Do Weather Conditions Affect COVID-19 Epidemic? Evidence Based on Panel Data of Prefecture-level Administrative Regions in China

Ruofei Lin
Tongji University School of Economics and Management
Xiaoli Wang
Tongji University School of Economics and Management
Junpei Huang (1830263@tongji.edu.cn)
Tongji University School of Economics and Management https://orcid.org/0000-0002-8357-0273

Research

Keywords: the COVID-19, weather conditions, causal inference, multiple linear regression, heterogeneity analysis, moderating effects

Posted Date: July 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-645202/v1

Do Weather Conditions Affect COVID-19 Epidemic？

Evidence Based on Panel Data of Prefecture-level Administrative Regions in China

Ruofei Lin[1a], Xiaoli Wang[1a],and Junpei Huang[*a]

a.School of Economics and Management, Tongji University, China

***Correspondence:** Junpei Huang, School of Economics and Management, Tongji University, China. Email: 1830263@tongji.edu.cn; Tel:+8613524597646; postal address: Building A, Tongji Building, 1500 Siping Road, Shanghai, China.

Declaration of interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1 Equal contribution.

Do Weather Conditions Affect COVID-19 Epidemic？ Evidence based on panel data of prefecture-level administrative regions in China

Abstract:

Background: Similar to other infectious diseases, weather conditions may affect the COVID-19 epidemic through changes to transmission dynamics, host susceptibility, and virus survival in the environment. It's critical to explore the relationship between weather variables and the spread of the COVID-19 for understanding seasonality and the possibility of future outbreaks, developing early warning systems, infection control methods, and public health measures. However, the influence of weather change on COVID-19 edidemic is still an emerging research field, and there is still relatively limited literature available.

Objectives: Our study aims to explore the causal relationship between weather conditions and COVID-19 epidemic, the regional heterogeneity of the influence of weathner conditions in east-middle-west and coastal-inland, the moderating effect of diurnal temperature difference, public health measures, and public opinion on the influence of weather conditions on the epidemic to investigate the effects of these factors on the intensity of weather conditions.

Methods: First, we theoretically explain the influence mechanism of weather conditions on the epidemic based on the epidemiological triangle model. Then, we collect COVID-19-related prefecture-daily panel data in mainland China from January 1, 2020, to February 19, apply two-way fixed effect model of multiple linear regression, and also take into account other influencing factors such as population movement, public health interventions of the local government, economic and social conditions, to explore the causal relationship between weather conditions and the COVID-19 epidemic.

Results: It is found that first, there is a conditional negative linear relationship between the weather conditions and the epidemic. When the average temperature is greater than -7℃, there is a significant negative causal relationship between the average temperature and the growth rate of the confirmed cases. Similarly, when the relative humidity is greater than 46%, the increase in the relative humidity significantly contain the epidemic. However, when the average temperature is less than -7℃ or the relative humidity is less than 46%, the effect of weather conditions is not significant. Further, from the perspective of weather conditions, prefecture-level administrative regions such as Chifeng, Zhangjiakou, and Ulanqab are more conducive to the outbreak of the epidemic in winter. Then, weather conditions have a greater influence in the east than in the middle and western regions, and it is better in coastal region than in the inland. Finally, increasing diurnal temperature differences will improve the impact of weather conditions on the confirmed cases. In dry and cold regions, higher diurnal temperature differences will increase the risk of spread of the disease; Strict public health measures and good public opinion can mitigate the adverse effects of cold and dry weather on the spread of the epidemic.

Discussion: In future research, it can adopt more detailed investigation methods. Under the legal framework of privacy protection, questionnaire surveys can be carried out with patients' consent to draw more accurate conclusions. At the same time, in terms of the mechanism of the role of weather variables, more in-depth interdisciplinary cooperation with epidemiologists is needed to study the specific impact of weather conditions on the survivability of the COVID-19 virus and the immunity of susceptible populations to obtain a clearer picture and compelling conclusions.

Keywords: the COVID-19; weather conditions; causal inference; multiple linear regression; heterogeneity analysis; moderating effects

1. Introduction

In December 2019, the first case of Corona Virus Disease 2019 (COVID-19) appeared in Wuhan, Hubei Province, China, and the same cases were subsequently found in other provinces of China. On January 23, the city of Wuhan went into lockdown, and from January 23 to 30, various provinces in China initiated *first-level response to major public health emergencies* [1]. Meanwhile, within two months, the COVID-19 had spread globally, and on March 11, 2020, the World Health Organization (WHO) declared the disease a global pandemic (Allam, 2020). As of April 18, 2021, the cumulative number of confirmed COVID-19 cases worldwide has exceeded 140.332 million, and the cumulative number of deaths has exceeded 3.004 million (WHO, 2021). The rapid spread of the COVID-19 poses a colossal challenge not only to human health but also to social and economic development.

The COVID-19 is a viral respiratory illness caused by the beta-coronavirus SARS-CoV-2, and it belongs to the same coronavirus family as infectious diseases such as severe acute respiratory syndrome (SARS) and Middle East Respiratory Syndrome (MERS), and other infectious diseases which spread rapidly through aerosolized droplets and virus-contaminated hands and surfaces (Sohrabi et al., 2020). The primary mechanism of action of SARS-CoV-2 is binding to the angiotensin-converting enzyme 2 (ACE2) receptors that exist on the surface of biological membranes predominantly found in the cells of the heart, lung, arteries, intestine, and renal tissues (Lan et al., 2020). After being infected with SARS-CoV-2, the incubation period may vary from 2-14 days before the onset of symptoms, mainly affecting the lower respiratory system, which the

1 China's "National Emergency Plan for Public Health Emergencies" divides public health emergencies into four levels: particularly serious (level I), major (level II), large (level III), and general (level IV). Among them, the provincial government is responsible for emergency response and preventive control measures for public health emergencies of level I and level II.

clinical manifestations are dry cough, fever, and fatigue (Keni et al., 2020). The severity of symptoms varies by individuals, ranging from asymptomatic manifestations to severe life-threatening symptoms, including myocardial dysfunction and acute respiratory failure (Lee et al., 2020). Those most at risk of severe COVID-19 manifestations are older individuals and those with pre-existing conditions and multi-morbidities, particularly cardiovascular disease or diabetes (Clark et al., 2020).

Similar to other infectious diseases, weather conditions may affect the COVID-19 epidemic through changes to transmission dynamics, host susceptibility, and virus survival in the environment. It's critical to explore the relationship between weather variables and the spread of the COVID-19 for understanding seasonality and the possibility of future outbreaks, developing early warning systems, infection control methods, and public health measures. Existing epidemiological literature shows that seasonal and weather changes affect the spread of respiratory pathogens. Existing literature studies the influence of weather changes on the transmission of respiratory pathogens such as influenza, SARS-COV, and MERS-COV. Weather changes have the potential to facilitate the emergence of new viruses and affect epidemic transmission, morbidity, and mortality (Van Doremalen et al., 2013; Barreca & Shimshack, 2012; Sobral et al., 2020). The COVID-19 is epidemiologically similar to influenza viruses in that both are highly spread through the respiratory tract and cause acute infections (Cobey, 2020). However, since the COVID-19 virus is very different from known viruses in terms of pathogenicity and transmission, the influence of weather changes on the COVID-19 epidemic is still an emerging research field, and the existing literature is still relatively limited. Moreover, there are still differences in research conclusions, and it is indispensable to carry out further related research.

The main contributions of this article are as follows. First, in this article, daily confirmed COVID-19 cases reported by prefecture-level administrative regions[2] of mainland China are used as the research sample, which helps to reach a more accurate conclusion. Regarding the influence of weather on the spread of COVID-19, most of the existing literature uses counrty-level samples (Iqbal et al., 2020; Huang et al., 2020) or limited region-level samples (Briz-Redón & Serrano-Aroca, 2020; Thu et al., 2020). However, the country-level sample fails to capture regional differences in weather in countries with large areas and uneven population distribution such as the United States, China and Brazil.

Second, compared with the existing literature that mainly uses approaches such as Spearman's rank correlation and time series analysis (Alkhowailed et al. 2020; Bashir et al. 2020; Menebo, 2020), in this paper, we carry out multiple linear regression using panel data which is more conducive to identifying the causal relationship between variables (Angrist & Pischke, 2008). In addition, we compare and analyze whether the relationship between weather variables and the COVID-19 epidemic is a nonlinear or conditional linear relationship, and further explore the regional heterogeneity and the moderating effects of diurnal temperature variation, public health measures, and social public opinion.

Third, different from the existing literature, we also consider the diurnal temperature difference while considering the temperature and humidity variables. So far, relevant research focuses on average temperature, minimum temperature, maximum temperature, relative humidity, and absolute humidity on the spread of COVID-19 (Huang et al., 2020; Pani et al., 2020). A few papers

[2] The prefecture-level administrative region is the second-level administrative region of China's administrative divisions. It is governed by the provincial administrative region, including 17 prefectures, 30 autonomous prefectures, 283 prefecture-level cities and three leagues. Prefecture-level cities (PLC) include both cities and counties, which cover rural areas with a vast land area.

consider variables such as wind speed and precipitation simultaneously, but there is limited evidence that these two variables are related to the spread of COVID-19 (McClymont & Hu, 2021). While focusing on the temperature and humidity variables, this paper for the first time considers the impact of diurnal temperature differences on the spread of COVID-19. The diurnal temperature difference is an essential factor affecting immunity and the incidence of infectious diseases (Epstein, 2010; Cheng et al., 2014), and it is also an important variable reflecting large dimensions cross north and south, significant differences in elevation between east, middle, and west, and geographic differences in coastal and inland locations of China, but ignored in the existing literature.

Fourth, compared with most of the existing literature, in addition to weather conditions, we also consider control variables such as human behavior patterns, public health measures, economic and social conditions, which is helpful to overcome the biased estimation caused by omitted variables. The principle of the spread of COVID-19 is complicated; in addition to weather conditions, it also involves some important factors. All other potential confounding factors must be controlled to analyze the role of weather in the spread of COVID-19 more effectively. Regarding human behavior patterns, it is related to population concentration, transportation convenience, and population mobility. In this paper, we refer to the method proposed by Brockmann and Helbing (2013) to calculate and incorporate the control variable "effective distance". So far, the role of socioeconomic conditions in the spread of COVID-19 is unclear. This article mainly considers factors such as economic development level as well as health and medical conditions. We also consider the influence of the incubation period which is ignored in the previous quantitative studies on the relationship between weather conditions and the spread of the virus. By incorporating some key variables ignored in the previous literature into the model in this article, the role of weather in the spread of the COVID-19 can be estimated more accurately.

Fifth, we pay more attention to the influence of the Chinese government's stringent public health measures on the estimated results than the existing literature. Different from other countries, Chinese governments adopted the most comprehensive and strict measures in the world to prevent and control the COVID-19 epidemic. Wuhan is not the only city that went into lockdown; many local governments of other prefecture-level cities also took such a measure, however, most of the existing literature does not consider the influence of these measures on analysis results. In the relevant statistical analysis using China as the research sample is required to carefully consider the influence of the public health measures taken by the Chinese government. This article collates and evaluates public health measures taken by each prefecture-level administrative region in China, including but not limited to school closures, travel restrictions, community control, social distancing, quarantine, isolation, close contact tracking. In addition, we use data mining technology to collect big data, such as the *Baidu search index*[3] and *Baidu migration data*[4].

