma_{5i}CO_{it} + \varepsilon_{it} \quad (5)$$

where we carry out Eq. (5) separately for each country ($i = 1, 2, \dots, 9$). Meanwhile, if the country does not contain air pollution variables, the corresponding parameter in the model is 0.

For analysis the difference within the country, considering the impact of multiple air pollution variables and the characteristics of the panel data composition of each variable, the RDD estimate has a certain deviation (Lee and Thomas 2010). And considering the goodness of fit and other statistical tests can help rule out overly restrictive specifications (Li et al. 2018; Vu 2020). Therefore, we use the threshold panel regression model to perform the effect analysis at the municipal level:

$$COVID_{it} = \mu_i + \alpha_1 d_{it} I(d_{it} \leq \gamma_1) + \alpha_2 d_{it} I(\gamma_1 < d_{it} < \gamma_2) + \dots + \alpha_{n+1} d_{it} I(d_{it} > \gamma_n) + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \beta_4 x_{4it} + \beta_5 x_{5it} + \varepsilon_{it} \quad (6)$$

where $i(i = 1, 2, \dots, n)$ and $t(t = 1, 2, \dots, T^*)$ denote the individuals and time, respectively. μ_i presents the individual effect; $COVID_{it}$ represents the new confirmed cases of COVID-19 in province i at time t ; d_{it} is both the threshold variable and the variable affected by the threshold, which is represented by 6 air pollution indicators ($PM_{2.5}$, PM_{10} , SO_2 , CO , NO_2 , O_3)

separately; $\gamma_1, \gamma_2, \dots, \gamma_n$ are the thresholds; $x_{1it}, x_{2it}, x_{3it}, x_{4it}, x_{5it}$ are the variables except the threshold of 6 air pollution indicators; $\alpha_1, \dots, \alpha_{n+1}, \beta_1, \dots, \beta_5$ are the parameters.

Empirical results

Descriptive analysis

Table 3 summarizes the descriptive statistics for air pollution variables and newly confirmed cases of COVID-19 infection. In general, this study consisted of two parts with different levels; on one side, this study included > 2.6 million confirmed cases during the observation period from the national level of unbalanced panel data (21 Jan 2020 to 20 May 2020), and the average number was 21,481 for the nine countries, with a minimum sample size of 121 and a maximum sample size of 1089. Average daily newly confirmed cases, PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O₃ were 2387, 51.49 $\mu\text{g}/\text{m}^3$, 28.21 $\mu\text{g}/\text{m}^3$, 2.7 $\mu\text{g}/\text{m}^3$, 3.81 $\mu\text{g}/\text{m}^3$, 16.7 $\mu\text{g}/\text{m}^3$, and 29.76 $\mu\text{g}/\text{m}^3$, respectively.

On the other side, this study included > 76,836 cases during the observation period from the municipal level in China (28 Jan 2020 to 31 May 2020), and the average number was 615; average daily newly confirmed cases, PM_{2.5}, PM₁₀, SO₂, CO, NO₂, and O₃ were 23, 38.8 $\mu\text{g}/\text{m}^3$, 63.37 $\mu\text{g}/\text{m}^3$, 31.69 $\mu\text{g}/\text{m}^3$, 101.52 mg/m, 0.73 $\mu\text{g}/\text{m}^3$, and 9.11 $\mu\text{g}/\text{m}^3$, respectively.

In the causality test, the lag order of NCC is 1 (Dickey-Fuller is -4.50), and the lag order of PM_{2.5} is also 1 (Dickey-Fuller, -5.94) by employing a battery of unit root tests (ADF); therefore, the lag order p can be determined as 1 and the maximum integration order of variables d can be determined as 1 also in Eq. (1) and Eq. (2). After iterating the number of frequencies K from 1 to 5, and comparing the Akaike information criterion (AIC), the smallest AIC which corresponds to K is selected as 2. Moreover, there is no causality in any direction between AP and COVID-19 under the

Toda-Yamamoto test framework (Wald are 0.57, 0.14, separately), while there exists unidirectional causal relationships from AP to COVID-19 when the structural breaks are taken into account with a Fourier approximation (Wald are 2.72, 0.51, separately). Therefore, the result for causality between air pollution (AP) and COVID-19 infection is also up to the test's method to some extent.

Relationship between AP and COVID-19 from the national level

The rating variable's histogram and density in Fig. 2 suggested that the variable of AP was in line with the conditions for internal validity, and the cut point was 35 $\mu\text{g}/\text{m}^3$; moreover, in the case analysis of this paper, intervention of AP has clear boundaries before and after the cut point, so it is of the sharp type, and the relationship between AP and newly confirmed cases of COVID-19 was non-linear significantly. Specifically, the relationship was approximately a positive linear in the range of 35 $\mu\text{g}/\text{m}^3$ and became opposite above 35 $\mu\text{g}/\text{m}^3$, indicating that the single threshold of the AP effect on COVID-19 was 35 $\mu\text{g}/\text{m}^3$.

