medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Preprint, Ver. 1 (May 5, 2020) - *the manuscript has not been peer-reviewed yet.*

Weather Conditions and COVID-19 Transmission: Estimates and Projections

Ran Xu, PhD[1,†], Hazhir Rahmandad, PhD[2,†], Marichi Gupta[3], Catherine DiGennaro[3], Navid Ghaffarzadegan, PhD[4], Mohammad S. Jalali, PhD[3,5,*]

[1]Department of Allied Health Sciences, University of Connecticut, Storrs, CT
[2]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA
[3]Institute for Technology Assessment, Massachusetts General Hospital, Boston, MA
[4]Department of Industrial and Systems Engineering, Virginia Tech, Falls Church, VA
[5]Harvard Medical School, Harvard University, Boston, MA

[†] These authors contributed equally to this study.

***Corresponding author:** msjalali@mgh.harvard.edu
Data, codes, and online simulator: https://projects.iq.harvard.edu/covid19

Abstract

Background: Understanding and projecting the spread of COVID-19 requires reliable estimates of how weather components are associated with the transmission of the virus. Prior research on this topic has been inconclusive. Identifying key challenges to reliable estimation of weather impact on transmission we study this question using one of the largest assembled databases of COVID-19 infections and weather.

Methods: We assemble a dataset that includes virus transmission and weather data across 3,739 locations from December 12, 2019 to April 22, 2020. Using simulation, we identify key challenges to reliable estimation of weather impacts on transmission, design a statistical method to overcome these challenges, and validate it in a blinded simulation study. Using this method and controlling for location-specific response trends we estimate how different weather variables are associated with the reproduction number for COVID-19. We then use the estimates to project the relative weather-related risk of COVID-19 transmission across the world and in large cities.

Results: We show that the delay between exposure and detection of infection complicates the estimation of weather impact on COVID-19 transmission, potentially explaining significant variability in results to-date. Correcting for that distributed delay and offering conservative estimates, we find a negative relationship between temperatures above 25 degrees Celsius and estimated reproduction number ($\hat{R}$), with each degree Celsius associated with a 3.1% (95% CI, 1.5% to 4.8%) reduction in $\hat{R}$. Higher levels of relative humidity strengthen the negative effect of temperature above 25 degrees. Moreover, one millibar of additional pressure increases $\hat{R}$ by approximately 0.8 percent (95% CI, 0.6% to 1%) at the median pressure (1016 millibars) in our sample. We also find significant positive effects for wind speed, precipitation, and diurnal temperature on $\hat{R}$. Sensitivity analysis and simulations show that results are robust to multiple assumptions. Despite conservative estimates, weather effects are associated with a 43% change in $\hat{R}$ between the 5th and 95th percentile of weather conditions in our sample.

Conclusions: These results provide evidence for the relationship between several weather variables and the spread of COVID-19. However, the (conservatively) estimated relationships are not strong enough to seasonally control the epidemic in most locations.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Introduction

The COVID-19 pandemic has significantly challenged the global community. High-stakes policy decisions require projections of the course of the pandemic across different geographic regions. Thus, it is critical to know how weather conditions impact the transmission of the disease [1]. Given that many related viral infections such as seasonal flu [2, 3], MERS [4-6], and SARS [7] show notable seasonality, one may expect the transmission of SARS-CoV-2 virus to be similarly dependent on weather conditions. Earlier works indicate that temperature, humidity, air pressure, ultraviolet light exposure, and precipitation potentially impact the spread of COVID-19 through changing the survival times of the virus on surfaces and in droplets, moderating the distances the virus may travel through air, and impacting individual activity patterns and immune responses [8-12].

Yet, there is limited agreement on the shape and magnitude of those relationships. While many studies find a correlation between variations in temperature [13, 14], relative and absolute humidity [15-23], ultraviolet light [24], and wind speed, visibility, and precipitation [25-27] with measures of pandemic severity [28], other works [24, 29-31] indicate weaker, inconsistent, or no relationships. A recent review finds inconclusive evidence for the relevance of weather in the transmission of COVID-19 [1].

It is unclear what explains the inconclusive results. Earlier samples based on datasets focused on China may be too narrow [16, 19, 24, 32]. Others have studied only a subset of weather components that complicate comparisons [18, 19, 33]. Most have not controlled for other important correlates such as government and public responses, population density, and cultural practices [15-17, 24-26, 32, 34]. Another especially under-studied factor is the delays between infection and official recording of cases. These delays, estimated to be close to 10 days [35, 36], confound attempts to associate daily weather conditions with recorded new cases. Extending time windows over which transmission trends are calculated may partially address this challenge, but significantly reduces the number of data points, complicates interpretation and projections, and increases the risk that results are driven by spurious correlation between weather and location specific factors and trends. Therefore, lack of corrections for these delays may partially explain the inconsistent and inconclusive findings to date.

This paper assembles one of the most comprehensive datasets of the global spread of COVID-19 pandemic until late April 2020, builds and validates a statistical method for the estimation of reproduction number controlling for detection delay, location-specific density, and time variant responses, estimates the association of weather conditions and the reproductive number of COVID-19, and provides year-round, global projections.

Data and Methods

Data

To track infections, we use the official case reports from various countries. Our starting point is the data collected and compiled by the Johns Hopkins Center for Systems Science [37]. We augment these data with those coming from the Chinese Center for Disease Control and Prevention, Provincial Health Commissions in China, and Iran's state-level reports. The data from the beginning of the epidemic (December 12, 2019) to April 22, 2020 are used for our analysis. We assemble disaggregate data on the spread of COVID-19 in Australia (8 states), Canada (10 states), China (34 province-level administrative units and also 301 individual cities),

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Iran (31 states), and the United States (3144 counties and 5 territories) and use country level aggregates for the rest of the world. Overall, our dataset includes cumulative infection data for 3739 distinct locations.