2. Literature Review

The relationship between climate, weather, and infectious disease epidemics has attracted people's attention since 2500 years ago when Hippocrates and his followers described the relationship between seasonal changes and the spread of infectious diseases (Fisman, 2007; Lloyd et al., 1983). Hippocratic treatise, *Airs, Waters, Places* describe the influence of the environment and seasons on the constitution and instructed physicians to observe the health of a community concerning sun

[3] Baidu Search Index website: http://index.baidu.com/v2/index.html#/。
[4] Baidu Migration website: http://qianxi.baidu.com/。

exposure, soil, elevation, climate, and geography (Miller, 1962). During the 16th and 18th centuries, interest in the effects of climate on health arose from the ability to measure environmental conditions with new instruments. For example, in the United States, both Thomas Jefferson and Noah Webster collected information about weather and disease (National Research Council, 2001). In the mid to late 19th century and most of the 20th century, people were no longer interested in the effects of seasons and climate due to the emergence of bacterial theory and the development of microbiology, and turned their attention to elucidating the risk factors for infectious diseases associated with host and pathogen. From the end of the 20th century to the present, attention to changes in climate and weather renewed interest in understanding the impact of environment, climate, and weather on the incidence of infectious diseases and other health-related diseases (Watts et al., 2017).

Epidemiological studies find that both the mortality and virulence of the Spanish flu from 1918-1919 are very high, which is related to low temperature and precipitation enhancement (More et al., 2020). Avian influenza A virus is the cause of Spanish flu (Taubenberger, 2006), which first appeared in the autumn and winter of 1917 and spread to Europe, North America, and Asia through troop mobilization and deployment during World War I (Taubenberger & Morens, 2006). Since the 1918 influenza pandemic, influenza A and B strains had continued to spread around the world. There are different seasonal outbreak patterns in different weather regions and in recent history, the emergence of new viruses such as the 2009 H1N1 (swine flu) pandemic, 2003 SARS, and 2012 MERS shows that they are related to the weather. Seasonal influenza outbreaks have prominent seasonal characteristics, and the peak of the annual outbreak is consistent with winter and related cold and dry weather patterns (Park et al., 2020). Seasonal outbreaks of subtropical and tropical weather show different patterns, usually with persistent low-level cases in the community, with multiple outbreaks throughout the year, most commonly in the shoulder season from autumn to spring (Tamerius et al., 2011; Shaman et al., 2009). The severity of these seasonal outbreaks varies, and weather is related to these changes since weather changes are conducive to increasing transmission or leading to increased morbidity and mortality (Liu & Zhang et al., 2020). When a cold winter is followed by a mild winter, with the weather changes, the weather changes more and more, severe and early seasonal outbreaks of influenza will occur (Roussel et al., 2016; Towers et al., 2013). As far as SARS is concerned, meteorological factors seem to affect the spread of the virus. Tan et al. (2005) find that there is a significant correlation between SARS cases and the environmental temperature 7 days before the attack, and the optimal environmental temperature for SARS cases is 16 C to 28 C. Lin et al. (2006) find that the incidence of SARS at lower temperatures is 18 times higher than that at higher temperatures and respiratory diseases are common in colder environments. At higher weather temperatures, the virulence of the pathogen will worsen because they may not withstand environmental changes. Gardner et al. (2019), based on the cases of MERS in Saudi Arabia, find that MERS is more likely to occur in relatively cold and dry conditions, similar to the seasonal patterns of other respiratory diseases in temperate regions. Altamimi & Ahmed (2020), based on the case of MERS in Riyadh, find that the incidence of MERS is affected by weather conditions, and it shouws an upward trend from April to August. High temperature and low relative humidity are the reasons for the increase in MERS cases. Although there are some achievements regarding the relationship between weather conditions and emerging infectious diseases, the relevant literature is still quite limited (Paraskevis et al., 2020).

Research on the influence of weather conditions on the spread of COVID-19 is in its infancy, and there are few related papers. Most of the literature focuses on temperature and humidity while a few involve other weather conditions such as wind speed and precipitation. However, the evidence for the correlation between wind speed and precipitation and the spread of COVID-19 is limited (McClymont & Hu, 2021). Therefore, we mainly focus on two weather conditions, that is temperature and humidity.

Regarding the relationship between temperature and the COVID-19 epidemic, the existing literature explores the role of minimum temperature, maximum temperature, or average temperature as variables for weather conditions, and the research conclusions are entirely different. Many works of literature conclude that temperature negatively correlated with the spread of COVID-19, that is the higher the temperature, the fewer people infected. Meyer et al. (2020) study samples from 100 countries worldwide and find that when the temperature rises above −15°C, there is a significant negative correlation between daily temperature and daily global cases. Shi et al. (2020) take samples of COVID-19 cases from 30 provincial administrative regions in China and find that the incidence rate varies with temperature, where the higher the temperature, the lower the incidence of the COVID-19; conversely, the lower the temperature regions, the more people will be infected. Liu & Zhou et al. (2020) use 30 provincial capital cities in China as the research sample, and they find the average temperature significantly negatively correlated with the number of the COVID-19 cases. Nevels et al. (2021) believe the temperature was negatively correlated with the transmission rate of COVID-19 in the early stage of the outbreak in Wuhan. However, some literature concludes that temperature positively correlated with the spread of COVID-19. For example, Iqbal et al. (2020) take 210 countries and territories worldwide as a sample and conclude that the average temperature and daylight hours have shown a positive association towards the spread rate of COVID-19. Islam et al. (2021) take cases from 206 countries/regions as samples and find that the COVID-19 cases positively correlated with the 14-day lag temperature. Paniet al. (2020) conduct a study on cases in Singapore and conclude that the average temperature and minimum temperature significantly positively correlated with the number of both new cases and the total cases. There is also some literature suggesting that temperature does not correlate with the spread of COVID-19 or the correlation relationship is uncertain. In the research of Jahangiri et al. (2020) on Iran and Briz-Redón & Serrano-Aroca (2020) on Spain, they find no correlation between temperature and the spread of COVID-19. Hossain et al. (2021) study cases in South Asian countries such as Afghanistan, Bangladesh, India, Pakistan, and Sri Lanka and conclude that the influence of temperature on the COVID-19 epidemic is different in different countries with some positive and other negative correlations.

Regarding the relationship between humidity and the spread of COVID-19, the existing literature studies the influence of absolute humidity or relative humidity as the variable of weather conditions, and the research conclusions are also entirely different. Most literature reports a negative correlation between humidity and the spread of COVID-19. For example, Wu et al. (2020) collect data from 166 countries other than China and find a negative correlation between relative humidity and the number of new cases and deaths per day, where a 1% increase in relative humidity leads to a 0.85% decrease of the new cases per day and a 0.67% decrease of new deaths. Qi et al. (2020) study cases in 30 provincial administrative regions in China and conclude that relative humidity significantly negatively correlated with the number of cases. When the temperature is between 5.04°C and 8.2°C, for every 1% increase in relative humidity, daily cases will decrease by 11-22%. Zhu et al. (2020) collect daily new cases in 8 hard-hit areas in 4 countries in South America and find that absolute humidity significantly negatively correlated with daily confirmed cases. While some other literature believes that humidity positively correlated with the spread of the COVID-19. For example, Chien and Chen (2020) study 50 counties in the United States with the highest cumulative confirmed cases and find that relative humidity has a significant positive correlation with the cumulative cases. Alkhowailed et al. (2020) report a weak positive correlation between average relative humidity and new cases in Saudi Arabia. Other literature suggests that humidity is not related to the spread of COVID-19 or cannot be determined to be related. Meyer et al. (2020) collect national data on the COVID-19 cases as of March 17, 2020, and find no correlation between relative humidity and the number of the COVID-19 cases. Pan et al. (2021) collect cases from 202 locations in 8 countries and conclude that weather conditions such as temperature, relative humidity, wind speed, and ultraviolet rays significantly did not correlate with the COVID-19 cases. Pahuja et al. (2021) study the number of cases in New Delhi, India, and did not observe a correlation between the number of cases and humidity or wind speed neither.

Regarding the relationship between weather conditions such as temperature and humidity and the COVID-19 epidemic, the different conclusions reached by existing literature mainly due to two aspects. The first reason is the problem of sample selection. Most of the existing literature selects national-level samples or limited large region and city-level samples. As Polgreen & Polgreen (2018) point out, to accurately verify the relationship between weather conditions and infectious diseases, it is needed not only the geographic area where the cases occurred but also a control group without infectious diseases. Selecting geographic areas with cases as samples for statistical analysis is prone to sample selection bias (Chen and Astebro, 2001). Moreover, national sample fails to capture regional differences in weather conditions between countries with large areas and uneven population distribution, such as the United States, China, and Brazil.

The second reason is the problem of statistical analysis methods. In our search of 28 studies using statistical models up to March 2021, most of them used Spearman's rank correlation, time series analysis. Pearson's correlation, generalized linear model (GLM), Generalized Linear Mixed Model (GLMM), and Generalized Additive Model (GAM) are used once. So far, these methods are more common in research projects on the relationship between weather and disease outside the field of infectious diseases, while there are few pieces of research in infectious diseases, and the methods are not yet mature (Polgreen & Polgreen, 2018). First, the limitations of these methods are apparent. For example, the time series regression model is built on the premise that the data satisfy linearity and staticity, and relies on a large amount of uninterrupted time series data. However, the COVID-19 case data is not necessarily linear, and the data is difficult to satisfy the requirements of uninterrupted time series. The applicable conditions of Spearman's correlation are that there are no repeated values in the data, and the two variables have a monotonic relationship, and the Spearman model calculates the grade correlation coefficient, based on the rank, which discards some vital information of the original data (Owen & Anil, 2009). The application of Pearson correlation has strict requirements on data, that is, normal distribution, but the actual situation of the COVID-19 cases is challenging to meet this requirement. Second, there is Omitted Variable Bias in most of the literature. The spread of COVID-19 is closely related to population movement and public health interventions. Studies show that public health interventions have a much more significant impact than the weather and climate variables (Oliveiros et al., 2020). Paraskevis et al. (2020) conclude that the seasonality of COVID-19 is very different from the common cold coronavirus or influenza, and in the absence of public health measures, climatic conditions cannot alleviate the spread of COVID-19. Most of the existing research does not consider the influence of population mobility or public health intervention. However, the intervention intensity of public health measures varies significantly from country to country, and most of the literature lacks a relatively accurate assessment of this difference. The omission of this critical variable can lead to a significant bias in the estimation results (Angrist & Pischke, 2008).

3. Conceptual Framework

The existing literature on the relationship between weather conditions and the spread of COVID-19 is preliminary, and the conclusions are opposite. However, based on the epidemiologic triangle [5] and related literature on respiratory viruses similar to the COVID-19, we believe that weather conditions such as temperature and humidity may impact the spread of COVID-19. First, weather conditions determine the viability and persistence of the virus in the air and on the surface (Aboubakr et al., 2021). Generally speaking, the continuous increase of COVID-19 and similar viruses is related to low temperature and low relative humidity. Chan et al., 2011 report that SARS can survive for more than 5 days on smooth surfaces at temperatures of 22-25°C and relative humidity of 40-50%. Some literature also proves that temperature and humidity are known factors affecting the survival of SARS, MERS, and influenza (Otter et al., 2016). A recent study shows that the COVID-19 can survive on glass, stainless steel, and paper currency for 28 days at the optimal temperature of 20°C while reduce survival time to 24 hours at 40°C (Riddell et al., 2020). Outside of the optimal ranges, the viability of the virus is limited, but it is sufficient to spread since people lack an adaptive immune response to the previously unknown coronavirus .