According to the results of RDD, a discontinuous linear regression was adapted with a cut point at 35 $\mu\text{g}/\text{m}^3$ to quantify the effect of air pollution (here is PM_{2.5} only) above and below the cut point. In Table 4, no matter what the power of the polynomial k is, the marginal coefficient τ of the treatment at the break point always passes the significance test, and is all negative from column (1) to column (5). This shows that the treatment D_{it} of PM_{2.5} on newly confirmed cases of COVID-19 infection is existing and with a significant threshold effect. Moreover, column (2) is the optimal estimate result for Eq. (3) after comparing the AIC from the 5 models, and then we can take k as 1 in Eq. (3) for the latter study. Based on the result of column (2), the coefficient of PM_{2.5} on newly confirmed cases of COVID-19 infection is significantly positive above 35 $\mu\text{g}/\text{m}^3$ (66.72). Judging from the overall results of the nine

Table 3 Descriptive statistics of newly confirmed cases and air pollution variables

Variables	From the national level				From the municipal level			
	Mean (SD)	Min	Max	Samples	Mean (SD)	Min	Max	Samples
NCC	2387 (6235.14)	0	45,251	1089	23 (310.64)	0	14,840	3375
PM _{2.5}	51.49 (35.57)	4	261	1089	38.8 (31.27)	3	906	3375
PM ₁₀	28.21 (18.92)	2	124	847	67.37 (41)	0	644	3375
SO ₂	2.7 (2.67)	1	12	121	31.69 (13.9)	4	119	3375
CO	3.81 (3.21)	0	23	484	101.52 (37.48)	7	283	3375
NO ₂	16.7 (10.58)	0	64	968	0.73 (0.27)	0.1	2.8	3375
O ₃	29.76 (13.84)	1	146	847	9.77 (5.94)	2	50	3375

Note: "NCC" is the newly confirmed cases, and PM_{2.5}, PM₁₀, SO₂, CO, NO₂, O₃ are the air pollution variables

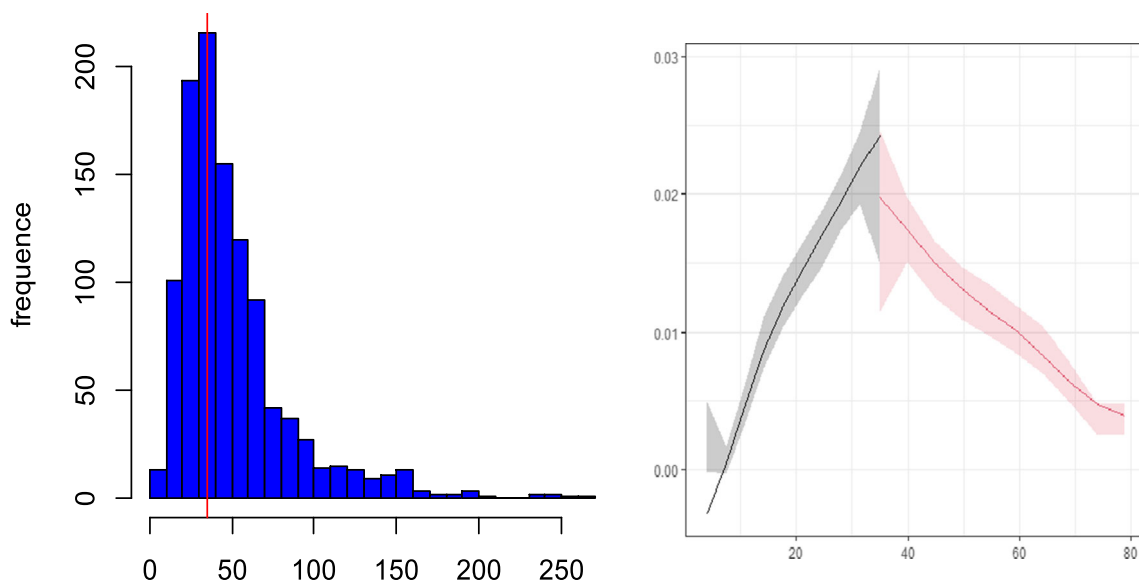


Fig. 2 Histogram and density of air pollution (take PM_{2.5} for example)

existing countries, this means that when the degree of air pollution continues to increase, it will aggravate the spread of COVID-19 infection to a certain extent. Therefore, controlling the air pollution within reasonable limits is one of the effective ways to slow down COVID-19 infection.

In order to compare the difference of cut-points that made by each country, we remove the individual influencing effect in Eq. (3), and then estimate the model for each country. As shown in Fig. 3, the histogram of each country’s rating variable (PM_{2.5}) presents a relatively obvious internal validity, which is suitable for RDD at first.

Secondly, there are great differences in the values of the cut points for the 9 countries; the cut points in Germany, England, Japan, America, and Russia are closer to the overall break point (it is 35) at the national level, while China has the largest deviation (it is 108.19), 3 times larger than it.

Table 5 presents the results of RDD for the relationship from air pollution only including PM_{2.5} to COVID-19. The marginal coefficients of τ of the treatment in the 9 countries showed great differences in size and sign at different cut points, but the coefficients with negative values as a whole passed the test, basically consistent with the conclusions in

Table 4 Results of RDD from the panel data of 9 countries (without control variables)

Variables	(1)	(2)	(3)	(4)	(5)
$AP_{it}-c$	2.36 (6.66)	66.72 [*] (37.79)	-3.54 (13.71)	81.80 (124.29)	7.85 (19.77)
$(AP_{it}-c)^2$			0.04 (0.08)	0.63 (4.79)	-0.18 (0.29)
$(AP_{it}-c)^3$					0.00 (0.00)
D_{it}	-1017.59 [*] (406.75)	-1588.65 ^{**} (523.58)	-875.33 [*] (428.91)	-1423.38 [*] (704.44)	-1111.60 [*] (501.95)
$(AP_{it}-c) \times D_{it}$		-66.28 [*] (38.32)		-97.44 (125.20)	
$(AP_{it}-c)^2 \times D_{it}$				-0.52 (4.79)	
df	4	5	5	7	6
R^2	2115.20	2101.87	2116.55	2103.33	2111.54
AIC	0.166	0.321	0.158	0.232	0.174