We use the Historical Weather database (World Weather Online and OpenWeather Ltd., 2020) to compile weather data. For each location for which we have infection data, we use the weather data for the longitude and latitude of the centroid of that location. We collected minimum and maximum daily temperature, humidity, precipitation, snowfall, moon illumination, sun hour, ultraviolet index, cloud cover, wind speed and direction, and pressure data; however, only a subset of these variables proved relevant in our analysis. We used population density data from Demographia (Cox., W, *Demographia*, The Public Purpose), the United States Census (U.S. Census Bureau, data.census.gov/cedsci), the Iran Statistical Centre, the United Nation's Projections, *City Population* (citypopulation.de), and official published estimates for countries not covered by these sources. We also projected contagion risk forward in our simulations for the list of highly-populated cities.

Methods

Estimation of the Reproduction Number

A critical parameter in understanding the spread of an epidemic is the basic reproduction number, R_0, which reflects the number of secondary cases generated by an index patient in a fully susceptible population. An epidemic is expected when R_0 is above 1 and will die out with R_0 values below 1.

Measuring R_0 directly requires data that are often not available at scale [38]. However, early in the epidemic before the stock of susceptible populations is depleted, R_0 is close to the reproduction number, R_e, which measures the number of secondary cases from each infected case. Reproduction number can be approximated ($\hat{R}$) based on the number of new infections (I_M) per currently infected individual, multiplied by the duration of illness (τ) (Equation 1). Actual new infections on any day (I_N) are not directly observable and should be estimated. Data on measured daily infections (I_M) lags actual new infections by both the incubation period and the delay between onset of symptoms and testing and recording of a case. We use published measures to quantify the distribution of incubation period (averaged between 5 to 6 days [25, 35, 36, 39]) and onset-to-detection delay (ranging between 4 to 6 days [35, 36]) to quantify the overall *detection delay*. Given the variance in detection delay, a simple shift of measured infection by the mean delay (about 10 days) offers an unreliable estimate for true infections (See Section S3 and S5.2.2.2 in Supplementary Document). We therefore use an optimization to find the daily estimated new infections ($\hat{I}_N$) that are consistent with the observed measured infections (I_M) and the overall detection delay distribution (See Section S3 for details).

$$\hat{R}(t) = \frac{\hat{I}_N(t)\tau}{\sum_{s=t-\tau-1}^{s=t-1} \hat{I}_N(s)} \quad (1)$$

We use the $\hat{R}(t)$ as our dependent variable. This formula is robust to existence of asymptomatic cases and under-counting (which is likely due to imperfect test coverage) as long as the changes in test coverage are not correlated with weather conditions 10 days ago (see Sections S3 and S5.2.2.3 in Supplementary Document). We use a delay of τ=20 days from exposure to resolution and results are robust to other durations of illness (see Section S4.1 in Supplementary Document). For each location we only include days with $\hat{I}_N$ values above one. $\hat{I}_N$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

values proceed actual detection, and thus the reliability of early values for each location is affected by irregularities in early testing. Moreover, an unbiased estimate for $\hat{R}(t)$ requires τ days of prior new infection estimates. Thus, to ensure robustness we separately exclude the first 20 days after the $\hat{I}_N$ reaches one in each location. Robustness to these exclusion criteria and exclusion of outliers based on $\hat{R}(t)$ values are discussed in Section S4.2 and S4.3 in Supplementary Document.

Independent Variables

Prior research [15, 23, 40] suggests various weather-related factors may have a role in the survival of the virus on surfaces, spread of droplets containing the virus, as well as behaviors and responses of human hosts relevant for understanding reproduction number. We therefore include the following daily variables as predictors: temperature (mean ($\bar{T}$) and diurnal temperature (difference between maximum and minimum daily temperature; ΔT), both in Celsius), relative humidity (H) (percentage), pressure (P) (millibars), precipitation (C) (millimeters), snowfall (S) (centimeter), wind speed (W) (km/hour), and number of hours of sun received per day (U). We also explore a few interactions among these variables.

Controls

Different locations vary in their reproduction number due to various location-specific factors. While location fixed effects can control for many of those variations, our simulation studies show that fixed effects, combined with measurement error inherent in inferring $\hat{I}_N$ from reported infections, may significantly attenuate and bias weather effects. Empirically we do not find such sensitivity (See supplement S4.2) and results are consistent with or without fixed effects. Nevertheless, we report fixed effect results in the supplement and in our preferred specification control for log-transformed population density. Moreover, public gathering bans, school closures, physical distancing, and other responses to the COVID-19 outbreak endogenously changes R_e in each location and could confound our findings. To account for these changes, we include a location-specific linear trend in predicting $\hat{R}(t)$.