Second, there is an influence of weather conditions on the susceptibility of the host. Cold and dry air suppresses the innate immune response by damaging the mucosa and slowing mucociliary clearance (Lowen et al., 2007). The innate immune response is essential to prevent initial infection, inhibit viral replication, and regulate the severity of immune response and inflammation (García, 2020). Exposure to the cold environment may cause hormonal changes, which directly or indirectly alters the immune system (van der Lans et al., 2015). Some researchers find that low temperature

[5] All infections involve pathogens, hosts, and the environment.

is associated with decreased lung function and worsened condition in patients with chronic obstructive pulmonary disease (Donaldson et al., 1999).

Third, weather conditions will affect human behavior patterns. Dai & Zhao (2020) point out that in subtropical or tropical climates, due to increased humidity and heat, people usually gather in air-conditioned buildings with reduced indoor ventilation and airflow, which may increase the risk of virus transmission. Menebo (2020) holds that increased sunshine and warmer weather lead to an increase in the number of people gathering in outdoor spaces which enhances the risk of the COVID-19 transmission. The conceptual framework of this article is as follows:

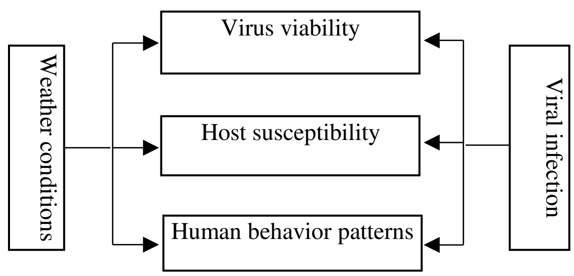

Figure 1 mechanism of the influence of weather conditions on the COVID-19 epidemic

China is the first country to be attacked, and cases occurred in all provincial-level administrative regions and most prefecture-level administrative regions (as of the end of February, there were 19 prefecture-level administrative regions without outbreaks). China's territory spans the tropical, temperate, and frigid zones from south to north, where the Qinling-Huaihe line as a geographic boundary is the 0 °C isotherm in January. The average temperature in January is above 0 in the south of the Qinling and below 0 in the north. There is a big difference in temperature between the northern and southern prefecture-level administrative regions. In January 2020, the average high temperature in the northern city of Heihe is -13°C, and the average low temperature is -25°C, while in the southern city of Sanya is 25°C and 16°C, respectively, with a difference of 38°C and 41°C. In February 2020, the average high temperature in Heihe is -9°C, and the average low temperature is -21℃, while in Sanya is 27℃ and 19℃, respectively, with a difference of 36℃ and 40℃. In China, January is usually the coldest month, but the temperature starts to rise in February and rises sharply in March when spring begins when most of the country's temperature is above 0, and the temperature difference between regions is greatly reduced. In March 2020, the highest temperature in Sanya reaches 30℃, and that in Heihe is 14℃.[6] The regional temperature difference between the north and south of China provides an excellent case for studying the influence of weather condition on the COVID-19 epidemic, especially the large temperature difference between the cold temperature below 0 in the north and the warm temperature above 0 in the south in January and February. At the same time, China also adopted stringent public health measures and basically brought the epidemic under control within more than two months. On February 12, the number of new cases reached a peak of 15,152 and on February 19, the number nationwide dropped to 394 when the number of cases in most provinces is already low except for a few provinces such as Hubei. There were no new confirmed cases on March 20 for the first time, after which, most provinces had no new cases or only sporadic cases. Considering the temperature and the number of confirmed cases simultaneously, in this paper, we select the data set of all prefecture-level administrative regions in mainland China from January 1 to February 19, 2020, as the research sample.

[6] The above temperature data is inquired from the National Meteorological Information Center (NIMIC) website, http://data.cma.cn.

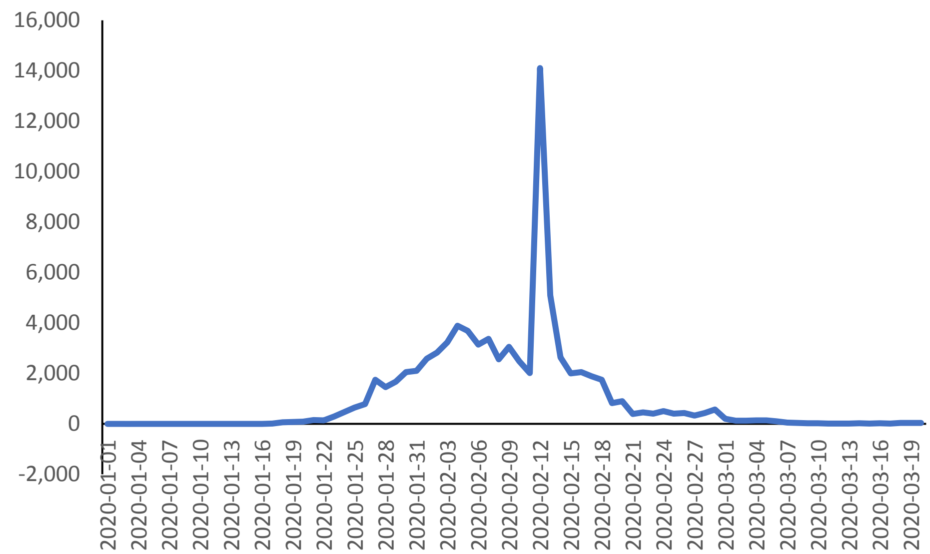

Figure 2 New confirmed cases of the COVID-19 in China from January 1, 2019 to March 20, 2020

Based on the above analysis,we proposes the research hypotheses as follows.

Hypothesis 1: There is a negative causal relationship between the average temperature and the COVID-19 epidemic;

Hypothesis 2: There is a negative causal relationship between the relative humidity and the COVID-19 epidemic.

4. Research Design

4.1 Empirical Model and Variable

Based on the theoretical analysis above, we apply econometrics approach and empirically test the influence of weather condition on the epidemic using panal data of mainland China. Econometric approach is commonly used to measure the effects of a factor on economic growth. Similar to early COVID-19 infections, economic output generally increases exponentially with a variable rate that can be affected by policies and other conditions (Hsiang et al., 2020). Therefore, it is appropriate to apply econometrics techniques to analyze the influence of weather condition on the outbreak of the epidemic. As mentioned above, compared to statistical methods such as Pearson's correlation coefficient to identify the correlation between weather and COVID-19 (Alkhowailed et al. 2020; Bashir et al. 2020; Menebo, 2020), the multiple linear regression approach of panel data is conducive to overcoming the limitations of the time-series regression and Spearman regression model adopted by most existing literature, and pays more attention to identifying the causal relationship between the variables (Angrist, 2008), that is, whether changes in temperature and humidity lead to changes in the epdidemic. The key of causal inference is to control the observable factors that interfere with the causal relationship. In order to avoid biased estimators led by omitting variables, we adopt the two-way fixed effect model to control the time-invariant individual heterogeneity and the individual-invariant time heterogeneity. The empirical model is as follows:

$$\mathrm{rate}_{it} = \alpha + \beta_1 AT_{it} + \beta_2 \mathrm{RH}_{it} + \beta_3 Measure_{it} + \beta_4 bed_{it} + \beta_5 pop_{it} + \beta_6 distance_{it} + t + \delta_i + \delta_t + \varepsilon_{it} \tag{1}$$

Where, $rate_{it}$ is the explained variable, representing the actual cumulative case growth rate of city i on date t.

$$Actual\ cumulative\ case\ growth\ rate\ _{it} = (Actual\ case_{it} - Actual\ \ case_{it-1})\ /Actual\ \ case_{it-1} \tag{2}$$

If there are no cases on the current day, $actual\ \ case_{it-1}$=0, let $Actual\ cumulative\ case\ growth\ rate\ _{it} = 0$.

If there are no confirmed cases on that day, then $Actual\ cumulative\ cases_{it-1} = 0$, Let the cumulative case growth rate be 0.

Considering that the average incubation period of COVID-19 is 5.2 days (Li et al., 2020), we take the fifth lead of reported cases as the proxy variable of the actual cases, namely:

$$reported\ case_{it} = actual\ case_{it+5} \tag{3}$$

Taking Chengdu, China as an example, the number of reported cumulative confirmed cases in this city on February 1, 2020, is 73. Considering the incubation period of 5 days, 73 should be the number of real cumulative cases five days previously, i.e., on January 27, 2020, while the number of real cases on February 1 should be the cumulative number of reported cases on February 6 (see table 1).

Table 1 Conversion between the actual cases and the reported cases

Date	**Reported**	**Actual**
24-Jan	12	59
25-Jan	22	69
26- Jan	33	72
27- Jan	37	73
28- Jan	46	77
29- Jan	59	87
30- Jan	69	92
31- Jan	72	97
1-Feb	73	102
2-Feb	77	109
3-Feb	87	120
4-Feb	92	123
5-Feb	97	124

6-Feb	102	125

Since the existing literature shows limited evidence for the correlation between wind speed, precipitation, and other weather variables (McClymont & Hu, 2021), temperature and humidity are selected as weather condition variables. AT_{it} denotes the average temperature of city i on date t, and RH_{it} denotes the relative humidity of city i on date t.

$Measure_{it}$ denoes the total score of public health measures of the city i on date t, which uses detailed assessment data from all prefecture-level regions. Compared with other countries, China's public health measures are stringent, but the central government does not uniformly stipulate them. Instead, regional governments decide when to start and what measures to take based on local conditions, therefore, there are significant differences between regions. In this paper, we conduct a very detailed evaluation through a manual collection of data. Hanage (2020) concludes that the comprehensive intervention measures implemented in China successfully alleviate the spread of COVID-19, especially in the early stages of the outbreak, so we include the factor into the model total score of public health measures.

The population movement is measured by effective distance which is proposed by Brockmann & Helbing (2013) as a new concept. They believe that population movement does not depend on the geographical distance between regions but the convenience of mobility. In this paper,we calculate the effective distance and incorporate it in the model ($distance_{it}$).

Regarding the socio-economic development, we take medical conditions, economic development level, and population size into consideration and also take number of beds in medical institutions (bed), GDP per capita (pergdp), and population size (pop) as control variables. Although the data structure is a wide panel, the time span is relatively long, and the development of the epidemic itself has a time trend, so we introduced the time trend "t" to control the variation trend of the explained variable over time. It is common in econometric studies to consider time trend in modeling. For example, when exploring the influence of economic development on the degree of democracy, Markus Bruckner et al. (2011) also introduced t into the econometric model considering that democratic development itself has time trend. δ_i is a region fixed effect to control the characteristics of provinces constant over time, δ_t is a time fixed effect to control to control the time factors that do not vary from individual to individual. ε_{it} is error term, we use cluster-robust standard error to estimate the standard deviation (Cameron & Miller, 2015).

4.2 Data

In the research data in this article, the maximum and minimum temperature data in selected areas comes from website of the *National Meteorological Information Center* (NMIC)[7]. the cumulative confirmed case of came from the official release of the National and local *Health Commission.* The original data on public health interventions is collected according to the information or announcements issued by the *prevention and control headquarters* of the prefecture-level administrative districts. The data of population movement used to calculate the effective distance comes from *Baidu Migration*, which is the data of migration and outflow between different regions collected by Baidu inc using big data technology. population size, GDP per capita, and number of beds in medical institutions are from the *China City Statistical Yearbook*. Due to the lack of statistical data in some prefecture-level administrative regions in the *China City Statistical Yearbook*, the number of regions returned to use was finally 279 after being eliminated.

