High Temperature and High Humidity Reduce the Transmission of COVID-19

Jingyuan Wang, Ke Tang, Kai Feng and Weifeng Lv*

March 9, 2020

Abstract. This paper investigates how air temperature and humidity influence the transmission of COVID-19. After estimating the serial interval of COVID-19 from 105 pairs of the virus carrier and the infected, we calculate the daily effective reproductive number, R, for each of all 100 Chinese cities with more than 40 cases. Using the daily R values from January 21 to 23, 2020 as proxies of non-intervened transmission intensity, we find, under a linear regression framework for 100 Chinese cities, high temperature and high relative humidity significantly *reduce* the transmission of COVID-19, respectively, even after controlling for population density and GDP per capita of cities. One degree Celsius increase in temperature and one percent increase in relative humidity lower R by 0.0383 and 0.0224, respectively. This result is consistent with the fact that the high temperature and high humidity significantly reduce the transmission of influenza. It indicates that the arrival of summer and rainy season in the northern hemisphere can effectively reduce the transmission of the COVID-19.

Keywords: COVID-19, Transmission, Effective Reproductive Number, Temperature, Humidity

1

Since December 2019, Wuhan, the capital of Hubei Province, China, has reported an outbreak of atypical pneumonia caused by COVID-19 (SARS-CoV-2 or 2019-nCov)[1,2], the virus has transmitted nationwide and internationally[3,4]. Compared with SARS, the range of the outbreak of COVID-19 is much wider. European Centre for Disease Prevention and Control shows that, as of 2 March 2020, more than 89,000 COVID-19 cases have been reported globally, from all provinces of China and 66 countries globally.[1] Global outbreaks of COVID-19 have posed major obstacles to public health and the world economy [5].

The transmission of viruses can be affected by a number of factors, including climate conditions (such as temperature and humidity), population density and medical care quality[6,7]. Therefore, understanding the relationship between weather and the transmission of COVID-19 is key to forecast the intensity and end time of this epidemic. However, up to now, it is still unknown whether such a relationship exists or not. For example, on March 06, 2020, Michael Ryan, the executive director of the WHO Health Emergencies Program, said that people still did not know the activity or behavior of the COVID-19 virus in different climatic conditions.[2]

Rough observations of outbreaks of COVID-19 outside China show a noteworthy phenomenon. In the early dates of the outbreak, countries with relatively lower air temperature and lower humidity (e.g. Korea, Japan and Iran) see severe outbreaks than warmer and more humid countries (e.g. Singapore, Malaysia and Thailand) do. Considering the natural log of the average number of cases per day from February 8 to 29 as a *rough* measure of the severity of the COVID-19 outbreaks[3], in Figure 1, we show that the severity is negatively related to temperature and relative humidity using 14 countries with more than 20 new cases during this period.[4]

[Figure 1 about here.]

---

[1]Refer to https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-outbreak-novel-coronavirus-disease-2019-covid-19-increased.
[2]Refer to http://www.xinhuanet.com/english/2020-03/07/c_138851282.htm
[3]If a country has its first case after February 8, we use the natural log of average number of cases per day between the first-case date and February 29.
[4]The temperature and relative humidity of the capital city of a country are chosen as a proxy for the country.

2

Inside China, the COVID-19 has spread widely to many cities, and the intensity of transmission and weather conditions in these cities vary largely (shown in Table SI 1), we can, therefore, analyze the determinants of COVID-19 transmission, especially the weather factors. In order to formally quantify the transmission of COVID-19, we first fit 105 samples of serial intervals with the Weibull distribution (a distribution commonly used to fit the serial interval of influenza[8]), then calculate the effective reproductive number, R, a quantity measuring the severity of infectiousness [9], for each of all 100 Chinese cities with more than 40 cases. Because we aim to study the influences of various factors on R under *natural* conditions,[5] we select our data before China's large-scale intervention in the spread of COVID-19 on 24 January, when the first-level response to major public health emergencies in many major cities and provinces including Beijing and Shanghai are announced. Moreover, after the statement of person-to-person transmission from Professor Nanshan Zhong on the evening of January 20 through a public television interview, Chinese hospitals of all provinces began serious case recording of COVID-19, we, therefore, take the daily R values from January 21 to January 23 to proxy the non-intervened R for each city.[6]

Figure 2 shows the average R values from January 21 to 23 for different Chinese cities geographically. Compared with the southeast coast of China, cities in the northern area of China show relatively larger R values and lower temperatures and relative humidity. The scatter plots in Figure 3 illustrate two negative relations between the daily air temperature and R value and between the daily relative humidity and R value, respectively.

[Figure 2 about here.]

[Figure 3 about here.]