Statistical Specification and Validation

Given the large variations in $\hat{R}$ values estimated in this method, we use a log transformation of $\hat{R}(t)$ and linear models to predict ($\ln(\hat{R})$). Thus, the exponential of each estimated effect is the multiplier changing $\hat{R}$ around its base-value indicated by the location-specific density, policies and responses. We designed and validated our statistical model for estimating $\ln(\hat{R})$ by testing its ability to identify true parameters in synthetic data. Specifically, we built a stochastic simulation model of COVID-19 epidemic, generated synthetic infection data using historical weather inputs and known impact functions, and fine-tuned the statistical model until it could reliability identify true effects in a large set of simulated epidemics. We found that: a) Whereas given actual infections (I_N), our method, with or without fixed effects, would identify the true functional form relating temperature to transmission rate, estimates become less reliable and significantly attenuated when true infections are inferred rather than exact; b) Our method of estimating true infections (see Sections S3 and S5.2.2.2 in Supplementary Document) offers significantly better results than simple shifting of official infection counts; and c) The most reliable results in synthetic epidemics are found when location-specific trends (but not fixed effects) are used in model specification.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Separately, authors NG and MG created a more realistic individual-based model of disease transmission and used that to generate a separate test dataset with synthetic epidemics. Three scenarios used actual temperatures from a sample of 100 regions and three different functions for temperature effect (including the placebo scenario of no effect). Then author RX, who was blinded to the true specification of temperature effect in the latter synthetic dataset, was able to approximate the true function with correct qualitative shape using the refined statistical method. Details on this test are reported in the Section S5.3 in Supplementary Document.

In short, the preferred specification includes location-specific trends, excludes days with $\hat{I}_N$<1, and the first 20 days after $\hat{I}_N$ exceeds one for the first time, and includes the following main and interaction effects: linear spline effect of average temperature with the knot at 25 degrees (see Section S4.5 for alternative knot values), diurnal temperature, relative humidity, quadratic effect of pressure (transformed: X - 1000), precipitation (transformed: ln(X+1)), snowfall (transformed: ln(X+1)), wind speed (transformed: ln(X+1)), number of hours of sun received per day, and interaction between relative humidity and average temperature.

While more complex specifications including nonlinear models could be used, we opted for a simpler and more theoretically-driven alternative for two reasons. First, location-specific trends account for much of the predictive power of the model, and thus, fine-tuning weather terms using cross-validation does not offer significant improvements in predictive power. Second, effects from more complex models would be harder to interpret and communicate, and less reliable to extrapolate.

Projection

After calculating a weather response function using existing data, we project the impact of weather conditions on the relative future risk of pandemic for all locations in our sample as well as the major urban areas of the world (those with more than 0.5 million population as of 2017; a total of 1072 cities) that constitute about 30% of the world population. A summary of results is provided in the main paper and the online appendix and an interactive online platform offer details.

Results

Estimated impact of weather factors on transmission rate

Table 1 reports our main regression results. This statistical model closely tracks $\ln(\hat{R})$ values (R^2=.534). Much of the accuracy is due to location-specific trends. As expected, (log) population density is strongly associated with $\ln(\hat{R})$ (β=.179, 95% CI, .159 to .199). Most locations show rapid reductions in reproduction number captured in the location-specific trends. Specifically, on average $\hat{R}$ goes down by 4.2% (SD=2.8%) per day since estimated number of new infections exceeds 1 for the first time in each location. Mean temperature ($\bar{T}$), diurnal temperature (ΔT), air pressure (P), wind speed (W), snowfall (S) and precipitation (C) are significant predictors of transmission.

Effect of mean daily temperature ($\bar{T}$) is best characterized within two regimes, below and above 25 degrees Celsius, and in interaction with humidity. Temperatures higher than 25 are associated with lower transmission rates (by 3.1% (95% CI, 1.5% to 4.8%) per degree Celsius; excluding interaction with H) while those below that threshold have a smaller impact (0.5% (95% CI, 0.29% to 0.77%) reduction in transmission rates per degree Celsius). The negative effect of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

temperatures above 25 degrees is strengthened at higher levels of relative humidity—10% increase in relative humidity is associated with an additional 1.2% (95% CI, 0.6% to 1.7%) decrease in transmission rates for each one degree increase in temperature above 25 degrees. Figure 1-A provides a graphical summary of the joint estimated relationship.

Air pressure (P) has a weak U-shaped effect on the reproduction number, with minimum around 1,004 millibars. One millibar of pressure increases the estimated reproduction number by approximately 0.8% (95% CI, 0.6% to 1%) at the median pressure (1016 millibars) in our sample. We also find weak but significant positive effects of diurnal temperature, precipitation, snowfall and wind speed on reproduction number. With one standard deviation increase in diurnal temperature (ΔT), log-transformed precipitation (P), log-transformed snowfall (S) and log-transformed wind speed (W), the estimated reproduction number increases by 1.9% (95% CI: 0.6% to 3.1%), 2.9% (95% CI: 1.6% to 4.2%), 1.4% (95% CI: 0.3% to 2.5%), and 3.8% (95% CI, 2.5% to 5.1%) respectively.

Overall, the association of various weather variables with COVID-19 transmission is moderate and potentially relevant to assessing the risk of contagion in different locations across the globe. Figure 1-B provides a histogram of the variations in relative $\hat{R}$ associated with the combined set of weather variables in our estimation sample. The ratio between any two values can be interpreted as the ratio in $\hat{R}$ due to differences in weather, all else equal. The gap between the 5th and 95th percentile in this distribution indicates a change in relative $\hat{R}$ by 43%. Given that the typical reproduction number estimated for COVID-19 is in the range of 2 to 3 [41, 42], estimated weather effects alone, may not provide a path to containing the epidemic in most locations, but could notably impact the relative transmission rates.