[7] NMIC website query, http://data.cma.cn.

Table 2 Variable Explanation

Attribute	Name	Explanation	Data Sources
Explained variable	Rate	Increased rate of confirmed cases	The official website of the National and local Health Commissions
Explanatory variable	RH	Relative humidity	NMIC
	AT	average temperature	NMIC
Control variable	Hospital_bed	Number of beds in medical institutions	China City Statistical Yearbook
	Measure	Total score of public health intervention	Relevant announcements from the prevention and control headquarters of the COVID-19 epidemic in various provinces and cities
	pop	registered population	China City Statistical Yearbook
	distance	effective distance	Baidu Migration

4.3 Calculation of Major Variables

a. Weather Variables of prefectural-level administrative regions

We take average temperature (AT) which is the average value of the highest temperature (HT) and the lowest temperature (LT), as the proxy variable of temperature:

$$AT = (HT+LT)/2 \tag{4}$$

Figure 3 shows the distribution of the mean of the average temperature from January 1 to February 19, 2020, from which it can be seen that the average temperature decreases in a step-like manner from south to north.

The calculation method relative humidity (RH) is as follows.

$$\mathrm{RH} = E_S/E_0 \tag{5}$$

$$E_0 = 6.11 * 10^{\frac{7.5T_c}{T_c+237.7}} \tag{6}$$

$$E_S = 6.11 * 10^{\frac{7.5T_d}{T_d+237.7}} \tag{7}$$

Where T_c is the air temperature (degrees Celsius), T_d is the dew point temperature (degrees Celsius). Figure 4 shows the distribution of the mean of the relative humidity from January 1 to February 19, 2020. It can be seen that the relative humidity in the southeast coastal area is higher, and it shows a decreasing trend in the northwest direction.

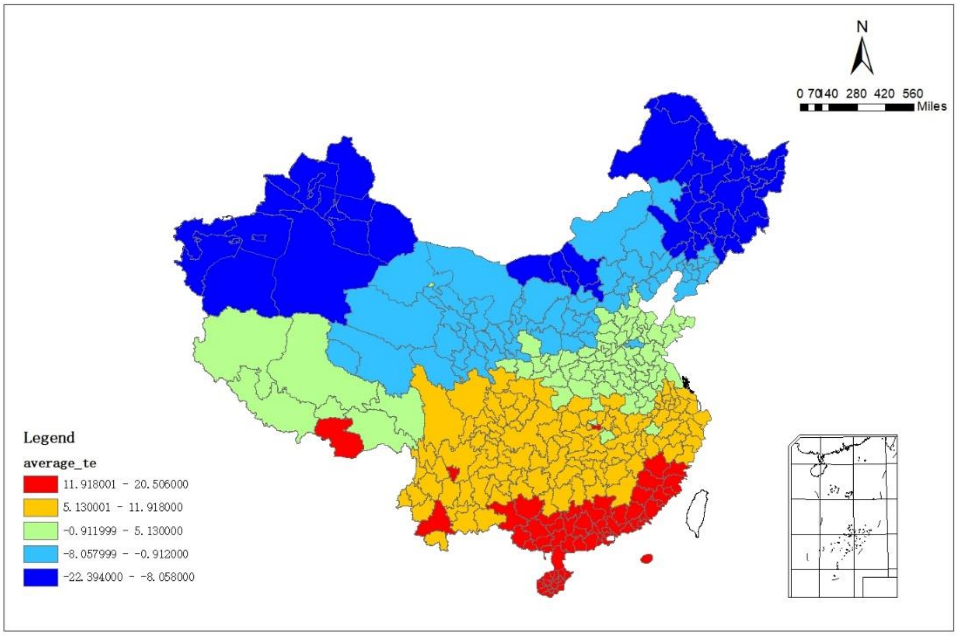

Figure 3 Mean of the Average temperature in China from January 1 to February 19, 2020

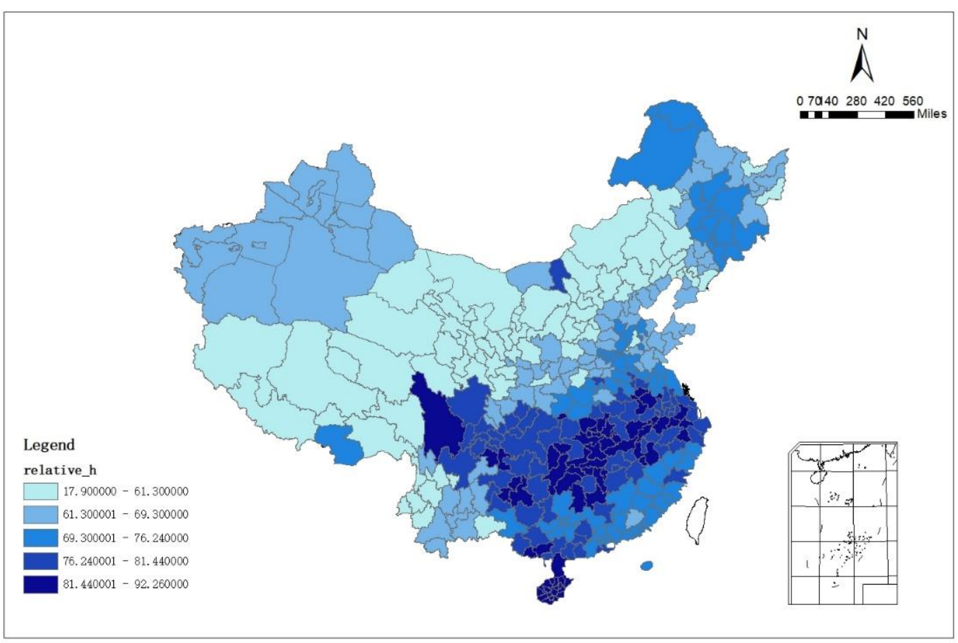

Figure 4 Mean of the relative humidity in China from January 1 to February 19, 2020

b. Total scores of public health intervention measures in prefecture-level administrative regions

We construct scoring data as proxy variable for public health intervention measures. According to the public health intervention measures taken by various prefecture-level administrative regions, we summarize 15 specific items (see table 3), each with a score of 1. Scoring starts until the measure is canceled. For example, on January 21, Shanghai began to implement "quarantining the contacts for 14 days, then the score of Shanghai from January 21 is 1. On January 24, Shanghai began to implement "closing part of the indoor urban public places", then the socre of Shanghai is added another 1 point since January 24, and so on, and finally, the points will be added up. The total score of public health intervention measures on February 19, 2020 is shown in the figure 5.

Table 3 Items of Public Health Intervention Measures

Launching level 1 response	Closing part the public places
Suspending all the cross-city passenger transport	Closed management of all the community
Suspending part of the cross-city passenger transport	Closed management of part of the community
Monitoring all the cross-city passenger transport	Quarantining returnees from key epidemic area (Hubei) for 14days
Monitoring part of the cross-city passenger transport	Quarantining all the returnees for 14days

Suspending all the public transport	Quarantining the contact for 14days
Suspending part of the public transport	Isolating and testing the suspected
Closing all the public places	

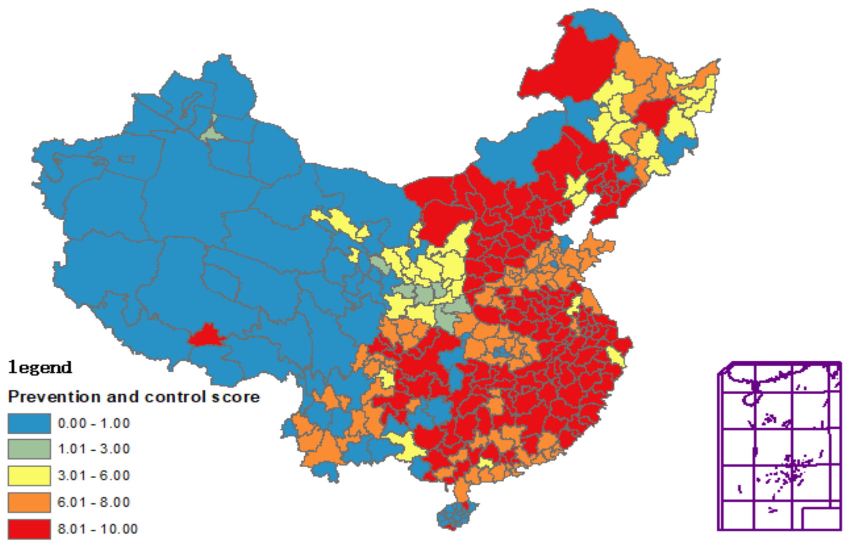

Figure 5 Total Scores of Public Health Measures in China on February 19, 2020

c. Effective distance to Wuhan

We draw on the concept of effective distance proposed by Brockmann and Helbing (2013) to better reflect the population movement between regions. They believe that the spread of the disease has nothing to do with the geographic distance between cities but is closely related to the effective distance. The effective distance between cities is the length that passengers choose between the various alternative routes from city i to city j. Passengers can be regarded as random particles, and they can visit the surrounding cities randomly according to traffic flow. After arriving at the next city, according to the traffic flow, it is converted into the probability of visiting the neighboring cities of the next city. Finally, the path that the particle are most likey to choose from city i to city j is the most probable path, and the length of the most probable path is the effective distance. This effective distance of probability is related to population movement and traffic convenience. The shorter the effective distance, the greater the probability of spreading the epidemic to the region, the greater the probability of increasing the input cases, and the earlier the outbreak time of the large-scale epidemic. The calculation of effective distance draws on the improved method of Lin et al. (2020). Assuming that P_{mn} is the proportion of the population from node N to node M, since the effective distance is cumulative, and the probability of multi-segment paths is calculated by multiplication, so P_{mn} take the logarithm, that is, the effective distance from node N to node M is expressed as: $d_{nm=}$1-log P_{mn}. Infectious diseases are more likely to spread to nodes with high connectivity in the network, where this kind of inequality is represented by $d_{mn} \neq d_{nm}$. The effective propagation distance depends only on the topological characteristics of the network, that is, the matrix P. In this paper, the probability in the effective distance P_{mn} (the proportion of the traffic flow from node M to node N to the total traffic flow from all nodes to node N) calculated using the traffic flow from the Baidu migration data . When there are multiple paths between two nodes, we can traverse all the paths $\Gamma = \{n_1, \Lambda\ , n_L\}$, and take the shortest one of the effective distance as the final path length between two nodes, $D_{mn} = \min_{\Gamma} \lambda(\Gamma)$ $D_{mn} \neq D_{nm}$. Among random paths starting at node N and ending at node M, the path closest to a straight line has the most significant probability and the shortest effective distance. Starting from a selected starting node N, the shortest paths to other nodes can form a shortest-path tree (Brockmann & Helbing, 2013). The effective distance between Wuhan and each other prefectural-level administrative regions on February 19, 2020 is shown in Figure 6.