Our finding is consistent with the evidence that high temperature and high humidity reduce the transmission of influenza[10-14], which can be explained by two possible

---

[5]If people stay at home for most of their time under the restrictions of the isolation policy, weather conditions are unlikely to influence the virus transmission due to no chance of contacts between people.
[6]Wuhan City imposed travel restriction at 10 a.m. on January 23, but a large amount of people left Wuhan before 10 a.m. on that day, therefore, our sample still includes January 23.

3

reasons: First, the influenza virus is more stable in cold temperature, and respiratory droplets, as containers of viruses, remain airborne longer in dry air[15, 16]. Second, cold and dry weather can also weaken the hosts' immunity and make them more susceptible to the virus[17, 18]. These mechanisms are also likely to apply to the COVID-19 transmission. Our result is also consistent with the evidence that high temperature and high relative humidity reduce the viability of SARS coronavirus [19,20].

We then regress the average R values of various cities on their average temperature and average humidity over the 3 days (January 21 to 23) and control variables including GDP per capita in 2018 and population density in 2018. Note that GDP per capita and population density can be considered as measures for medical conditions and the crowdness of people inside a city, respectively, which might influence the spread of the COVID-19.

Table 1 shows that the air temperature has a quite strong influence on R with significance levels of 1% for all specifications. The influence of relative humidity on R is relatively smaller compared to that of temperature, with significance levels of 5% to 10% for different specifications. The control variables are not as significant as the temperature and relative humidity, but with expected signs. For example, cities with larger GDP per capita are likely to have better health care facilities, which tend to reduce the transmission of COVID-19. In cities with higher levels of population density, the virus is expected to spread faster than that in less crowded cities.

We then run a panel regression of daily R values on daily temperatures, relative humidity and control variables with both fixed and random effects models. Temperature and relative humidity have quite strong influences on R values, with 1% significant levels for both. Relative humidity has a stronger significance than the temperature does, due to its larger daily fluctuation relative to the temperature. Note that since GDP per capita and population density do not change from January 21 to 23, their effects are, therefore, absorbed in the fixed effects dummies in the fixed-effects panel regressions. One degree Celsius increase in temperature and one percent increase in relative humidity lower the R value by 0.0383 and 0.0224, respectively, in the panel regression

with fixed effects. We run a Hausman test with a null hypothesis that the random-effects model is preferred to the fixed-effects one, and get the test's p value of 0.06.

If omitting control variables,[7] the fixed-effects model of Table 2 provides an estimation of the R value for a certain city given its temperature and relative humidity:

$$R = 3.968 - 0.0383\,Temperature - 0.0224\,RelativeHumidity \quad (1)$$

Assuming that the same relationship of Equation (1) applies to cities outside China and that the temperature and relative humid of 2020 are the same as those in 2019, we can draw a map of R values for worldwide cities in Figure 4 by plugging the average March and July temperatures and relative humidity of 2019 into Equation (1). This figure cautions people of the risk of COVID-19 outbreak worldwide, in March and July of 2020, respectively. As expected, the R values are larger for temperate countries and smaller for tropical countries in March. In July, the arrival of summer and rainy season in the northern hemisphere can effectively reduce the transmission of the COVID-19; however, risks remain in some countries in the southern hemisphere (e.g. Australia and South Africa). If we plug the normal summer temperature and relative humidity of Tokyo (28ºC and 85%, respectively) into Equation (1), the transmission of the COVID-19 in Tokyo will be seriously reduced between March and the Olympics: the estimated R value decreases from 1.914 to 0.992, a 48% drop!


[Table 1 about here.]

[Figure 4 about here.]

---

[7]Even though GDP per capita has a significance level about 10% in the panel regression with random effects, we still prefer to not include it in our estimation on R values for worldwide cities, because many countries outside China have different health care systems than China, hence the impact of GDP per capital on other countries may not be the same as it is in China.

5

References

1. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* https://doi.org/10.1038/s41586-020-2008-3 (2020).

2. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* https://doi.org/10.1038/s41586-020-2012-7 (2020).

3. Li, Q. et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2001316 (2020).

4. Callaway, E . Time to use the p-word? Coronavirus enters dangerous new phase. *Nature* https://doi 10.1038/d41586-020-00551-1 (2020).

5. Phelan, A.L., Katz, R. & Gostin LO. The Novel Coronavirus Originating in Wuhan, China: Challenges for Global Health Governance. *JAMA* **323,** 709-710 (2020).

6. Hemmes, J.H. et al. Virus survival as a seasonal factor in influenza and poliomyelitis. *Nature* **188**, 430-431 (1960).

7. Dalziel, B.D. et al. Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities, *Science* **362**, 75-79 (2018).

8. Cowling, B. J. et al. Estimation of the serial interval of influenza. *Epidemiology* **20**, 344 (2009).

9. Wallinga, J. & Teunis, P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* **160***,* 509–516 (2004).

10. Lowen, A., Samira, M., & Peter, Palese. High temperature (30ºC) blocks aerosol but not contact transmission of influenza virus. *J. Virol.* **82**, *5650*–5652 (2018).