Table 1. Impact of weather on COVID-19. Outcome variable: $ln(\hat{R})$

Weather variables	Mean (SD)	Coefficient (95% CI)	p-value	Standardized coefficient*
Wind speed	2.53 (.472)	.078 (.052, .104)	<.001	.0369
Snowfall	.08 (.352)	.040 (.010, .070)	.01	.0139
Precipitation	.779 (1.03)	.027 (.015, .039)	<.001	.0282
Air Pressure	15.36 (6.03)	-.003 (-.008, .003)	.359	-.0154
Air pressure^2	272.23 (191.17)	.0003 (.0002, .0005)	<.001	.0633
Humidity	66.97 (16.33)	-.001 (-.003, .0001)	.068	-.0234
Mean temperature below 25	11.11 (7.22)	-.016 (-.023, -.008)	<.001	-.1236
Humidity* Mean temperature below 25		.0001 (.00004, .0002)	.004	.0854
Mean temperature above 25	28.29 (2.53)	.025 (-.007, .056)	.129	.0292
Humidity* Mean temperature above 25		-.001(-.002, -.0005)	<.001	-.0671
Sun hour	9.53 (2.15)	.002 (-.004, .008)	.463	.0048
Diurnal temperature	9.13 (3.43)	.005 (.002, .009)	.004	.0184
N=19,140; R^2=.534				

*Standardized coefficients were obtained by first standardizing all of the weather variables (mean=0, SD=1) and then re-running the analysis with our main model specification.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

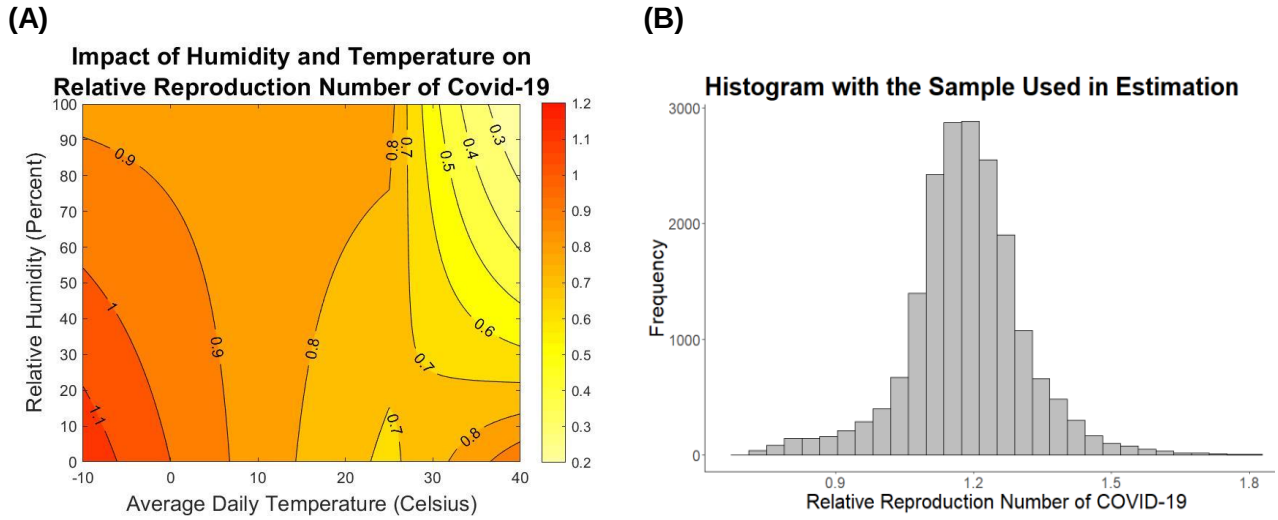


Figure 1. Summary of results. **A)** Association of average temperature, humidity, and relative $\hat{R}$ (excluding other effects). **B)** Histogram for predicted relative $\hat{R}$ as a result of the combined effect of weather variables in our sample.

Robustness

Validation of our statistical method using synthetic data showed that our method is capable of identifying the correct sign and shape for the impact of weather variables and that those estimates are potentially conservative (i.e., downplaying the true impacts). For example, the estimated curvatures in our synthetic data are about 20% of true values (See Section S5.2.2 in Supplementary Document). The conservatism is mainly due to unavoidable errors in estimating daily infections from lagged official data. As a result, weather variables are not perfectly matched with the true infection rates on the corresponding days, weakening any estimated relationship. If not carefully corrected for, this challenge can lead to incorrect and misleading estimates across a variety of specifications (see Section S5.2.2.2).

We also conducted six different tests to assess the robustness of our finding. First, our results do not change with the use of different illness durations to calculate $\hat{R}$ (Table S1 in section S4.1). Second, our main findings are robust to excluding extreme values of the dependent variables, last few days of data, or including location fixed effects in the analysis (Table S2 in section S4.2). Third, our results are largely insensitive to different exclusion criteria for initial periods (Tables S3 and S4 in section S4.3). Fourth, a placebo test where weather variables in each location are permuted and shifted by a random number of days shows no effect of the weather variables in most of our model specifications, adding confidence that results cannot be attributed to mechanical features of the statistical model (Table S5 in section S4.4). Finally, we found a significant and qualitatively relevant negative effect of moon illumination on reproduction number (which also did not change other weather effects; see Table S7 in section S4.6). Absent theoretical explanations, we decided not to include that effect in the main specification, but we find it worth further exploration.