Note: cut point c is 35, the power of the polynomial k is 1 from the smallest AIC, and the standard errors are shown in brackets

^{*} $P < 0.1$; ^{*} $P < 0.05$; ^{**} $P < 0.01$; ^{***} $P < 0.001$

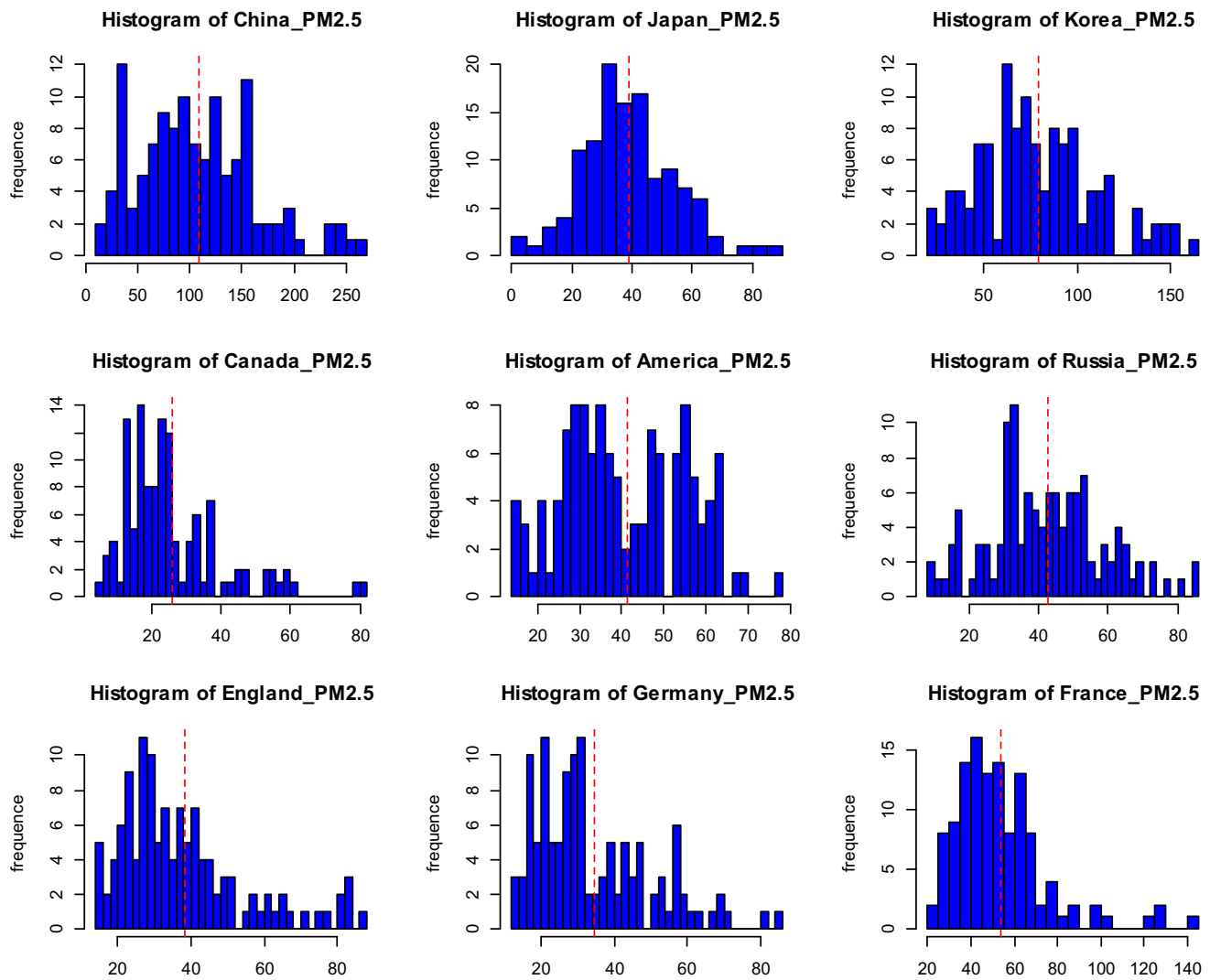


Fig. 3 Histogram of PM_{2.5} determined by cut points for each country. Note: the red dotted lines are the vertical lines of PM_{2.5} for each country at the cut points

Table 5 Estimation of RDD for each country without considering other environmental variables

Variables	Asia			North America		Europe			
	China	Japan	Korea	Canada	America	Russia	England	Germany	France
Intercept	148.7* (227.0)	122.4*** (35.65)	75.69* (34.4)	843.7*** (134.3)	17,126*** (3177.3)	4822.4*** (869.11)	1762.47*** (520.25)	1554.92*** (451.44)	1399*** (364.81)
AP _{it} - c	-6.24 (6.60)	1.87 (2.60)	-0.02 (1.23)	-2.28 (14.29)	146.9 (143.0)	110.24 (55.95)	-8.50 (41.87)	9.00 (7.84)	35.61 (23.56)
D _{it}	165.26 (301.3)	66.87 (51.18)	54.27* (26.73)	-578.2* (221.3)	-3890.4 (5141.2)	-2738.8* (1179.67)	1219.98 (721.49)	-1031.33* (450.73)	240.59 (166.01)
(AP _{it} - c) × D _{it}	14.65** (5.62)	-2.92* (1.37)	-0.68 (1.58)	-1.64 (16.57)	-421.3 (355.9)	-142.53 (72.39)	-27.40 (47.44)	46.91 (44.37)	-40.31* (15.79)
k	1	1	1	1	1	1	1	1	1
Cut point	108.19	39.17	79.36	26.20	41.26	42.55	38.09	34.67	53.96
R ²	0.269	0.142	0.102	0.374	0.189	0.383	0.112	0.151	0.201