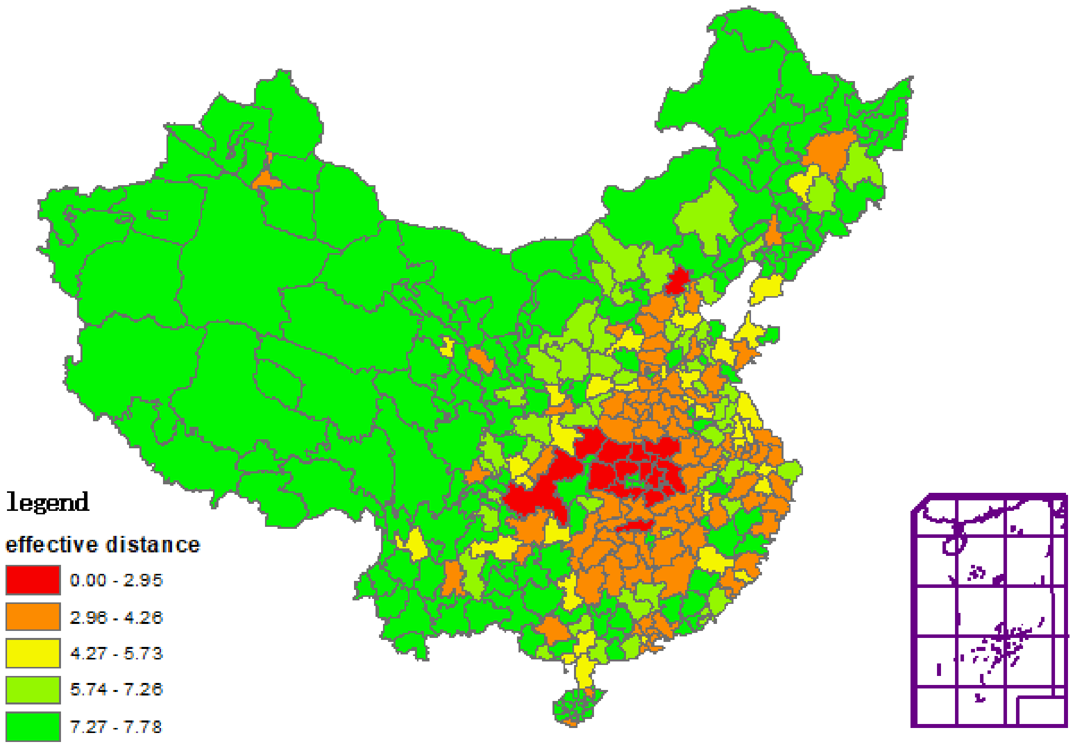

Figure 6 Effective Distance between Wuhan and Each Other Prefectural-level Administrative Regions on February 19, 2020

4.4 Statistical Description

The data samples in this paper is composed of balance panel data of 279 prefecture-level administrative regions from January 1 to February 19, 2020, and the descriptive statistics of related variables are shown as table 4.

Table 4 Statistical Description

Variables	**Implication**	**Notation**	**N**	**mean**	**sd**	**min**	**max**
Explained variable	Increased rate of confirmed cases	Rate	13,950	0.0851	0.443	0	19

Explanatory variable	average temperature	AT	13,950	3.789	9.0667	-31.2	26.4
	Relative humidity	RH	13,850	71.17	17.15	6	102
Control variable	Public health measures score	Measure	13,950	3.729	3.912	0	10
	Number of registered population	Pop	13,950	171.3	226.3	16	2,451
	Number of hospital beds	Hospital_bed	13,950	12,906	17,135	920	142,708
	GDP per capita	Per GDP	13,950	92,348	379,890	17,890	6.400e+06
	effective distance	Distance	13,950	5.722	1.874	0	7.785

5. Results

5.1 Does Weather Condition Matter？

5.1.1 Baseline regression

Table 5 reports the results of the baseline regression. The explanatory variable is the growth rate of cumulative cases, and the explanatory variables are average temperature and relative humidity. Column (1) only introduces the explanatory variables, and column (2) - column (5) adds the control variables in sequence based on column (1). The results show that the average temperature coefficient is significantly negative, which indicates that there is a significant negative casual relationship between temperature and the growth rate of the confirmed cases. In cities with higher temperatures, the transmission rate of the epidemic is slow, and the growth rate of confirmed cases is lower; on the contrary, the virus transmission ability is stronger in cold conditions. Similarly, the coefficient of relative humidity is significantly negative, which indicates that there is a significant negative casual relationship between the relative humidity and the growth rate of confirmed cases. In cities with higher humidity, the growth rate of confirmed cases is lower, while in cities with low humidity, the epidemic spread rate is faster, and the growth rate of confirmed cases is higher.

Table 5 Bseline regression

	(1)	(2)	(3)	(4)	(5)
	rate	rate	rate	rate	rate
AT	-0.0036***	-0.0032***	-0.0033***	-0.0033***	-0.0036***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
RH	-0.0017***	-0.0015***	-0.0015***	-0.0015***	-0.0015***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Measure		-0.0220***	-0.0222***	-0.0220***	-0.0223***
		(0.0030)	(0.0030)	(0.0030)	(0.0030)
Pop			0.0299***	0.0415***	0.0316***
			(0.0065)	(0.0109)	(0.0112)
Hospital				-0.0085	-0.0132**
				(0.0064)	(0.0065)
Distance					-0.0133***
					(0.0038)
Constant	0.0709***	0.0162	-0.1226***	-0.1007**	0.0663

	(0.0262)	(0.0271)	(0.0407)	(0.0438)	(0.0651)
Observations	12,555	12,555	12,555	12,555	12,555
R-squared	0.033	0.038	0.039	0.039	0.040
Time Trend	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

We also analyze control variables. It is concluded that the coefficient of public health measures is significantly negative, indicating that taking public health measures is essential to mitigate the epidemic. The better the public health measures are, the lower the growth rate of the number of confirmed cases is. The coefficient of population size is significantly positive. The vast population size will increase the difficulty of isolating person-to-person contact which has an adverse effect on blocking the further spread of infectious diseases. The coefficient of both the number of beds in health institutions and effective distance are significantly negative. The number of beds in health institutions represents the medical resources condition of the city, and cities with richer medical resources are more capable of mitigating the deterioration of the epidemic. The shorter the effective distance to Wuhan, the more severe the outbreak of the epidemic, which is in line with theoretical expectations.

5.1.2 Nonlinear relationship or conditional linear relationship?

We have preliminarily verified that the average temperature and relative humidity negatively affects the epidemic, however, we're also concerned that whether this negative relationship only holds up over a certain interval, or whether there is the possibility of a nonlinear relationship.

Therefore, we try to introduce the quadratic terms of the average temperature and relative humidity respectively, to further explore the influence of weather conditions on the epidemic.

Table 6. Is there a Nonlinear Relationship

	(1) Non-linear relationship between AT and rate	(2) Non-linear relationship between RH and rate
AT	-0.0030***	-0.0034***
	(0.0011)	(0.0011)
AT*AT	-0.0002***	
	(0.0001)	
RH	-0.0015***	0.0032*
	(0.0003)	(0.0019)
RH*RH		-0.00003***
		(0.0000)
Measure	-0.0216***	-0.0223***
	(0.0030)	(0.0030)
Pop	0.0304***	0.0313***
	(0.0112)	(0.0112)
Hospital	-0.0126*	-0.0138**

	(0.0065)	(0.0065)
Distance	-0.0130***	-0.0135***
	(0.0038)	(0.0038)
Constant	0.0832	-0.0840
	(0.0652)	(0.0874)
Observations	12,555	12,555
R-squared	0.041	0.041
Time Trend	YES	YES
Province FE	YES	YES
Time FE	YES	YES
-b/2a	-7	46

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 6 reports the regression results of the nonlinear model and figure 7 shows the nonlinear relationship between the average temperature, relative humidity and the epidemic. We find that although the coefficients of quadratic terms of both average temperature and relative humidity are significant, it can be seen from the figure 7 that the positive correlation is almost insignificant on the left side of the inflection point. In order to further confirm whether the negative relationship is within a certain value range of explanatory variables, we perform sub-sample regression based on the value of the inflection point (-b/2a), the inflection point of the average temperature is -7°C, and the inflection point of the relative humidity is 46%.

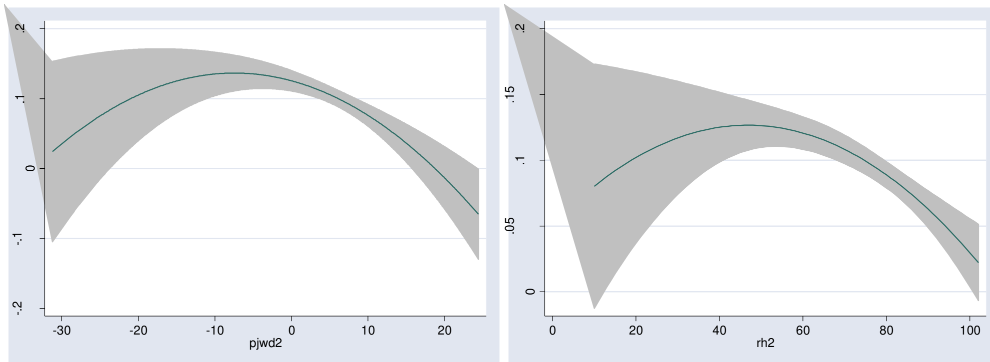

Figure 7 Whether There is a Nonlinear Relationship.

Table 7 reports the results of the sub-sample regression. It can be seen that when the average temperature is more than -7 ℃, it has a negative correlation with the growth rate of the cases; when it is lower than - 7 ℃, then there is no significant correlation between the average temperature and the epidemic; Similarly, when the relative humidity is higher than 46%, there is a negative correlation between it and the cumulative case growth rate, but when it is lower than 46, the decrease of relative humidity will not affect the epidemic. Therefore, there is a conditional linear relationship between weather conditions and the COVID-19 epidemic.

Table 7. Conditional Negative Linear Relationship

	(1)	(2)	(3)	(4)
	AT<-7℃	AT>=-7℃	RH<46%	RH>=46%
AT	-0.0011	-0.0049***	-0.0003	-0.0033***
	(0.0014)	(0.0013)	(0.0021)	(0.0012)
RH	-0.0005	-0.0018***	-0.0002	-0.0019***

	(0.0006)	(0.0004)	(0.0006)	(0.0004)
Constant	0.0041	0.1015	0.1487	0.0775
	(0.1414)	(0.0733)	(0.1357)	(0.0720)
Observations	1,541	11,014	1,087	11,468
R-squared	0.036	0.043	0.026	0.044
Control Variables	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

5.1.3 Robustness

The purpose of robustness check is to demonstrate whether the results change with the adjustment of parameter setting. If the results show that the signs or significance changes after the parameter setting is adjuested, it indicates that the results are not robust and the problem needs to be found out. There are generally three methods of robustness check: 1) In terms of data, adjust the classification according to different standards, and check whether the results are still significant. 2) In terms of variables, replace the orign variables with other proxy. For example, the average temperature can be changed to the maximum temperature. 3) In terms of econometrics identification method, try a variety of identification strategies, such as ordinary least square (OLS), limited information maximum likelihood (LIML), and generalized method of moments (GMM).

First, considering that the epidemic first broke out in city of Wuhan and had a severe impact on other cities in Hubei Province, we eliminate the sample of Hubei Province; Second, we adjust the hypothesis about the length of the incubation period, assuming that the incubation period is 6 days and 7 days respectively, and then perform the regression again. The results of the robustness are reported in Table 8. It can be seen that the sign and significance of the coefficients are consistent with the baseline regression, indicating that the conclusion is still robust after changing the sample selection basic assumptions.