Projections

The associations between weather variables and transmission rates highlight the potential to project the risk of COVID-19 spread as a function of weather conditions. Of course, our results

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

are associative and extrapolating out-of-sample includes unknown risks. With that caveat in mind, one could calculate the weather components' contribution for any vector of weather variables based on Table 1. Exponentiating that value (as we do in Figure 1-B) leads to a multiplicative weather score which is associated with reproduction rate independent of the location-specific characteristics and responses. Such scores are therefore comparable across locations and the ratio between two scores offers a measure of relative transmission risk for two vectors of weather variables. We define "Relative COVID-19 Risk Due to Weather" (CRW) as the relative predicted risk of each weather vector against the 90th percentile of predicted risk in our estimation sample, 1.33 (Figure 1-B). The choice of this reference point is somewhat arbitrary and is made to make a value of 1 a rather high risk of transmission due to weather. A CRW of 0.5 reflects a 50% reduction in the estimated reproduction number compared to the reference weather condition. Formally:

$$CRW = \exp[0.0782W + 0.0395S + 0.0273C - 0.00255P + 0.000331P^2 - 0.00143H - 0.0156 * \min[\bar{T}, 25] + 0.000145 \min[\bar{T}, 25]\, H + 0.0247(\max[\bar{T}, 25] - 25) - 0.00116(\max[\bar{T}, 25] - 25)H + 0.00537\Delta T + 0.00224U]/1.33 \quad (2)$$

It is important to note that these scores do not reveal the actual values of reproduction number; that value is highly contingent on location-specific factors and policies, for which we have no data outside of the estimation sample. For example, the COVID-19 reproduction number in New York City is likely larger than a rural district in upstate New York. Our projections cannot inform the absolute risks for either location. CRW scores could be compared to inform relative risks due to weather (i.e., assuming all else equal) across locations or within a location over time.

Figure 2 provides a visual summary of global CRW scores, averaged over the first half of June 2020. The color-coded scores suggest much variation could be expected in the risk of COVID-19 transmission across different locales, with many regions in the Middle East, Africa, Asia, and North America potentially benefiting from rather low CRW values while Canada, Europe, and Australia show continued excess risk due to weather. In Section S6 in the Supplementary Document, we provide additional snapshots of global CRW scores at different times of the year, and the website (projects.iq.harvard.edu/covid19) offers week-by-week risk measures year-round.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

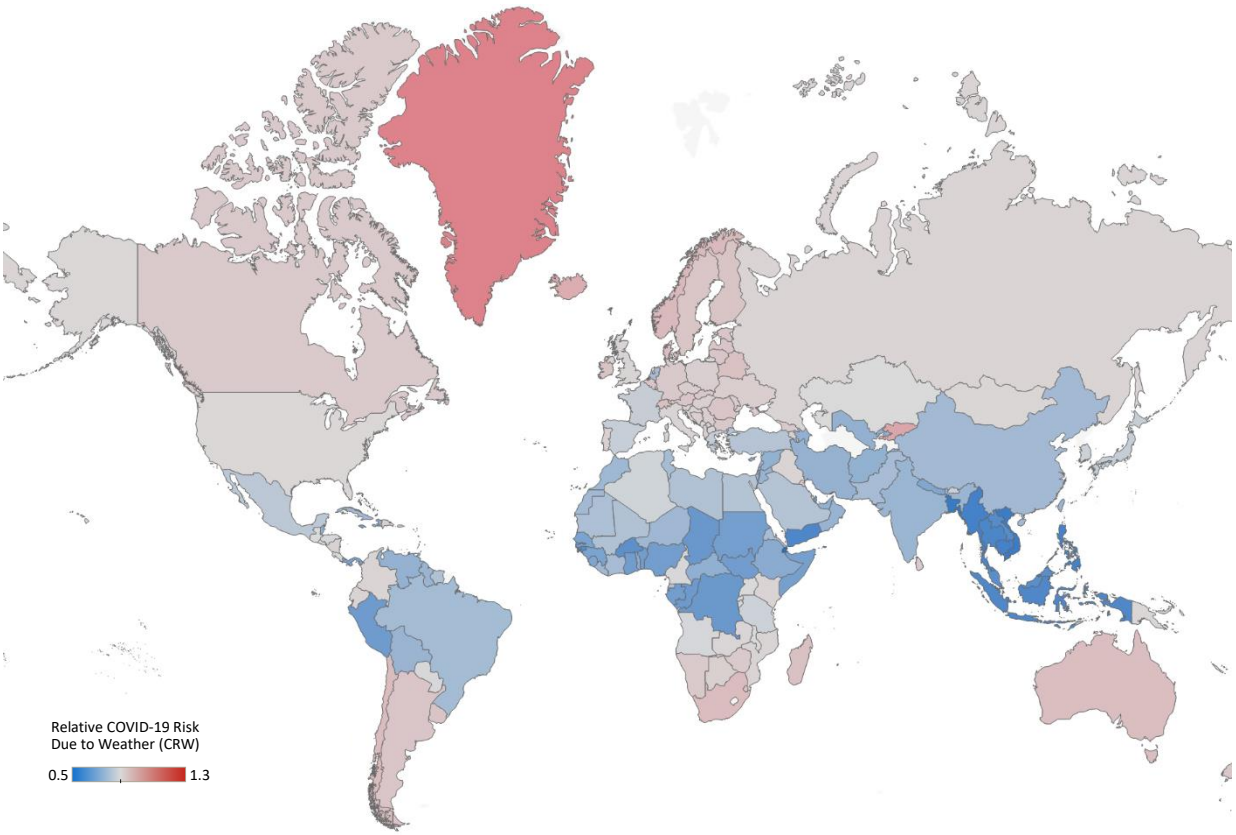


Figure 2. Relative COVID-19 Risk Due to Weather (CRW) for different regions of the world, averaged over the first half of June 2020.