Note: air pollution (AP) presents by PM_{2.5} in this table, and the standard errors are shown in brackets

· P < 0.1; * P < 0.05; ** P < 0.01; *** P < 0.001

Table 4. When $PM_{2.5}$ exceeds the break points, the positive increase of air pollution to COVID-19 infection is obvious in America, Russia, Germany, and Japan, while the remaining 5 countries showed inhibitory effects, but none of them passed the test. This shows that air pollution has a certain aggravating effect on the spread of COVID-19 infection. When the value of $PM_{2.5}$ is greater than the cut points, this aggravating effect will be abrupt, which needs to be paid attention to by governments of various countries. In addition, considering the role of $PM_{2.5}$ on COVID-19 infection only, it is not only difficult to fully understand the effect of air pollution on coronavirus disease but also difficult to ensure the estimated effect of the model, with goodness of fit being less than 0.4 for the nine national models, which is not the standard of statistical tests.

The solution in the above respects, we include other indicators PM_{10} , SO_2 , CO , NO_2 , and O_3 that represent air pollution into the model, and estimate Eq. (5) for the 9 countries. As shown in Table 6, marginal coefficients of τ in the 9 countries are detected for a larger set of countries and more significance when the other 5 air pollution indicators are taken into account on one hand. On the other hand, the critical effect of $PM_{2.5}$ on COVID-19 infection is significantly weakened, but the overall effect of air pollution on COVID-19 infection

does not decrease but increases. Specifically, PM_{10} plays a stronger role in accelerating the spread of COVID-19 infection in China, England, Germany, and France; O_3 presents a more pronounced positive effect on COVID-19 infection in more countries (such as Japan, Canada, America, Russia, France, etc.); the aggravating effect of NO_2 on COVID-19 infection appears in Canada and France; SO_2 and CO will increase the propagation speed of COVID-19 infection, which is significant in Korea and China, respectively.

Overall, we can sum up that (i) accounting for air pollution has an important effect on COVID-19 infection and (ii) the sensitivity of COVID-19 infection in different countries to 6 environmental pollutants is very different, but the threshold effect on $PM_{2.5}$ is significant when the other 5 air pollution indicators are taken into account. To summarize further, COVID-19 infection in European countries is more sensitive to environmental pollutants PM_{10} and $PM_{2.5}$, North American countries are more sensitive to O_3 and $PM_{2.5}$, while Asian countries have no rules. Specifically, China is more sensitive to PM_{10} , CO , and $PM_{2.5}$; Japan is more sensitive to O_3 and $PM_{2.5}$; while Korea is more sensitive to SO_2 and $PM_{2.5}$. Hence, policy implications may differ based on the relationship analysis with or without the cut point, as presented in Fig. 4.

Table 6 Estimation of RDD for each country with consideration for other environmental variables

Variables	Asia			North America		Europe			
	China	Japan	Korea	Canada	America	Russia	England	Germany	France
Intercept	950.2** (311.9)	83.27 (81.7)	267.6* (104.3)	1413*** (393.47)	-5458.8 (4941.5)	7703.9*** (2005.6)	1120.75* (536.74)	439.06 (688.33)	-739.63 (701.9)
$PM_{2.5}-c$	-3.77 (4.12)	0.58 (2.43)	0.80 (1.26)	3.78 (12.21)	120.50 (216.70)	127.72** (48.45)	-46.24 (33.01)	-13.15 (35.09)	30.02 (19.59)
D_{it}	60.37 (263.5)	29.02 (48.0)	42.44 (48.37)	-518.1** (186.97)	-901.1* (616.0)	-1171.93* (1043.12)	1009.99* (563.28)	-1256.7* (597.43)	-332.73* (282.3)
PM_{10}	13.5*** (3.54)	-1.28 (2.81)	-0.42 (1.26)			-23.15 (20.76)	183.31*** (25.25)	80.42*** (16.40)	30.36* (15.25)
O_3	-4.71 (3.73)	5.22** (1.67)	-4.08* (3.01)	32.26** (11.15)	706.20*** (126.60)	115.89* (48.93)			85.08*** (11.71)
NO_2	-13.11 (13.80)	-4.29 (4.44)	2.25 (1.78)	71.39** (22.78)		265.33* (263.06)	-221.5 (228.34)	-33.39 (29.03)	44.04** (15.21)
SO_2			1.64* (0.77)						
CO	86.08** (29.19)	-17.65 (18.8)	-19.06 (13.63)	-97.47 (142.86)					
$(PM_{2.5}-c) \times D_{it}$	4.53 (5.18)	-1.81 (3.20)	-1.24 (1.60)	-0.97 (14.11)	-302.40 (318.10)	-170.19** (62.11)	10.38 (37.98)	55.30 (40.78)	-17.73 (21.88)
Cut point	108.19	39.17	79.36	26.20	41.26	42.55	38.09	34.67	53.96
R^2	0.457	0.273	0.234	0.642	0.257	0.538	0.623	0.276	0.483