11. Park, J. et al. Effects of temperature, humidity, and diurnal temperature range on influenza incidence in a temperate region. *Influenza. Other. Respir. Viruses* **14**, 11–18 (2020).

12. Shaman, J. & Kohn, M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc. Natl. Acad. Sci. USA* **106**, 3243-3248 (2009).

13. Steel, J., Palese, P. & Lowen, A. Transmission of a 2009 pandemic influenza virus shows a sensitivity to temperature and humidity similar to that of an H3N2 seasonal strain. *J. Virol.* **85**, 1400-1402 (2011).

14. M. Lipsitch, C. Viboud, Influenza seasonality: Lifting the fog, *Proc. Natl. Acad. Sci.* **106**, 3645-3646 (2009).

15. Lowen, A, & John, S. Roles of humidity and temperature in shaping influenza seasonality. *J. Virol.* **88**, 7692–7695 (2014).

16. Tellier, R. Aerosol transmission of influenza A virus: a review of new studies. *J. R. Soc. Interface* **6**, S783–S790 (2009).

17. Eccles, R**.** An explanation for the seasonality of acute upper respiratory tract viral infections. *Acta Otolaryngol.* **122**, 183-191 (2002).

18. Kudo, E. et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. *Proc. Natl. Acad. Sci. USA* **116**, 10905–10910 (2019) .

19. Chan, K.H. et al. The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus, *Adv. Virol.*, 734690-734696 (2011).

20. Yuan, J. et al, A climatologic investigation of the SARS-CoV outbreak in Beijing, China, *Am. J. Infect. Control* 34, 234-236 (2006).
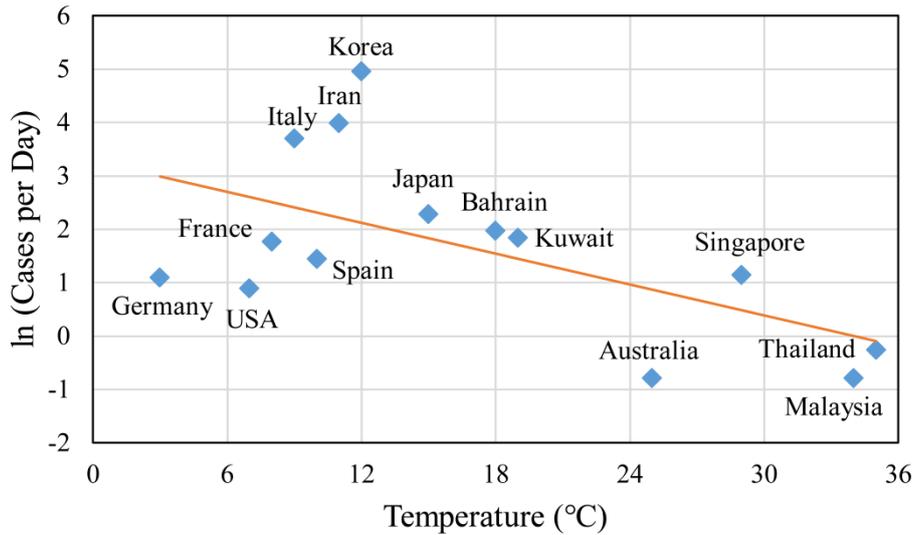
(a)
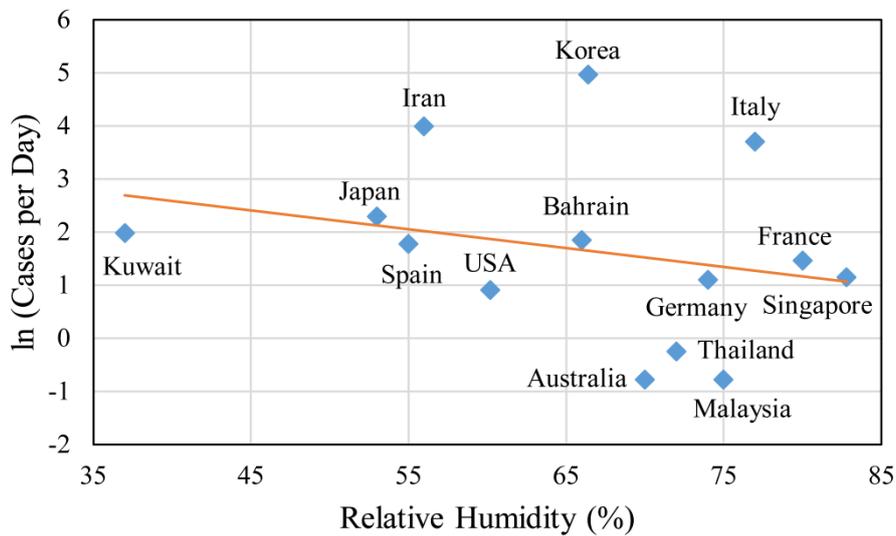
(b)

Figure 1: Severity of COVID-19 outbreaks v.s. temperature and relative humidity for countries outside China.