Table 8 Robustness

Panel A without Hubei Province				
	(1)	(2)	(3)	(4)
	AT>-7°C		RH>=46%	
AT	-0.0048***	-0.0048***	-0.0029***	-0.0031***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)
RH	-0.0017***	-0.0015***	-0.0018***	-0.0016***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Observations	10,474	10,474	10,845	10,845
R-squared	0.042	0.052	0.041	0.053
Control Variables	No	YES	NO	YES
Time Trend	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

	(1)	(2)	(3)	(4)

Panel B Adjust the incubation period				
	The incubation period is 6 days	The incubation period is 7 days	The incubation period is 6 days	The incubation period is 7 days
	AT>=-7℃		RH>=46%	
AT	-0.0057***	-0.0065***	-0.0036***	-0.0047***
	(0.0014)	(0.0014)	(0.0013)	(0.0013)
RH	-0.0018***	-0.0009**	-0.0020***	-0.0011***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Observations	10,757	10,510	11,101	10,826
R-squared	0.043	0.041	0.044	0.042
Control Variables	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

5.1.4 Endogenity Treatment

Since the spread of the epidemic may be affected by some unobservable factors, the problem of omitted variables may not be avoided in the regression. In this case, the influence of the omitted variables is included in the error term where when it is related to other explanatory variables, the endogenous problems arise. In order to address the endogenous problems caused by omitted variables, we use the instrumental variable approach to re-estimate. The selection of instrument variables needs to meet two requirements: the first is that there is a significant relationship between the instrumental variables and the endogenous explanatory variables; and the second is that the instrument variable must be exogenous. We take the 1st lag of the explanatory variable as an instrument variable, and in order to ensure the robustness of the results, we respectively use two-stage least squares (2SLS), limited information maximum likelihood (LIML), and generalized method of moments (GMM). Table 9 reports the results of the instrumental variable approach. We can say that there is no obvious change in the sign and significance of the coefficients. The results of the instrumental variables approach are consistent with the baseline regression, indicating that the conclusion is still robust after considering endogeneity.

Table 9. IV Results

	(1)	(2)	(3)	(4)	(5)	(6)
		AT>=-7°C			RH>=46%	
	2SLS	LIML	GMM	2SLS	LIML	GMM
AT	-0.0062***	-0.0062***	-0.0062***	-0.0033***	-0.0033***	-0.0033***
	(0.0010)	(0.0010)	(0.0010)	(0.0008)	(0.0008)	(0.0008)
RH	-0.0019***	-0.0019***	-0.0019***	-0.0032***	-0.0032***	-0.0032***
	(0.0004)	(0.0004)	(0.0004)	(0.0007)	(0.0007)	(0.0007)
Observations	10,732	10,732	10,732	11,101	11,101	11,101
R-squared	0.042	0.042	0.042	0.042	0.042	0.042

Control Variables	YES	YES	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

5.2 Further Exploration

5.2.1 Where is more dangerous?

Since we have proved that when the average temperature is more significant than -7°C, the average temperature is negatively correlated with the spread of the epidemic, and when the relative humidity is greater than 46%, the relative humidity is negatively correlated with the epidemic, then it can be inferred that when the temperature is -7°C, the spread of the epidemic is the most serious; when the relative humidity is 46%, the spread of the epidemic is the most serious. So, in China, a county with vast territory and wide difference in climatic conditions, which cities have more favorable climate conditions for the development of the epidemic?

We make statistics of the prefecture-level administrative regions which meet the requirements that the average temperature is -7℃ ± one standard deviation (9.0667℃), the relative humidity is 46% ± one standard deviation (17.15%), and both meet the two conditions simultaneously from January 1, 2020 to February 19. We construct a dummy variable of whether the city falls into the interval for regression and count the number of days that each city meets the conditions. Table 10 reports the impact of the dangerous weather on the epidemic. It can be seen that the epidemic is indeed more severe in the dangerous weather.

Table 10. The Impact of Dangerous Weathere on the Epidemic

	(1)	(2)
	rate	rate
City whith AT -7°C±9.0667°C	0.0472***	
	(0.0131)	
City with RH 46%±17.15%		0.0313***
		(0.0110)
Observations	12,555	12,555
R-squared	0.039	0.038
Control Variables	YES	YES
Time Trend	YES	YES
Province FE	YES	YES
Time FE	YES	YES

Table 11, Table 12, and Table 13 report the top dangerous areas in the country under three conditions. It can be seen from the results that from a weather perspective, the winter in cities such as Chifeng and Zhangjiakou is more conducive to the outbreak of the epidemic.

Table 11 the Number of Days of Complying with Dangerous Temperature

City	**number of days**
Chengde	45

Hohhot	45
Wuhai	45
Bayannur	45
LanZhou	45
Baiyin	45
Zhangye	45
Dingxi	45
Xining	45
Shizuishan	45
Urumqi	45

Table 12 the Number of Days Complying with Dangerous Relative Humidity

Rank	**City**	**number of days**
1	Chifeng	47
2	Zhangjiakou	47
3	Longnan	46
4	Ulanqab	44
5	Jinzhou	42
6	Lijiang	42
7	Chaoyang	40
8	Chengde	40
9	Zhangye	38

10	Baoshan	37
11	Lincang	37
12	Panzhihua	37

Table 13 the Number of Days that Meet Both the Two Conditions

Rank	City	number of days
1	Zhangjiakou	42
2	Chifeng	39
3	Ulanqab	37
4	Chengde	36
5	Jinzhou	35
6	Zhangye	33
7	Xinzhou	33
8	Chaoyang	32
9	Liaoyang	31
10	Anshan	31

5.2.2 What is the difference between heterogeneous regions?

The difference in weather conditions among different regions in China is affected not only by the large latitude crossing between north and south and also by the significant difference in altitude and the location of coastal and inland. For this reason, we carry out sub-sample regression according to geographical location.

The results are shown in Table 14. Panel A in Table 14 reports the results according to the samples in the east, middle, and west. The eastern region includes Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The middle region includes Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan and Guangxi; the western region includes Sichuan, Guizhou, Yunnan, Tibet, Shanxi, Gansu, Qinghai, Ningxia, and Xinjiang. [8] It can be seen that the coefficient of average temperature is still significantly negative in the eastern and western regions, in which the influence of average temperature in the east is greater than that in the west, but it doesn't work in the middle. The effect of relative humidity is the most significant in the middle, followed by the east and the weakest in the west. Panel B reports the sub-sample results of the coastal and inland areas. According to the *China Marine Statistical Yearbook*, coastal areas are defined as areas with coastlines, which are divided into coastal provinces, autonomous regions and municipalities. At present, there are 53 coastal cities and 242 coastal counties. It shows that the influence of both average temperature and relative humidity is greater in the coastal areas, and the role of weather conditions is more important in the coastal areas than inland.

Table 14 Sub-sample Results

Panel A East Middle and West						
	(1)	(2)	(3)	(4)	(5)	(6)
		AT>=-7°C			RH>=46%	
	East	Middle	West	East	Middle	West
AT	-0.0059***	-0.0025	-0.0056***	-0.0045**	0.0000	-0.0048***
	(0.0020)	(0.0035)	(0.0015)	(0.0020)	(0.0026)	(0.0014)

[8] According to the classification of the *National Bureau of Statistics of China*, http://www.stats.gov.cn/.

RH	-0.0019***	-0.0027***	-0.0013***	-0.0021***	-0.0025***	-0.0015***
	(0.0006)	(0.0009)	(0.0004)	(0.0006)	(0.0008)	(0.0004)
Observations	4,128	3,604	3,257	4,082	4,255	3,040
R-squared	0.054	0.041	0.040	0.054	0.041	0.047
Control Variables	YES	YES	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES

Panel B Coastal and Inland

	(1)	(2)	(3)	(4)
	Coastal	Inland	Coastal	Inland
	AT>=-7℃		RH>=46%	
AT	-0.0098**	-0.0044***	-0.0101**	-0.0025*
	(0.0040)	(0.0015)	(0.0042)	(0.0013)
RH	-0.0018*	-0.0017***	-0.0021*	-0.0017***
	(0.0010)	(0.0004)	(0.0011)	(0.0004)
Observations	2,191	8,798	2,101	9,276
R-squared	0.051	0.043	0.051	0.044
Control Variables	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES
Province FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

5.2.3 What affects the influence of weather conditions on the epidemic

Both Polgreen & Polgreen (2018) and Paraskevis et al. (2020) point out that studying the impact of weather conditions on the spread of the epidemic cannot separate public health interventions and human behavior patterns. In order to further analyze the effects of other important factors, we introduce the interation term of the explanatory variables with diurnal temperature variation , public health measures, and social public opinion to explore the moderating effects of these factors on weather conditions affecting the epidemic.

Jaagus et al. (2014) believe that the main influencing factors of the temperature difference between day and night are latitude, altitude, and location of the land and sea. A large temperature difference between day and night weakens the immune system and makes people more susceptible to infection under equal conditions. The temperature difference between day and night (TD) is the difference between the highest temperature (HT) and the lowest temperature (LT) each day. The calculation method is as follows.

$$TD = HT - LT \tag{8}$$

The temperature difference between day and night in China on February 19, 2020 is shown in figure 8.

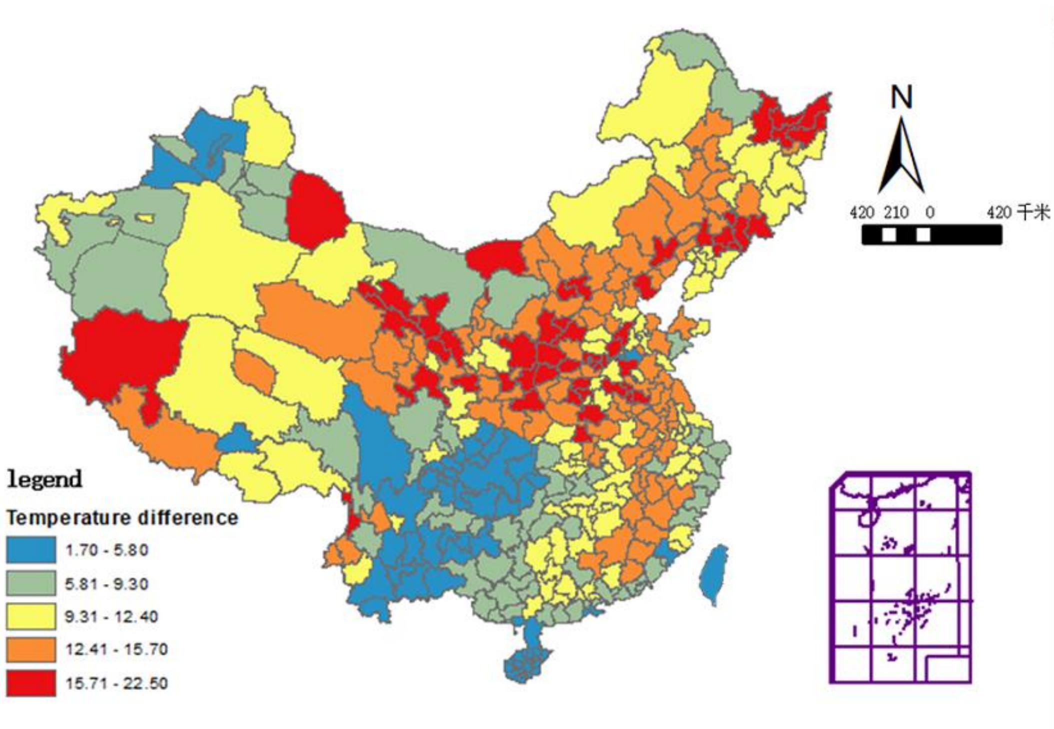

Figure 8 The temperature Difference between Day and Night in China on February 19, 2020

Rapid and strict public health measures can effectively prevent the further spread of the epidemic, and good public opinion can enhance the public's attention to the epidemic to improve the awareness of prevention. Theoretically, both the factors are conducive to reducing the impact of weather conditions on the epidemic. The data of social public opinion is based on the search service provided by *Baidu Index*. We select six epidemic-related terms: the COVID-19, pneumonia, Zhong Nanshan, pneumonia symptoms, mask, and correct wearing of masks to reflect the public's response to the epidemic, apply big data mining technology to collect data, and add up to obtain daily data. Social public opinion in China on February 19, 2020 is shown in figure 9.