Note: the standard errors are shown in brackets

* $P < 0.1$; ** $P < 0.05$; *** $P < 0.01$; **** $P < 0.001$

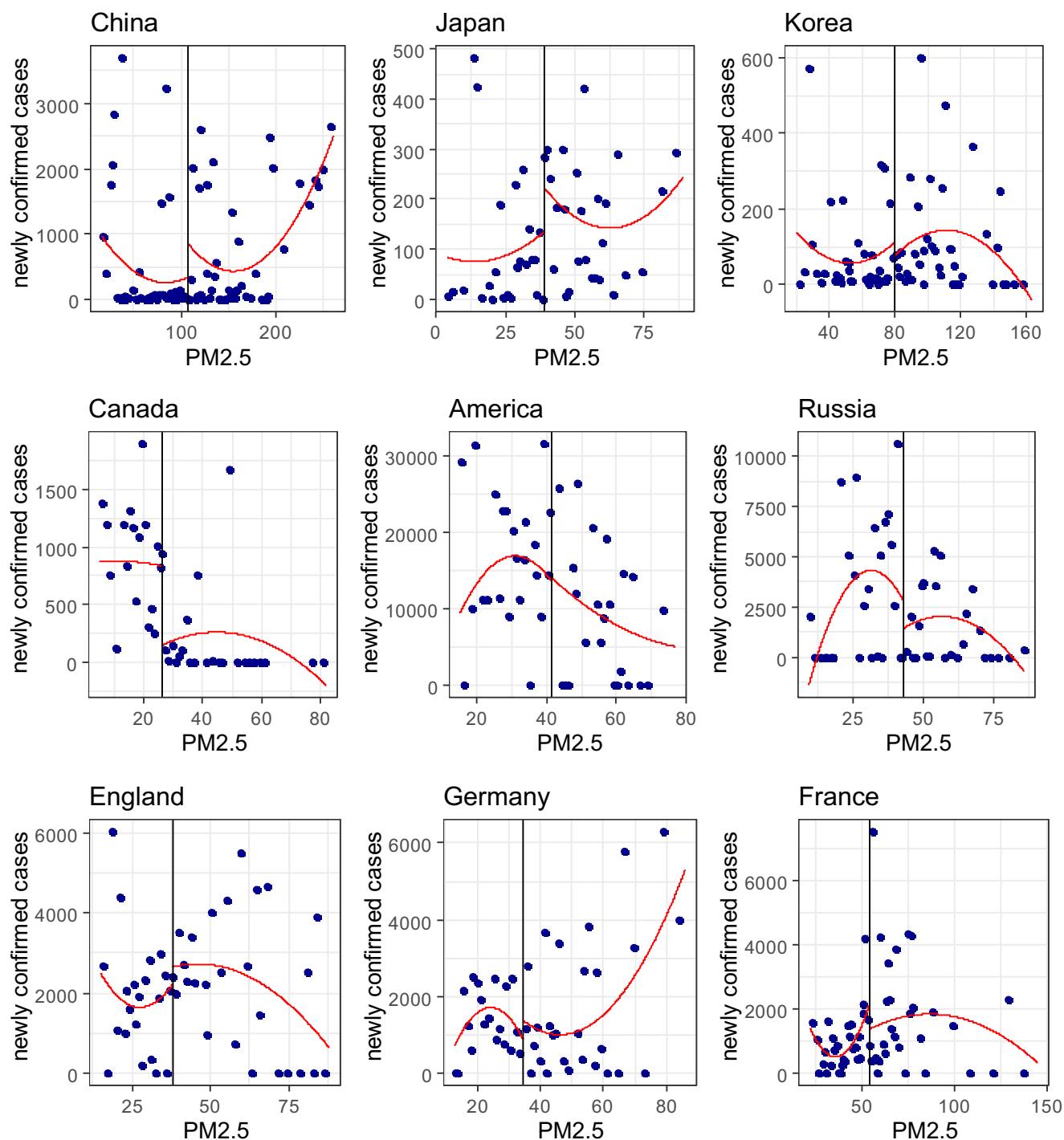


Fig. 4 Non-linear relationship between air pollution and COVID-19 infection (take $PM_{2.5}$ as an example). Note: The black vertical line in the middle is the cut point, and the red curve on the left and right sides of the cut point is the non-linear fitting line

Relationship between AP and COVID-19 from the municipal level in China

In order to analyze the impact of air pollution on COVID-19 infection within the country, we conducted threshold tests on 6 different pollutants separately at the municipal level in China. Figure 5 presents the results of single threshold

parameter from AP to COVID-19 by the method of likelihood ratio (such as $PM_{2.5}$, PM_{10} , SO_2 , CO , NO_2 , O_3 , etc.). Obviously, air pollution has a threshold effect on COVID-19 infection in example analysis.

Table 7 presents the results for threshold relationship between AP and COVID-19. First, the threshold value of $PM_{2.5}$ indicates different size from the municipal level (it is 26) and