The natural log of the average number of cases per day from February 8 to 29 is used to proxy the severity of the COVID-19 outbreak for a certain country. If a country has its first case after February 8, we use the natural log of the average number of cases per day between the first-case date and February 29. The temperature and relative humidity of the capital city are used as proxies for the country. Negative relationships between temperature and severity and between humidity and severity are shown in (a) and (b), respectively.
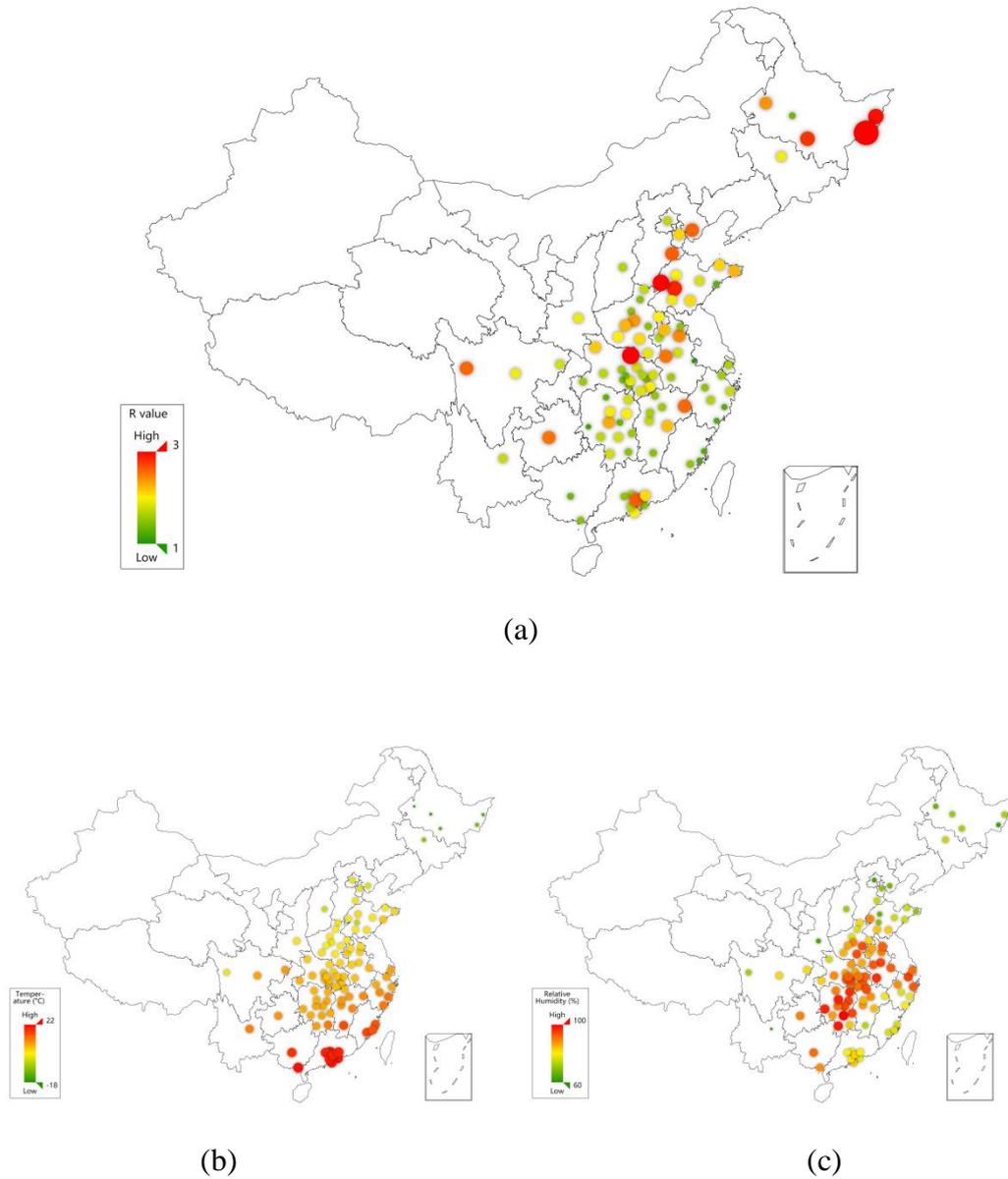
8

(a)

(b)

(c)

**Figure 2: A city-level visualization of the COVID-19 transmission (a), temperature (b) and relative humidity (c).**

Average *R* values from January 21 to 23, 2020 for 100 Chinese cities are used in subplot (a). The average temperature and relative humidity for the same period are plotted in (b) and (c). Subplots (a), (b) and (c) together inform that the *R* values are larger in the cold and dry northern regions of China.