Figure 3 shows CRW projections for five major cities in each of the four regions of America (panel A), Europe (B), Africa and Oceania (C), and Asia (D). These projections use weather data from 2019-2020, averaged over a 15-day moving window, for 2020-2021 dates; as such, they include historical noise despite averaging. Many of these large cities go through periods of higher risk as well as reduced risk during the year. As discussed before, we cannot associate these risks with absolute reproduction numbers, and our estimates are likely conservative. Nevertheless, assuming typical reproduction rates in the 2-3 range, one needs CRWs below 0.3 to contain the epidemic based on weather factors, a condition rarely observed in our data. The website projects.iq.harvard.edu/covid19 provides these projections for the 1,072 largest global cities.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

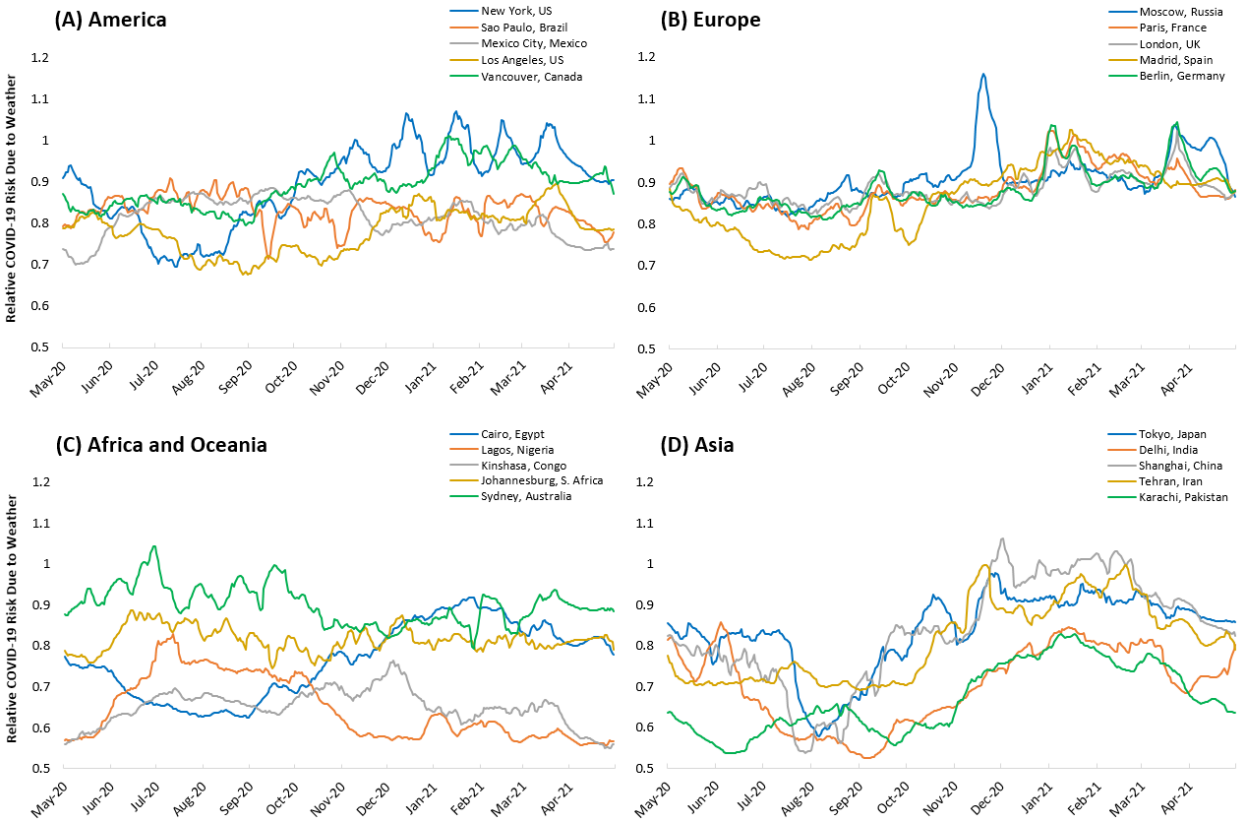


Figure 3. CRW measures over the year for major cities around the world.

Discussion

Combining one of the most comprehensive datasets of COVID-19 transmission to date with weather data across the world, this paper provides evidence for the association of various meteorological variables with the spread of COVID-19.

While we find a stronger effect (A 3.1% (95% CI, 1.5% to 4.8%) reduction in $\hat{R}$ for every degree increase) for temperatures above 25 degrees Celsius, the relatively mild slope of temperature effect below 25 degrees suggests many temperate zones with large population density may face larger risks, while some warmer areas of the world may experience slower transmission rates. For example, the estimated associations may partially explain the smaller sizes of outbreaks in southern Asia and Africa to date.

We also show that the challenge of estimation due to detection delay is significant and at best conservative estimates may be expected from standard regression methods. This observation may partially explain inconclusive and inconsistent prior results. Overcoming this challenge not only requires careful estimation of true infections based on reported cases, but also may benefit from methods that correct for the resulting bias, e.g., indirect inference [43]. Other limitations of the study include: the variance in reliability and availability of transmission data across the globe; oversampling from U.S. locations; use of last year's weather data to project next year's outcomes; and use of correlational evidence to inform out-of-sample projections.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Despite these limitations, consistent results using various conservative specifications and validation tests are promising indications of true impacts of weather conditions. These effects may offer partial relief to some regions of the world during the summer, but it is important for policymakers and citizens to remain vigilant in their responses to the pandemic, rather than assuming that summer climate naturally prevents transmission. In fact, much of the variation in reproduction number in our sample is explained by location-specific responses, not weather. Ultimately, weather much more likely plays a secondary role in the control of the pandemic.

Declaration of Interest:

The authors declare no competing interests.

Acknowledgements:

We thank Peiyi Li, Hongjin Xu, and Wenhan Dai who assisted us in collecting and evaluating data from China, and Maedeh Moghadam who helped us with assembling data from Iran.

Funding:

No funding was used to conduct this study.