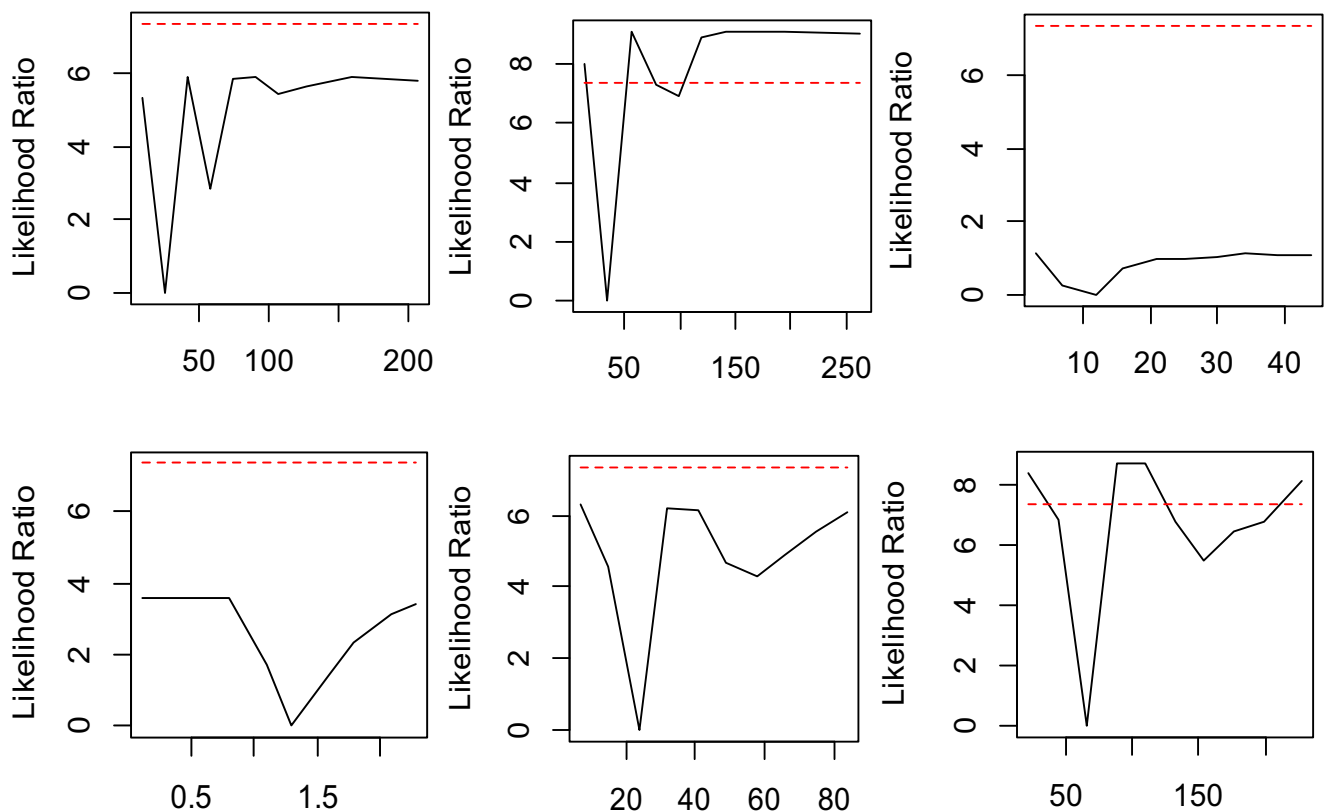


Fig. 5 Single threshold parameter from AP to COVID-19. Note: The first row from left to right is PM_{2.5}, PM₁₀, and SO₂; the second row from left to right is CO, NO₂, and O₃; and the red dotted line is the optimal likelihood ratio given by bootstrap replication

Table 7 Threshold relationship from air pollution to COVID-19 infection

Variables	Dependent variable: newly confirmed cases of COVID-19 infection					
	PM _{2.5}	PM ₁₀	SO ₂	CO	NO ₂	O ₃
Threshold	26	35	12	1.3	24	66
α_1	1.594* (0.764)	1.895** (0.676)	1.766** (0.742)	109.414** (38.322)	1.618* (0.815)	0.841* (0.414)
α_2	-0.045* (0.022)	0.106 (0.232)	-0.062* (0.024)	66.029* (33.008)	-0.331* (0.165)	-0.166*** (0.048)
PM _{2.5}		-0.095 (0.326)	-0.092 (0.326)	-0.037 (0.327)	-0.116 (0.326)	-0.076 (0.326)
PM ₁₀	0.048 (0.230)		-0.021 (0.228)	-0.019 (0.228)	-0.029 (0.228)	-0.006 (0.228)
SO ₂	-0.491 (1.451)	-0.677 (1.449)		-0.508 (1.453)	-0.616 (1.449)	-0.358 (1.453)
CO	81.224* (33.257)	75.777* (33.082)	75.025* (33.162)		76.082* (33.106)	65.137* (33.180)
NO ₂	-0.937* (0.423)	-0.843 (0.526)	-1.080* (0.523)	-1.167* (0.526)		-1.139 (0.722)
O ₃	-0.404** (0.159)	-0.345* (0.160)	-0.424* (0.159)	-0.427* (0.159)	-0.410** (0.159)	
LR test	5.912	9.105	1.114	3.577	6.324	8.812

Note: 6 air pollution indicators for threshold variable, respectively; the standard errors are shown in brackets
 * $P < 0.1$; ** $P < 0.05$; *** $P < 0.01$; **** $P < 0.001$

national level in China (108.19); we can conduct the reason from three aspects: (i) span of the time is inconsistent; the national level is from 21 Jan to 20 May, while the municipal level is 28 Jan to 31 May; the disturbance effects in the time dimension may exist; (ii) objects are different; China as a whole from the average degree at the national level, while it reflects the differences of 27 individuals at the municipal level, the disturbance effects in spatial dimension may exist; and (iii) the measurement standards are different; they are measured by America in the national level, while they are measured by China in the municipal level; measurement errors may exist.

Second, coupled with the conclusions before, $PM_{2.5}$ has a significant non-linear effect on COVID-19 infection no matter at the national level or municipal level. It is manifested that the increase in $PM_{2.5}$ will aggravate the speed of COVID-19 infection, but this relationship has the same effect and will weaken with the emergence of the threshold. Moreover, the most sensitive ones in China are still $PM_{2.5}$ and CO when air pollution is taken into account on COVID-19 infection, among the 6 indicators of air pollution. Hence, according to air pollution on COVID-19 infection with or without internal differences, the policy implications may differ in China.