9
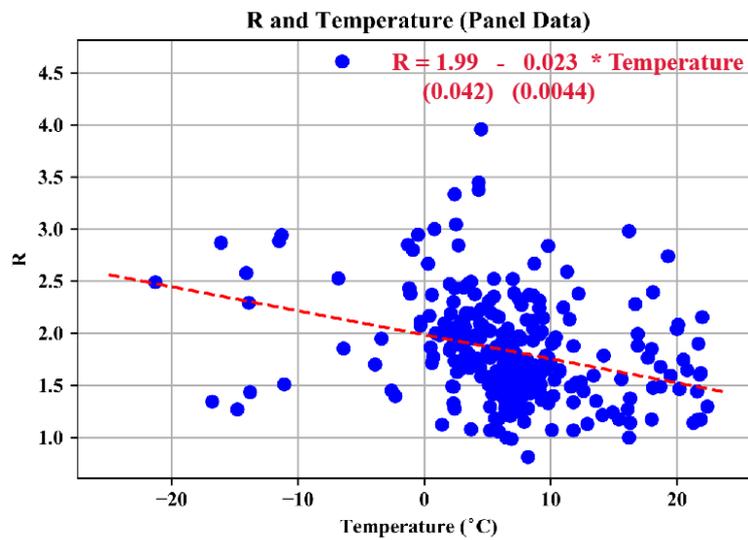
(a)
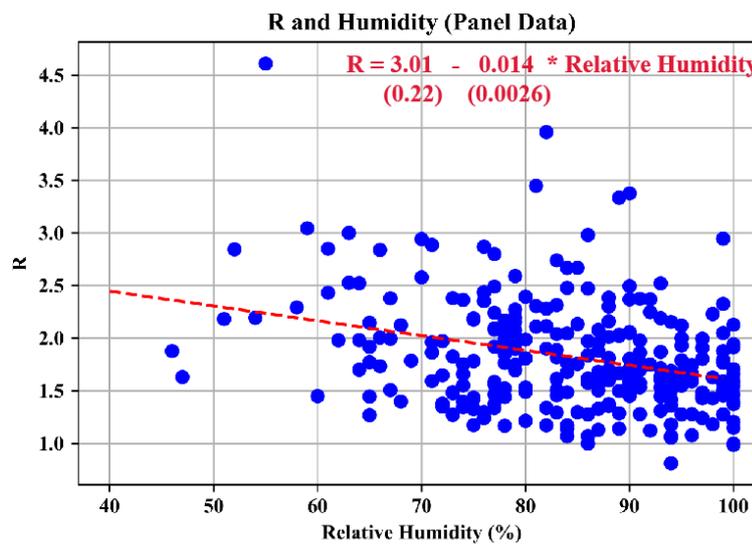
(b)

Figure 3: Effective reproductive number R v.s. temperature and relative humidity for 100 Chinese cities

Daily *R* values, temperature and relative humidity from January 21 to 23, 2020 for 100 Chinese cities are used in this figure. Negative relationships between temperature and *R* and humidity and *R* are shown in (a) and (b), respectively. The quantities in parentheses are the standard deviations of the numbers above them.