References:

1. National Academies of Sciences, E., and Medicine, *Rapid Expert Consultation on SARS-CoV-2 Survival in Relation to Temperature and Humidity and Potential for Seasonality for the COVID-19 Pandemic*. 2020: Washington, DC.
2. Shaman, J., E. Goldstein, and M. Lipsitch, *Absolute Humidity and Pandemic Versus Epidemic Influenza.* American Journal of Epidemiology, 2010. **173**(2): p. 127-135.
3. Lowen, A.C., et al., *Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature.* PLOS Pathogens, 2007. **3**(10): p. e151.
4. Nassar, M.S., et al., *Global seasonal occurrence of middle east respiratory syndrome coronavirus (MERS-CoV) infection.* Eur Rev Med Pharmacol Sci, 2018. **22**(12): p. 3913-3918.
5. Altamimi, A. and A. Ahmed, *Climate factors and incidence of Middle East respiratory syndrome coronavirus.* J Infect Public Health, 2019. **In press.**
6. van Doremalen, N., T. Bushmaker, and V.J. Munster, *Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions.* Euro Surveill, 2013. **18**(38).
7. Yuan, J., et al., *A climatologic investigation of the SARS-CoV outbreak in Beijing, China.* American Journal of Infection Control, 2006. **34**(4): p. 234-236.
8. Casanova, L.M., et al., *Effects of air temperature and relative humidity on coronavirus survival on surfaces.* Applied and environmental microbiology, 2010. **76**(9): p. 2712-2717.
9. van Doremalen, N., et al., *Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1.* N Engl J Med, 2020.
10. Kampf, G., et al., *Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents.* Journal of Hospital Infection, 2020. **104**(3): p. 246-251.
11. Chin, A.W.H., et al., *Stability of SARS-CoV-2 in different environmental conditions.* The Lancet Microbe.
12. Chin, A., et al., *Stability of SARS-CoV-2 in different environmental conditions.* medRxiv, 2020.
13. Notari, A., *Temperature dependence of COVID-19 transmission.* medRxiv, 2020: p. 2020.03.26.20044529.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

14. Ficetola, G.F. and D. Rubolini, *Climate affects global patterns of COVID-19 early outbreak dynamics.* medRxiv, 2020.
15. Sajadi, M.M., et al., *Temperature, Humidity and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19.* Preprint, 2020.
16. Shi, P., et al., *The impact of temperature and absolute humidity on the coronavirus disease 2019 (COVID-19) outbreak - evidence from China.* medRxiv, 2020: p. 2020.03.22.20038919.
17. Oliveiros, B., et al., *Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases.* medRxiv, 2020: p. 2020.03.05.20031872.
18. Ficetola, G.F. and D. Rubolini, *Climate affects global patterns of COVID-19 early outbreak dynamics.* medRxiv, 2020: p. 2020.03.23.20040501.
19. Qi, H., et al., *COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis.* Science of The Total Environment, 2020. **728**: p. 138778.
20. Baker, R.E., et al., *Susceptible supply limits the role of climate in the COVID-19 pandemic.* medRxiv, 2020: p. 2020.04.03.20052787.
21. Wang, J., et al., *High temperature and high humidity reduce the transmission of COVID-19.* Available at SSRN 3551767, 2020.
22. Islam, N., S. Shabnam, and A.M. Erzurumluoglu, *Temperature, humidity, and wind speed are associated with lower Covid-19 incidence.* medRxiv, 2020: p. 2020.03.27.20045658.
23. Luo, W., et al., *The role of absolute humidity on transmission rates of the COVID-19 outbreak.* medRxiv, 2020.
24. Merow, C. and M.C. Urban, *Seasonality and uncertainty in COVID-19 growth rates.* medRxiv, 2020: p. 2020.04.19.20071951.
25. Li, Q., et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia.* N Engl J Med, 2020. **382**(13): p. 1199-1207.
26. Bu, J., et al., *Analysis of meteorological conditions and prediction of epidemic trend of 2019-nCoV infection in 2020.* medRxiv, 2020: p. 2020.02.13.20022715.
27. Araujo, M.B. and B. Naimi, *Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate.* medRxiv, 2020: p. 2020.03.12.20034728.
28. Araujo, M.B. and B. Naimi, *Spread of SARS-CoV-2 Coronavirus likely to be constrained by climate.* medRxiv, 2020.
29. Lipsitch, M., *Seasonality of SARS-CoV-2: Will COVID-19 go away on its own in warmer weather?*, in *https://ccdd.hsph.harvard.edu/will-covid-19-go-away-on-its-own-in-warmer-weather/*. 2020.
30. Briz-Redón, Á. and Á. Serrano-Aroca, *A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain.* Science of The Total Environment, 2020. **728**: p. 138811.
31. Briz-Redón, Á. and Á. Serrano-Aroca, *A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain.* Science of The Total Environment, 2020: p. 138811.
32. Guo, X.-J., H. Zhang, and Y.-P. Zeng, *Transmissibility of COVID-19 and its association with temperature and humidity*. 2020.
33. Coelho, M.T.P., et al., *Exponential phase of covid19 expansion is not driven by climate at global scale.* medRxiv, 2020: p. 2020.04.02.20050773.
34. Bukhari, Q. and Y. Jameel, *Will coronavirus pandemic diminish by summer?* Available at SSRN 3556998, 2020.
35. Lauer, S.A., et al., *The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.* Ann Intern Med, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