Third, when PM_{10} , SO_2 , CO, NO_2 , and O_3 are used as threshold variables, their non-linear threshold effects on COVID-19 infection are also significant. Specifically, different from the other 5 indicators, the positive effect of CO on COVID-19 infection is the strongest with a coefficient of 109.414 before the threshold (13 mg/m) is reached; moreover, the coefficient of CO on COVID-19 infection reduced to 66.029 after the threshold, and still aggravated the spread of COVID-19, but the spread will slow down. For PM_{10} , its increase intensifies COVID-19 infection only before the threshold ($35 \mu\text{g}/\text{m}^3$) is taken. For the remaining 4 indicators of air pollution, they are positive before the threshold and negative after the threshold on COVID-19 infection, but also have differences; among them, the positive effect of O_3 on COVID-19 infection is the weakest (0.814) before the threshold ($66 \mu\text{g}/\text{m}^3$) is reached. Overall, the threshold effect of air pollution on COVID-19 infection is significant, and the increase in air pollution will not continue to accelerate the spread of COVID-19 infection. Therefore, the policy of air pollution indicators should be further adjusted according to the thresholds, which has achieved the purpose of effectively controlling the external environment of COVID-19 infection in China.

Table 8 presents the results for the deviation of air pollution indicators to thresholds in China at the municipal level. Among the 6 indicators, only the threshold of PM_{10} on COVID-19 infection (1.3) is greater than the mean value (0.73); this shows that the aggravating effect of CO on COVID-19 has a certain mitigating effect in China from the provincial aspect, and it is difficult to control the spread of

Table 8 Deviation of air pollution indicators to thresholds in China

	$PM_{2.5}$	PM_{10}	SO_2	CO	NO_2	O_3
Threshold	26	35	12	1.3	24	66
Mean	38.80	67.37	9.77	0.73	31.69	101.50
Min	3	0	2	0.1	4	7
Max	906	644	50	2.8	119	283

COVID-19 by reducing CO emissions. For the remaining 5 indicators, thresholds of air pollution are less than mean values in all; the effects of $PM_{2.5}$, PM_{10} , SO_2 , NO_2 , and O_3 on COVID-19 infection are in the accelerated stage before the attached thresholds, which indicates that the relative external environment that slows down the spread of COVID-19 infection can be provided by controlling the emissions of these 5 air pollution indicators.

Discussion and conclusion

In this study, the concentrations of air pollution on COVID-19 infection were used in order to explain the spatial variation of newly confirmed cases in 9 countries at the national level from three regions (such as Asia, North America, and Europe), and explain the internal difference in 27 provinces at the municipal level from China. The data at the national level and municipal level shows three main air pollution hotspots on COVID-19 infection over countries and provinces: (i) there is a threshold non-linear effect of air pollution on the spread of COVID-19 infection; the increase of air pollution will greatly intensify the newly confirmed cases of COVID-19 before the threshold is reached, but this aggravating effect will be weakened after the threshold; (ii) the impact intensity of the 6 air pollution indicators on COVID-19 is significantly different; $PM_{2.5}$ is the key air pollution indicator for preventing and controlling COVID-19 at the national level, and PM_{10} , $PM_{2.5}$, and SO_2 are the key points at the municipal level; (iii) the impact of different air pollution indicators on COVID-19 is different not only between countries but also within countries.

According to these results, high air pollution concentration accompanied by a decline in air quality causes an increase in newly confirmed cases of COVID-19 infection. Considering the complex and volatile nature of the geographical location at the national and municipal levels, this topographical structure combined with reversed atmospheric conditions can prevent the spread of air pollutants (such as $PM_{2.5}$, PM_{10} , SO_2 , CO, NO_2 , O_3 , etc.), which can lead to a high incidence of respiratory problems and inflammation in the local population. Long-term exposure could be an important reason for newly confirmed cases of high COVID-19 infection in these regions. As earlier studies have shown, the six indicators of air

pollution can hardly reach the threshold during the sample period, whether it is at the national or municipal level. Therefore, it is believed that there is a very sensitive relationship between serious air pollution and the newly confirmed cases of highly COVID-19 infection. In other words, the exposure to air pollution causes inflammation in the lungs, by $PM_{2.5}$ and PM_{10} for European countries and $PM_{2.5}$ and O_3 for North American countries in particular; it is now necessary to examine whether the presence of an initial inflammatory condition is related to the response of the immune system to the coronavirus. Hence, increasing air pollution means poisoning our own body and when it experiences a chronic respiratory stress, its ability to defend itself from infections is limited.

In addition, based on the results of this study, more studies should be conducted which focus on additional factors such as terrain structure and composition of main environmental pollutants along with the different threshold effects of pre-exposure to $PM_{2.5}$, PM_{10} , SO_2 , CO , NO_2 , and O_3 on hypercytokinemia in order to verify their impact on newly confirmed cases due to the COVID-19 pandemic.

Author contributions Qiang Liu is responsible for the overall inspection and article frame design (30%); Xu Shengxia is responsible for the model construction, program implementation, and thesis writing (40%); Lu Xiaoli is responsible for the data collection and article content verification (30%).

Data availability Sources for all the data in the paper have been given and these data can be viewed on the corresponding website, so no supplementary information is submitted.

Declarations

Ethics approval

· The manuscript was not submitted to more than one journal for simultaneous consideration.

· The submitted work is original and has not been published elsewhere in any form or language (partially or in full).

· This study has not been split up into several parts to increase the quantity of submissions and submitted to various journals or to one journal over time.

· No data, text, or theories by others are presented as if they were the authors' own ("plagiarism"). Proper acknowledgements to other works are given (this includes material that is closely copied (near verbatim), summarized, and/or paraphrased), quotation marks (to indicate words taken from another source) are used for verbatim copying of material, and permissions secured for material that is copyrighted.

Consent to participate Yes.