10
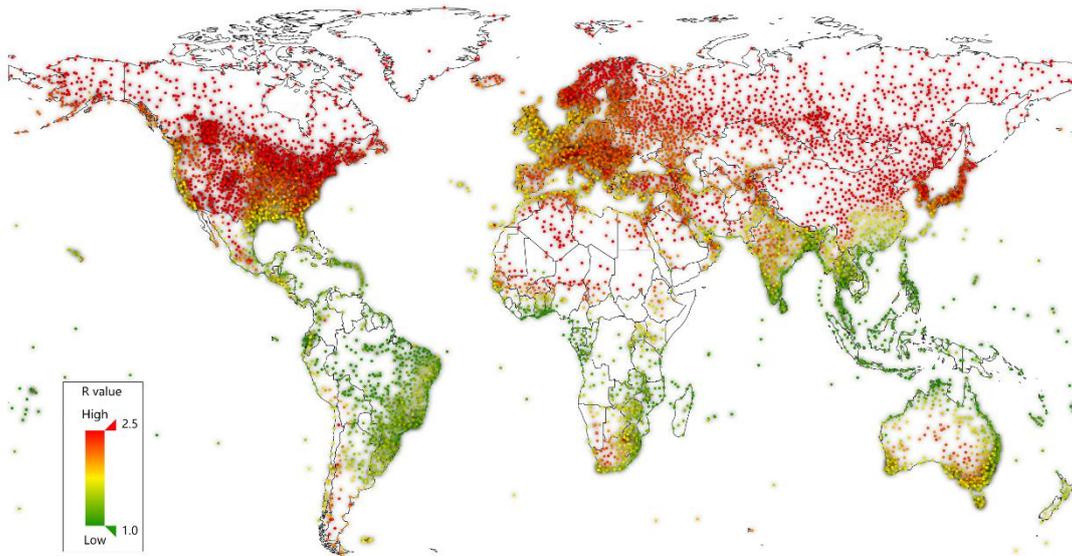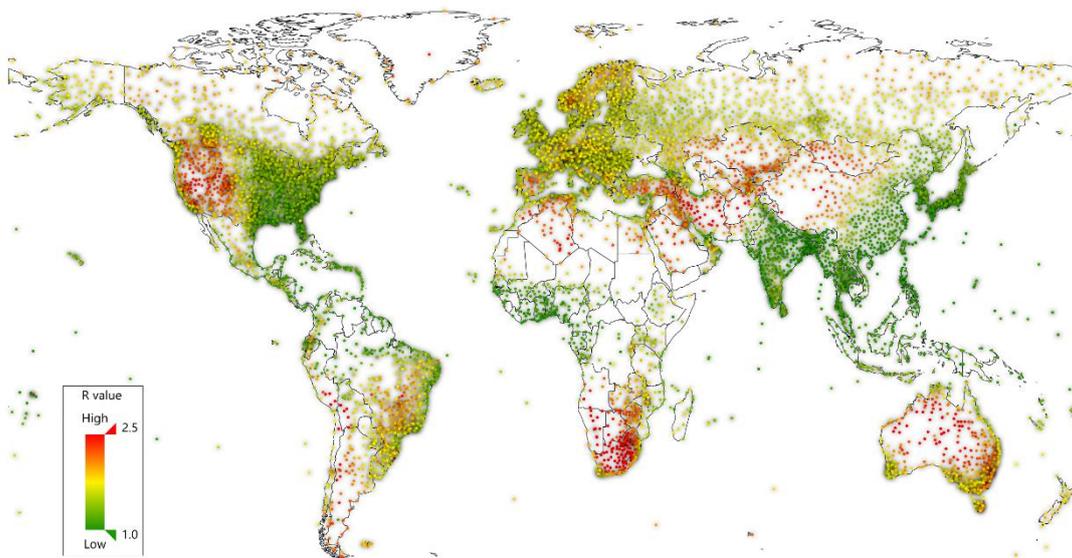
(a) *R* values in March

(b) *R* values in July

Figure 4: Worldwide risks of COVID-19 outbreak in March and July 2020

We use coefficients from the fixed-effects model of Table 2 to estimate *R* values of worldwide cities (represented by dots) for March and July 2020, respectively: $R = 3.968 - 0.0383\,Temperature - 0.0224\,RelativeHumidity$, where temperatures and relative humidity of March and July 2019 are obtained from https://www.ncdc.noaa.gov/ and assumed to be the same as those of 2020.

Table 1: Temperature, relative humidity and effective reproductive number: A cross-sectional regression analysis

This table reports the cross-sectional regression coefficients of the average effective reproductive number, *R,* on an intercept, average temperature, average relative humidity and two control variables (GDP per capita and population density). Average *R* values, average temperatures and relative humidity from January 21 to 23, 2020 are used in this regression. The regression is estimated by an Ordinary Least Square (OLS) method with robust standard errors. T-statistics are in the italic format with *, ** and *** representing significance at the 10%, 5% and 1% levels, respectively.

	OLS	OLS robust standard errors	OLS	OLS robust standard errors
Temperature	-0.0287	-0.0287	-0.0266	-0.0266
t-statistics	*-4.25****	*-3.14****	*-3.55****	*-2.91****
Relative Humidity	-0.00892	-0.00892	-0.0106	-0.0106
t-statistics	*-1.74**	*-1.65**	*-2.01***	*-1.85**
GDP per Capita			-0.0272	-0.0272
t-statistics			*-1.66**	*-1.66**
Population Density			0.0747	0.0747
t-statistics			*1.08*	*1.18*
const	2.802	2.802	3.061	3.061
t-statistics	*6.66****	*5.74****	*6.66****	*5.52****
R^2	22%	22%	24%	24%

Table 2: Temperature, relative humidity and effective reproductive number: A panel regression analysis

This table reports the panel regression coefficients of the effective reproductive number, *R,* on an intercept, temperature, relative humidity and two control variables (GDP per capita and population density). Daily *R* values, temperature and relative humidity from January 21 to 23, 2020 are used in this regression. Fixed and random effects models are both performed with robust standard errors. T-statistics are in the italic format with *, ** and *** representing significance at the 10%, 5% and 1% levels, respectively.