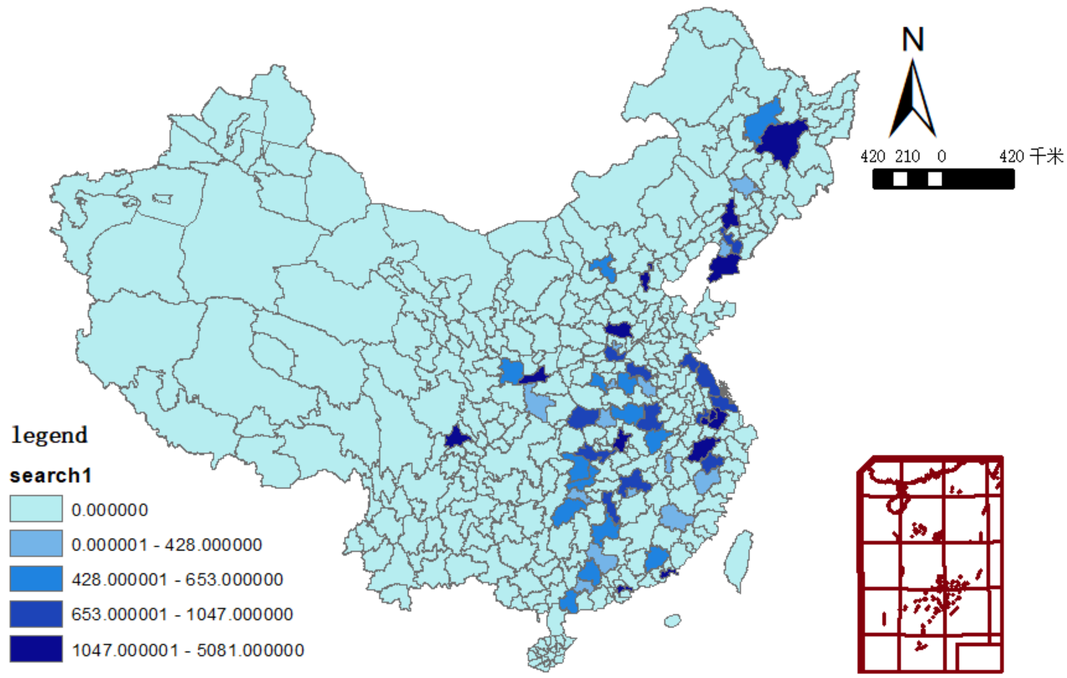

Figure 9 Social Public Opinion in China on February 19, 2020

We divide the three variables of diurnal temperature difference, public health measures, and social public opinion into high, medium, and low, respectively, generate dummy variables, and construct the interaction terms between dummy variables and the weather condition variables. The results are shown in Table 15. Column (1) introduces the interaction between average temperature and high diurnal temperature variation (hightf), and column (2) introduces the interaction between relative humidity and the item; similarly, column (3) and column (4) presents the results of public health measures, and column (3) and column (4) presents the results of social publicopinion.

It can be seen that, whether it is average temperature or relative humidity, the coefficient of the interaction with high diurnal temperature difference (hightf) is significantly negative, that is, the increase in diurnal temperature differences lead to a stronger impact of weather conditions on the increase in the rate of confirmed COVID-19 cases, especially in dry and cold regions, where higher diurnal temperature differences will increase the risk of the spread of the epidemic. On the contrary, the coefficients of the interaction with high public health measures (highpolicy) and high social public opinion (highopinion) are both significantly positive, indicating that the improvement of public health measures and social public opinion can weaken the influence of average temperature and relative humidity on the COVID - 19 confirmed cases growth rate. It can be concluded that strict public health measures and sound public opinion can mitigate the adverse effects of cold and dry weather on the spread of the epidemic, which reinforces the importance of public health measures and attention to public response.

Table 15 Exploration of Moderating Effects

	(1) rate	(2) rate	(3) rate	(4) rate	(5) rate	(6) rate
AT	-0.0033**	-0.0033***	-0.0063***	-0.0027**	-0.0056***	-0.0032***
	(0.0015)	(0.0012)	(0.0014)	(0.0012)	(0.0014)	(0.0012)
RH	-0.0020***	-0.0018***	-0.0021***	-0.0034***	-0.0020***	-0.0022***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
AT*hightf	-0.0034***					
	(0.0011)					
RH* hightf		-0.0005***				
		(0.0001)				
AT*highpolicy			0.0040***			
			(0.0013)			
RH* highpolicy				0.0028***		
				(0.0002)		
AT*highopinion					0.0030***	
					(0.0011)	
RH*highopinion						0.0004***
						(0.0001)
Observations	10,989	11,377	10,989	11,377	10,989	11,377
R-squared	0.038	0.037	0.038	0.050	0.037	0.036

Control Variables	YES	YES	YES	YES	YES	YES
Time Trend	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

6. Conclusion and Discussion

In this paper, we collect the COVID-19 related prefecture-daily data of mainland China from January 1, 2020, to February 19, calculate indicators such as growth rate of the confirmed cases, average temperature, relative humidity, the score of public health measures, and effective distance, and empirically test the influence of weather conditions on the COVID-19 epidemic applying two-way fixed effect model of multiple linear regression. The main conclusions are as follows.

First, we analyze the effects of the average temperature on the growth rate of the confirmed cases, and we find that there is a conditional negative linear relationship between the weather conditions and the epidemic. When the average temperature is greater than -7℃, there is a significant negative causal relationship between the average temperature and the growth rate in the confirmed cases, while when the average temperature is less than -7℃, it has no significant effect on the epidemic. Similarly, when relative humidity is greater than 46%, it has a negative impact on the spread of the epidemic, while when relative humidity is less than 46%, the reduction inrelative humidity will no longer affect the epidemic. In robustness checks, we try to remove data of Hubei province from the whole sample, which is most affected by the epidemic in China, and to adjust the assumption of incubation period length in the calculation of actual cumulative case growth rate; considering the possible endogeneity, we take the first-order lag of the main explanatory variables as an instrument variable, and to ensure the robustness of the results, we respectively use two-stage least squares (2SLS), limited information maximum likelihood (LIML), and generalized method of moments (GMM) to reestimate the coefficient.

Second, according to the conditional negative causality between weather conditions and the COVID-19 epidemic, we make statistics of the prefecture-level administrative regions which meet the requirements that the average temperature is -7℃ ± one standard deviation (9.0667℃), the relative humidity is 46% ± one standard deviation (17.15%), and both meet the two conditions simultaneously from January 1, 2020, to February 19. It is concluded that from the perspective of weather conditions, prefecture-level administrative regions such Chifeng, Zhangjiakou, and Ulanqab are more conducive to the outbreak of the epidemic in winter.

Third, we explore the heterogeneity of the influence of weather conditions. Based on the geographical characteristics of China, we conduct sample regression according to the eastern, middle, west, and inland , coastal regions, respectively. We find that the coefficient of average temperature is still significantly negative in the eastern and western regions, in which the influence of average temperature in the east is greater than that in the west, but it doesn’t work in the middle. The effect of relative humidity is the most significant in the middle, followed by the east and the weakest in the west. The influence of both average temperature and relative humidity is greater in the coastal areas, which indicating that the role of weather conditions is more important in the coastal areas than inland.

Finally, by introducing interaction terms, we explore the moderating effect of diurnal temperature difference, public health measures, and public opinion on the influence of weather conditions on the epidemic to investigate the effects of these factors on the intensity of weather conditions. We find that the coefficient of the interaction between weather conditions and the high diurnal temperature difference is significantly negative, suggesting that the increase in diurnal temperature differences lead to a stronger impact of weather conditions on the increase in the growth rate of the COVID-19 cases, especially in dry and cold regions, where higher diurnal temperature differences will increase the risk of the spread of the epidemic; the coefficients of the interaction with high public health measures and high social public opinion are both significantly positive, indicating that the improvement of public health measures and social public opinion can weaken the influence of weather conditions on the COVID - 19 confirmed cases growth rate, that is, strict public health measures and sound public opinion can mitigate the adverse effects of cold and dry weather on the spread of the epidemic.

Using panel data, this paper applies the two-way fixed effect model of multiple linear regression to explore the causal relationship between weather conditions and the COVID-19 epidemic and takes into account important influencing factors such as human behavior patterns and public health measures, which draws new conclusions. In future research, it can adopt more detailed investigation methods. Under the legal framework of privacy protection, questionnaire surveys can be carried out with patients' consent to draw more accurate conclusions. At the same time, in terms of the mechanism of the role of weather variables, more in-depth interdisciplinary cooperation with epidemiologists is needed to study the specific impact of weather conditions on the survivability of the COVID-19 virus and the immunity of susceptible populations to obtain a clearer picture and compelling conclusions.

References

Aboubakr, H. A., Sharafeldin, T. A., & Goyal, S. M. (2021). Stability of SARS-CoV-2 and other coronaviruses in the environment and on common touch surfaces and the influence of climatic conditions: A review. *Transboundary and emerging diseases*, *68*(2), 296-312.

Alkhowailed, M., Shariq, A., Alqossayir, F., Alzahrani, O. A., Rasheed, Z., & Al Abdulmonem, W. (2020). Impact of meteorological parameters on COVID-19 pandemic: A comprehensive study from Saudi Arabia. *Informatics in medicine unlocked*, *20*, 100418.

Altamimi, A., & Ahmed, A. E. (2020). Climate factors and incidence of Middle East respiratory syndrome coronavirus. *Journal of Infection and Public Health*, *13*(5), 704-708.

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Barreca, A. I., & Shimshack, J. P. (2012). Absolute humidity, temperature, and influenza mortality: 30 years of county-level evidence from the United States. *American journal of epidemiology*, *176*(suppl_7), S114-S122.

Bashir, M. F., Ma, B., Komal, B., Bashir, M. A., Tan, D., & Bashir, M. (2020). Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Science of the Total Environment*, *728*, 138835.

Briz-Redón, Á., & Serrano-Aroca, Á. (2020). A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Science of the total environment*, *728*, 138811.

Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *science*, *342*(6164), 1337-1342.

Brückner, M., & Ciccone, A. (2011). Rain and the democratic window of opportunity. *Econometrica*, *79*(3), 923-947.

Cameron, A. C., & Miller, D. L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317-372.

Chan, K. H., Peiris, J. M., Lam, S. Y., Poon, L. L. M., Yuen, K. Y., & Seto, W. H. (2011). The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Advances in virology*, *2011*.

Chen, G. G., & Astebro, T. (2001). The economic value of reject inference in credit scoring. *Department of Management Science, University of Waterloo*.

Cheng, J., Xu, Z., Zhu, R., Wang, X., Jin, L., Song, J., & Su, H. (2014). Impact of diurnal temperature range on human health: a systematic review. *International journal of biometeorology*, *58*(9), 2011-2024.