Consent for publication Yes.

Competing interests The authors declare not competing interests.

Foundation National Office for Philosophy and Social Science (20BTJ027); National Natural Science Foundation of China (11901508).

References

- Buhlmann P, Meneil AJ (2002) An algorithm for nonparametric GARCH modelling. *Computational Statistics and Data Analysis* 4(10):665–683
- Cai Q.-C., Lu J., Xu Q.-F., Guo Q., Xu D.-Z., Sun Q.-W., Yang H., Zhao G.-M., Jiang Q.-W. (2007) Influence of meteorological factors and air pollution on the outbreak of severe acute respiratory syndrome. *Public Health*. 121, 258–265. <https://doi.org/10.1016/j.puhe.2006.09.02>
- Durusu-Ciftci D., Soytaş U., Nazlıoğlu S., 2020. Financial development and energy consumption in emerging markets: smooth structural shifts and causal linkages. *Energy economics*. 104729.
- Daniele Fattorini, Francesco Regoli, 2020. Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. *air pollution*. <https://doi.org/10.1016/j.envpol.2020.114732>
- Federico A, BugniaIvan AC (2020) Testing continuity of a density via g-order statistics in the regression discontinuity design. *Journal of Econometrics*. <https://doi.org/10.1016/j.jeconom.2020.02.004>
- Jiang F., Deng L., Zhang L., Cai Y., Cheung C.W., Xia Z., 2020. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). *Journal of general internal medicine*.
- Kraemer M.U., Yang C.-H., Gutierrez B., Wu C.-H., Klein B., Pigott D.M., du Plessis L., Faria N.R., Li R., Hanage W.P., 2020. The effect of human mobility and control measures on the COVID-19 APidemic in China. *Science*. <https://doi.org/10.1126/science.abb4218>
- Lee DS, Card D (2008) Regression discontinuity inference with specification error. *Journal of Econometrics* 142(2):655–674
- Lee DS, Thomas L (2010) Regression discontinuity designs in economics. *Journal of Economic Literature* 48(6):281–355
- Li H, Li KY, Wu J (2018) Local political chief turnover and economic growth: evidence from China. Available at SSRN 3:1–36
- Monks PS, Archibald AT, Colette A, Cooper O, Coyle M, Derwent R et al (2015) Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. *Atmospheric chemistry and physics*. 15:8889–8973
- Nazlıoğlu S, Gormuş NA, Soytaş U (2016) Oil prices and real estate investment trusts (REITs): gradual-shift causality and volatility transmission analysis. *Energy economics*. 60:168–175
- Ogen Y (2020) Assessing nitrogen dioxide (NO₂) levels as a contributing factor to the coronavirus (COVID-19) fatality rate. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.138605>
- Rodriguez-Morales A.J., Bonilla-Aldana D.K., Tiwari R., Sah R., Rabaan A.A., Dhama K., 2020. COVID-19, an emerging coronavirus infection: current scenario and recent developments-an overview. *Journal of pure and applied microbiology*. 9.
- Saaha S, Li B, Oiamo Tor H, Burnett Richard T, Scott W, Michael J et al (2020) Association between road traffic noise and incidence of diabetes mellitus and hypertension in Toronto, Canada: a population-based cohort study. *Journal of the American heart association*. 9: e013021
- Tang YG, Chen Y, Shi SB (2019) Valuing the urban Hukou in China: evidence from a regression discontinuity design for housing price. *Journal of Development Economics*. <https://doi.org/10.1016/j.jdeveco.2019.102381>
- Thistlethwaite DL, Campbell DT (1960) Regression-discontinuity analysis: an alternative to the ex post facto experiment. *Journal of Educational Psychology* 51(6):309–317
- Vu HD (2020) The threshold of absorptive capacity: the case of Vietnamese manufacturing firms. *International Economics* 163(10):44–57
- Wang T., Xue L.K., Brimblecombe P., Lam Y.F., Li L., Zhang L., 2016. Ozone pollution in China: a review of concentrations,

- meteorological influences, chemical precursors, and effects. *Science of the total environment*. S0048969716322471.
- Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y (2020) Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *The Journal of the American Medical Association*. <https://doi.org/10.1001/jama.2020.1585>
- World Health Organization. 2020. Coronavirus disease 2019 (COVID-19) Situation RAPort-51 (<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-rAPorts>).
- Xie J., Zhu Y., 2020. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Science of the Total Environment*. 138201. <https://doi.org/10.1016/j.scitotenv.2020.138201>
- Xu Q., Li X., Wang S., Wang C., Huang F., Gao Q., Wu L., Tao L., Guo J., Wang W., 2016. Fine particulate air pollution and hospital emergency room visits for respiratory disease in urban areas in Beijing, China, in 2013. *PLoS One*. 11. <https://doi.org/10.1371/journal.pone.0153099>
- Yang Z, Hao J, Huang S, Yang W, Zhu Z, Tian L, Lu Y, Xiang H, Liu S (2020) Acute effects of air pollution on the incidence of hand, foot, and mouth disease in Wuhan, China. *Atmospheric Environment*. 225:117358. <https://doi.org/10.1016/j.atmosenv.2020.117358>
- Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine*. 367(19): 1814–1820
- Zhang Z, Xue T, Jin X (2020) Effects of meteorological conditions and air pollution on COVID-19 transmission: evidence from 219 Chinese cities. *Science of the total environment*. <https://doi.org/10.1016/j.scitotenv.2020.140244>
- Zhu YJ, Xie JG, Huang FM, Cao LQ (2020) Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. *Science of the Total Environment*. <https://doi.org/10.1016/j.scitotenv.2020.138704>
- Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.