	Fixed Effects	Fixed Effects with Robust Errors	Random Effects	Random Effects with Robust Errors
Temperature	-0.0383	-0.0383	-0.024	-0.024
t-statistics	*-3.27****	*-2.16***	*-3.97****	*-3.06****
Relative Humidity	-0.0224	-0.0224	-0.020	-0.020
t-statistics	*-10.18****	*-10.15****	*-10.15****	*-9.72****
GDP per Capita			-0.031	-0.031
t-statistics			*-1.95**	*-1.92**
Population Density			0.078	0.078
t-statistics			*1.14*	*1.23*
const	3.968	3.968	3.877	3.877
t-statistics	*19.04****	*16.78****	*19.60****	*15.83****
R^2	17%	17%	19%	19%

Methods

1. Data

We hand-collect 4,711 cases from the epidemiological survey data available online published by the Center for Disease Control and Prevention of 11 provinces and municipalities including Beijing, Shanghai, Jilin, Sichuan, Hebei, Henan, Hunan, Guizhou, Chongqing, Hainan and Tianjin. By analyzing the records of each patient's contact history with other patients, we match 209 close contacts. Among them, if an infected patient comes into contact with several COVID-19 carriers, we choose the earliest contact; if a group of patients go to Wuhan City together, we hence cannot distinguish between the carrier and the infected, and, therefore, remove such samples from the data. We finally screen out 105 pairs of virus carriers and the infected, which are used to estimate serial intervals of COVID-19. We also construct epidemic curves for all100 Chinese cities with more than 40 cases from their first-case dates to February 20. The epidemic curves are used to estimate the daily effective reproductive number, R, for different cities.

Temperature and relative humidity data are obtained from 699 meteorological stations in China. If a city does not have a meteorological station inside it, the closest station is used instead. Population density and GDP per capita of 2018 for different cities are obtained from https://data.cnki.net.

2. Distribution of the serial interval

The serial interval, defined as the time span between symptom onset dates of a primary case to a successive case, is calculated based on the 105 samples of the carrier and the infected. Specifically, we fit the Weibull distribution[1, 2] using the Maximum Likelihood Estimation (MLE) method[8] and obtain the parameters of the mean and standard deviation of 7.4 and 5.2 days, respectively, which are consistent with the

---

[8] We fitted the Weibull distribution by Python package 'Scipy' and R package 'MASS', which can be found at https://www.scipy.org/ and https://cran.r-project.org/web/packages/MASS/index.html. The two results are consistent to each other.

14

preliminary estimation[3] using 10 cases (7.5 days average with 95% confidence interval of 5.3 to 19). Compared to SARS[2], the COVID-19's serial interval has a smaller average but a larger standard deviation. The fitted Weibull distribution is shown in Figure SI 1.

[Figure SI 1 about here.]

3. Estimation of the effective reproductive number

We estimate the daily effective reproductive number, R, for 100 cities with more than 40 cases from the first-case date to February 20 by employing a time-dependent method[4]. The inputs of the model are epidemic curves, i.e. the historical numbers of patients with symptom onset of each day for a certain city. We estimate the R values using a package 'R0'[4] (https://cran.r-project.org/web/packages/R0/index.html). In this package, we particularly use a function named 'est.R0.TD' in our estimation. We calculate the three-day averaged R from January 21 to 23 as the proxy of natural non-intervened effective reproductive number for each city. If there is no case on a certain day for a city, we skip that day and use the remaining days to get the average value. The average R value of these 100 cities is 1.9 with the minimum and maximum values of 1.0 and 4.6, respectively. Table SI 1 provides summary statistics of the variables used in this paper.

[Table SI 1 about here.]

4. Linear Regression

Table 1 reports the cross-sectional regression coefficients of the average effective reproductive number, R, on an intercept, the average temperature, the average relative humidity and two control variables (GDP per capita and population density). We estimate the parameters in the regression using the Ordinary Least Square (OLS) method. In Table 2, we run a panel regression with both fixed and random effects, respectively, using daily observations. All regressions are performed with the econometrics software *Stata*. Figure 3 shows slightly larger fitting errors in the low

15

temperature and low humidity range, as a robustness check, we, therefore, utilize the White's robust standard errors [5] to estimate the t-statistics in the regression.

Furthermore, among these 100 cities, Wuhan is a special sample because of the double standards for the confirmation of cases. For example, there was a sudden increase of more than 13,000 cases in a single day (February 12, 2020) in Wuhan, and the majority of them were previously left unable to seek medical treatment. Therefore, as a robustness check, we remove Wuhan city in our sample and redo both the cross-sectional and panel regressions. The results of robustness checks, presented in Table SI 2, are consistent with those in Table 1 and 2.