Chien, L. C., & Chen, L. W. (2020). Meteorological impacts on the incidence of COVID-19 in the US. *Stochastic Environmental Research and Risk Assessment*, *34*(10), 1675-1680.

Cobey, S. (2020). Modeling infectious disease dynamics. *Science*, *368*(6492), 713-714.

Dai, H., & Zhao, B. (2020, December). Association of the infection probability of COVID-19 with ventilation rates in confined spaces. In *Building Simulation* (Vol. 13, No. 6, pp. 1321-1327). Tsinghua University Press.

Donaldson, G. C., Seemungal, T., Jeffries, D. J., & Wedzicha, J. A. (1999). Effect of temperature on lung function and symptoms in chronic obstructive pulmonary disease. *European respiratory journal*, *13*(4), 844-849.

Epstein, P. (2010). The ecology of climate change and infectious diseases: comment. *Ecology*, *91*(3), 925-928.

Fisman, D. N. (2007). Seasonality of infectious diseases. *Annu. Rev. Public Health*, *28*, 127-143.

García, L. F. (2020). Immune response, inflammation, and the clinical spectrum of COVID-19. *Frontiers in immunology*, *11*, 1441.

Gardner, E. G., Kelton, D., Poljak, Z., Van Kerkhove, M., Von Dobschuetz, S., & Greer, A. L. (2019). A case-crossover analysis of the impact of weather on primary cases of Middle East respiratory syndrome. *BMC infectious diseases*, *19*(1), 1-10.

Hossain, M. S., Ahmed, S., & Uddin, M. J. (2021). Impact of weather on COVID-19 transmission in south Asian countries: An application of the ARIMAX model. *Science of The Total Environment*, *761*, 143315.

Hsiang, S., Allen, D., & Annan-Phan, S., et al. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, *584*(7820), 262-267.

Huang, Z., Huang, J., Gu, Q., Du, P., Liang, H., & Dong, Q. (2020). Optimal temperature zone for the dispersal of COVID-19. *Science of The Total Environment*, *736*, 139487.

Iqbal, M. M., Abid, I., Hussain, S., Shahzad, N., Waqas, M. S., & Iqbal, M. J. (2020). The effects of regional climatic condition on the spread of COVID-19 at global scale. *Science of the Total Environment*, *739*, 140101.

Islam, N., Bukhari, Q., Jameel, Y., Shabnam, S., Erzurumluoglu, A. M., Siddique, M. A., ... & D'Agostino Sr, R. B. (2021). COVID-19 and climatic factors: A global analysis. *Environmental research*, *193*, 110355.

Jaagus, J., Briede, A., Rimkus, E., & Remm, K. (2014). Variability and trends in daily minimum and maximum temperatures and in the diurnal temperature range in Lithuania, Latvia and Estonia in 1951–2010. *Theoretical and applied climatology*, *118*(1), 57-68.

Jahangiri, M., Jahangiri, M., & Najafgholipour, M. (2020). The sensitivity and specificity analyses of ambient temperature and population size on the transmission rate of the novel coronavirus (COVID-19) in different provinces of Iran. *Science of the total environment*, *728*, 138872.

Li, Q., Guan, X., & Wu, P., Wang, X., et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England journal of medicine*.

Lin, K. U. N., Fong, D. Y. T., Zhu, B., & Karlberg, J. (2006). Environmental factors on the SARS epidemic: air temperature, passage of time and multiplicative effect of hospital infection. *Epidemiology & Infection*, *134*(2), 223-230.

Liu, H., Zhang, Y., Tian, Y., Zheng, Y., Gou, F., Yang, X., He, J., Liu, X., Meng, L., & Hu, W. (2020). Epidemic features of seasonal influenza transmission among eight different climate zones in Gansu, China. *Environmental research*, *183*, 109189.

Liu, J., Zhou, J., Yao, J., et al. (2020). Impact of meteorological factors on the COVID-19 transmission: A multi-city study in China. *Science of the total environment*, *726*, 138513.

Lloyd, G. E. R., Chadwick, J., & Mann, W. N. (Eds.). (1983). *Hippocratic writings* (Vol. 451). Penguin UK.

Lowen, A. C., Mubareka, S., Steel, J., & Palese, P. (2007). Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog*, *3*(10), e151.

McClymont, H., & Hu, W. (2021). Weather Variability and COVID-19 Transmission: A Review of Recent Research. *International journal of environmental research and public health*, *18*(2), 396.

Menebo, M. M. (2020). Temperature and precipitation associate with Covid-19 new daily cases: A correlation study between weather and Covid-19 pandemic in Oslo, Norway. *Science of the Total Environment*, *737*, 139659.

Meyer, A., Sadler, R., Faverjon, C., Cameron, A. R., & Bannister-Tyrrell, M. (2020). Evidence That Higher Temperatures Are Associated With a Marginally Lower Incidence of COVID-19 Cases. *Frontiers in public health*, *8*, 367.

Miller, G. (1962). " Airs, Waters, and Places" in History. *Journal of the history of medicine and allied sciences*, 129-140.

More, A. F., Loveluck, C. P., Clifford, H., Handley, M. J., Korotkikh, E. V., Kurbatov, A. V., McCormick, M., & Mayewski, P. A. (2020). The impact of a six-year climate anomaly on the “Spanish flu” pandemic and WWI. *GeoHealth*, *4*(9), e2020GH000277.

National Research Council. (2001). *Under the weather: climate, ecosystems, and infectious disease*. National Academies Press.

Nevels, M., Si, X., Bambrick, H., Zhang, Y., Cheng, J., McClymont, H., Bonsall, M. B., & Hu, W. (2021). Weather variability and transmissibility of COVID-19: a time series analysis based on effective reproductive number. *Experimental Results*, *2*.

Otter, J. A., Donskey, C., Yezli, S., Douthwaite, S., Goldenberg, S., & Weber, D. J. (2016). Transmission of SARS and MERS coronaviruses and influenza virus in healthcare settings: the possible role of dry surface contamination. *Journal of Hospital Infection*, *92*(3), 235-250.

Owen, S., & Anil, R. (2009). Mahout in Action, MEAP Edition.

Pahuja, S., Madan, M., Mittal, S., & Pandey, R. M. (2021). Weather Parameters and COVID-19: A Correlational Analysis. *Journal of Occupational and Environmental Medicine*, *63*(1), 69.

Pan, J., Yao, Y., Liu, Z., Meng, X., Ji, J. S., Qiu, Y., Wang, W., Zhang, L., Wang, W., & Kan, H. (2021). Warmer weather unlikely to reduce the COVID-19 transmission: An ecological study in 202 locations in 8 countries. *Science of The Total Environment*, *753*, 142272.

Pani, S. K., Lin, N. H., & RavindraBabu, S. (2020). Association of COVID-19 pandemic with meteorological parameters over Singapore. *Science of the Total Environment*, *740*, 140112.

Paraskevis, D., Kostaki, E. G., Alygizakis, N., Thomaidis, N. S., Cartalis, C., Tsiodras, S., & Dimopoulos, M. A. (2020). A review of the impact of weather and climate variables to COVID-19: In the absence of public health measures high temperatures cannot probably mitigate outbreaks. *Science of the Total Environment*, 144578.

Park, J. E., Son, W. S., Ryu, Y., Choi, S. B., Kwon, O., & Ahn, I. (2020). Effects of temperature, humidity, and diurnal temperature range on influenza incidence in a temperate region. *Influenza and other respiratory viruses*, *14*(1), 11-18.

Polgreen, P. M., & Polgreen, E. L. (2018). Infectious diseases, weather, and climate. *Clinical Infectious Diseases*, *66*(6), 815-817.

Qi, H., Xiao, S., Shi, R., Ward, M. P., Chen, Y., Tu, W., Su, Q., Wang, W., Wang, X., & Zhang, Z. (2020). COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis. *Science of the total environment*, *728*, 138778.

Riddell, S., Goldie, S., Hill, A., Eagles, D., & Drew, T. W. (2020). The effect of temperature on persistence of SARS-CoV-2 on common surfaces. *Virology journal*, *17*(1), 1-7.

Roussel, M., Pontier, D., Cohen, J. M., Lina, B., & Fouchet, D. (2016). Quantifying the role of weather on seasonal influenza. *BMC Public Health*, *16*(1), 1-14.

Shaman, J., & Kohn, M. (2009). Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences*, *106*(9), 3243-3248.

Shi, P., Dong, Y., Yan, H., Zhao, C., Li, X., Liu, W., ... & Xi, S. (2020). Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Science of the total environment*, *728*, 138890.

Tamerius, J., Nelson, M. I., Zhou, S. Z., Viboud, C., Miller, M. A., & Alonso, W. J. (2011). Global influenza seasonality: reconciling patterns across temperate and tropical regions. *Environmental health perspectives*, *119*(4), 439-445.

Tan, J., Mu, L., Huang, J., Yu, S., Chen, B., & Yin, J. (2005). An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. *Journal of Epidemiology & Community Health*, *59*(3), 186-192.

Taubenberger, J. K. (2006). The origin and virulence of the 1918 “Spanish” influenza virus. *Proceedings of the American Philosophical Society*, *150*(1), 86.

Taubenberger, J. K., & Morens, D. M. (2006). 1918 Influenza: the mother of all pandemics. *Revista Biomedica*, *17*(1), 69-79.

Thu, T. P. B., Ngoc, P. N. H., & Hai, N. M. (2020). Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Science of The Total Environment*, *742*, 140430.

Towers, S., Chowell, G., Hameed, R., Jastrebski, M., Khan, M., Meeks, J., Mubayi, A., & Harris, G. (2013). Climate change and influenza: the likelihood of early and severe influenza seasons following warmer than average winters. *PLoS currents*, *5*.

van der Lans, A. A., Boon, M. R., Haks, M. C., Quinten, E., Schaart, G., Ottenhoff, T. H., & van Marken Lichtenbelt, W. D. (2015). Cold acclimation affects immune composition in skeletal muscle of healthy lean subjects. *Physiological reports*, *3*(7), e12394.

Van Doremalen, N., Bushmaker, T., & Munster, V. J. (2013). Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. *Eurosurveillance*, *18*(38), 20590.

Watts, N., Adger, W. N., Ayeb-Karlsson, S., Bai, Y., Byass, P., et al. (2017). The Lancet Countdown: tracking progress on health and climate change. *The Lancet*, *389*(10074), 1151-1164.

Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., Du, M., & Liu, M. (2020). Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Science of the Total Environment*, *729*, 139051.

Zhu, L., Liu, X., & Huang, H., et al. (2020). Meteorological impact on the COVID-19 pandemic: A study across

Declarations

- Ethics approval and consent to participate

Not applicable

- Consent for publication

Not applicable

- Availability of data and material

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request

- Competing interests

The authors declare that they have no competing interests

- Funding

the National Social Science Foundation of China (Grant No. 13BJY091)

the National Natural Science Foundation of China (Grant No.71773083).

- Authors' contributions

Ruofei Lin：Conceptualization, Data curation, Visualization, Formal analysis

Xiaoli Wang：Writing- Reviewing and Editing; Validation, Supervision

Junpei Huang：Methodology, Software, Writing - Original Draft

- Acknowledgements

The authors gratefully acknowledge the financial support of the National Social Science Foundation of China (Grant No. 13BJY091) and the National Natural Science Foundation of China (Grant No.71773083).

eight severely affected regions in South America. *Science of The Total Environment*, *744*, 140881.