[Table SI 2 about here.]

References

1. Cowling, B. J. et al. Estimation of the serial interval of influenza. *Epidemiology* **20**, 344 (2009).

2. Lipsitch, M. et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science* **300**, 1966–1970 (2003).

3. Li, Q. et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMoa2001316 (2020).

4. Wallinga, J. & Teunis, P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am. J. Epidemiol.* **160**, 509–516 (2004).

5. Wooldridge, J, M. Introductory Econometrics: A Modern Approach, 3rd edition. 276 (*Nelson Education,* 2016).
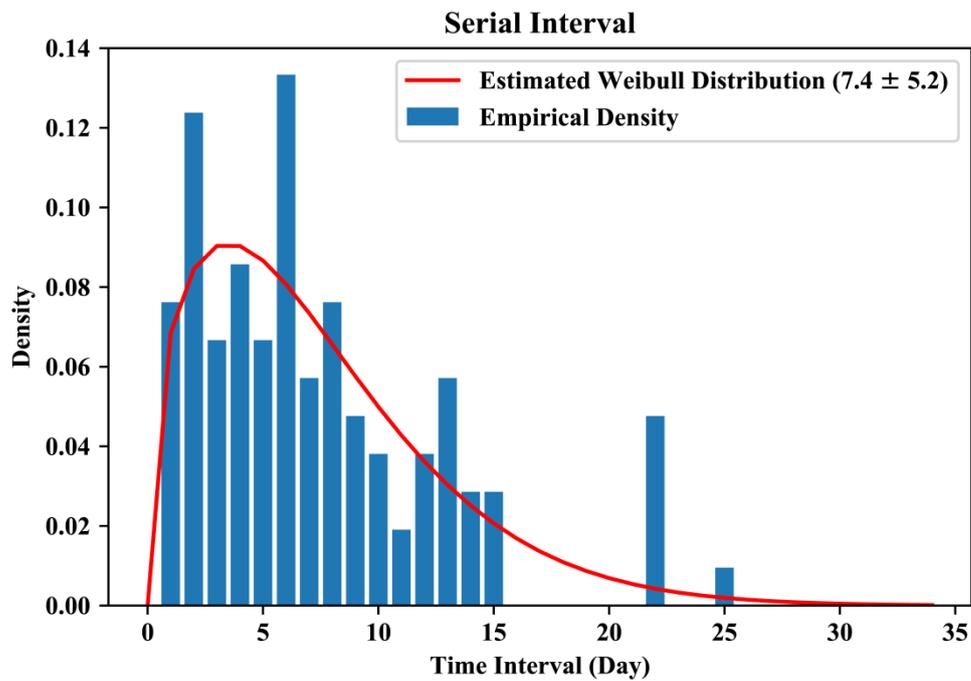
Figure SI 1: Estimation of the serial interval with the Weibull distribution

Bars denote the probability of occurrences in specified bins, and the red curve is the density function of the estimated Weibull distribution.

17

Table SI 1: Data Summary

This table summarizes variables for 100 cities: the average *R* value, average temperature and relative humidity from January 21 to 23, 2020, and the GDP per capita and the population density of 2018.

	Mean	**Std**	**Min**	**Max**
Average *R*	1.881	0.527	1.031	4.609
Average Temperature (Celsius)	6.183	7.283	-17.933	21.033
Average Relative Humidity (%)	83.340	9.567	46.667	100.0
GDP per Capital (RMB 10k)	6.771	3.762	1.387	18.594
Population Density (k/km^2)	0.723	0.879	0.00657	6.671

Table SI 2: Relationship between Temperature, relative humidity, and effective reproductive number for samples without Wuhan

The table reports the linear regression coefficients of the effective reproductive number, *R,* on an intercept, temperature, relative humidity and two control variables (GDP per capita and population density) for samples without Wuhan. Both cross-sectional and panel regressions in Table 1 and 2 are performed. T-statistics are in the italic format with *, ** and *** representing significance at the 10%, 5% and 1% levels, respectively.

	Cross-sectional	Cross-sectional	Panel Fixed Effects	Panel Random Effects
Temperature	-0.0287	-0.0266	-0.0380	-0.0244
t-statistics	*-4.24****	*-3.52****	*-3.23****	*-3.93****
Relative Humidity	-0.00876	-0.0106	-0.0224	-0.0203
t-statistics	*-1.70**	*-1.97***	*-10.13****	*-10.10****
GDP per Capita		-0.0269		-0.0316
t-statistics		*-1.60*		*-1.94**
Population Density		0.0745		0.0782
t-statistics		*1.07*		1.13
const	2.791	3.056	3.965	3.879
t-statistics	*6.59****	*6.54****	*18.95****	*19.44****
R^2	22%	24%	17%	19%