36. Linton, N.M., et al., *Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data.* J Clin Med, 2020. **9**(2).
37. Dong, E., H. Du, and L. Gardner, *An interactive web-based dashboard to track COVID-19 in real time.* The Lancet Infectious Diseases.
38. Paul, L.D., et al., *Complexity of the Basic Reproduction Number (R0).* Emerging Infectious Disease journal, 2019. **25**(1): p. 1.
39. Guan, W.J., et al., *Clinical Characteristics of Coronavirus Disease 2019 in China.* N Engl J Med, 2020.
40. Cai, Y., et al., *The Effects of "Fangcang, Huoshenshan, and Leishenshan" Makeshift Hospitals and Temperature on the Mortality of COVID-19.* medRxiv, 2020: p. 2020.02.26.20028472.
41. Zhang, S., et al., *Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis.* Int J Infect Dis, 2020. **93**: p. 201-204.
42. Li, R., et al., *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2).* Science, 2020: p. eabb3221.
43. Gourieroux, C., A. Monfort, and E. Renault, *Indirect inference.* Journal of applied econometrics, 1993. **8**(S1): p. S85-S118.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Online Supplementary Document

Weather Conditions and COVID-19 Transmission: Estimates and Projections

Contents

1. Data and replication instructions 14

2. Estimating the detection delay distribution 15

3. Algorithmic estimation of true infection rate 16

4. Statistical sensitivity analyses and robustness checks 18

4.1. Sensitivity to duration of disease (15, 20, 25) 19

4.2. Sensitivity to excluding outliers (extreme large R0 and last couple of days) and including fixed effects 19

4.3. Sensitivity to shifting the data start date after first exposure 20

4.4. Placebo tests (random shifts of weather data) 22

4.5. Different knot for linear spline effect of mean temperature 23

4.6 Inclusion of moon illumination 24

5. Specification and validation using synthetic data 26

5.1. Summary of our approach and findings 26

5.2. Statistical specification using synthetic data 26

5.2.1. Simulation model 27

5.2.2. Synthetic experiments 28

5.3. Blinded study specification and results 37

6. Global projections over time 39

References: 41

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

1. Data and replication instructions

All code and data for this research is available at https://github.com/marichig/weather-conditions-COVID19.

Case and coordinate data were first taken from JHU's published case reports, available at https://github.com/CSSEGISandData/COVID-19, which covered all locations in the United States and 258 of the 590 locations from outside the U.S., including breakdowns for Canada into 10 states/territories, and Australia into 8 states. The remaining locations were made up of 301 Chinese cities and 31 Iranian states, and for these case and coordinate data were taken from the Chinese Center for Disease Control and Prevention, Provincial Health Commissions in China, and Iran's state level reports.

Some locations included in the U.S. case reporting data were dropped from the main analysis. Namely:

- Cases from the cruise ships Diamond Princess, Grand Princess, and MS Zaandam were discarded.
- Cases labelled as "Out of [State]" or "Unassigned, [State]" were discarded.
- Cases from Michigan Department of Corrections and Federal Correctional Institute, Michigan were dropped since they reflect unique spread dynamics and carried no coordinate data.
- Cases attributed to the Utahn Local Health Departments (Bear River, Central Utah, Southeast Utah, Southwest Utah, TriCounty, and Weber-Morgan) were discarded; as of 4/22/2020, only 291 cases were reported from these sources compared to 3154 from all Utah counties. These health departments span several counties and reporting from them only began on 4/19/2020.

Errors in the reported coordinate data were also identified and resolved manually. (For instance, Congo-Brazzaville was reported to have the same coordinates as Congo-Kinshasa.) With this coordinate data, weather data is collected primarily through World Weather Online (WWO), which provides an API for data collection – the Python "wwo-hist" package <https://pypi.org/project/wwo-hist/> was used to access this API. Historical weather data were collected for each day between 1/23/2019 thru 4/22/2020, with data from 2019 being used for future projection.

The following variables were collected: maximum daily temperature, Celsius; minimum daily temperature, Celsius; average daily temperature, Celsius; precipitation, millimeters; humidity, percentage; pressure (atmospheric), millibars; windspeed, kilometers per hour; sun hours (i.e., hours of sunshine received); total snowfall, centimeters; cloud cover, percentage; ultraviolet (UV) index; moon illumination, percent (i.e., percentage of moon face lit by the sun); local sunrise and sunset time; local moonrise and moonset time; dew point, Celsius; "Feels Like", Celsius; wind chill, Celsius; wind gust (i.e., peak instantaneous speed), kilometers per hour; visibility, kilometers; and wind direction degree, clockwise degrees from due north. A description of each variable is available at https://www.worldweatheronline.com/developer/api/docs/historical-weather-api.aspx. The ultraviolet (UV) index data was not consistently reported from WWO, and was instead gathered using OpenWeatherMap <https://openweathermap.org/> and the Python "pyowm" package <https://pypi.org/project/pyowm/>.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

We interpolated over any missing entries in the temperature and UV data provided. The reported temperature data were missing for most (but not all) locations 9/15-17/2019, 10/22/2019, 11/27/2019, and 12/15/2019, which were then interpolated using five-day moving averages. UV data were missing for less than 0.1% of date-location pairs, with the main gaps occurring on 6/2/2019, 8/13/2019, 12/2/2019, 2/18/2020, and 2/21/2020, which were interpolated using three-day moving averages. This averaging should not impact the analysis given that most of the above dates fall outside the pandemic's date-range.

Population density data was sourced from Demographia (Cox., W, *Demographia World Urban Areas, 15th Edition*, The Public Purpose), which provided data for urban areas with population greater than 500,000; the United States Census (U.S. Census Bureau, data.census.gov/cedsci); the Iran Statistical Centre; the United Nation's Projections; *City Population* (citypopulation.de); and official published estimates for countries not covered by these sources. For data sourced from Demographia, the population densities reported are urban densities, whereas other sources primarily reported overall density (spanning urban and non-urban areas). The urban and overall densities are largely on different orders, which weakens the inclusion of population density as an independent variable.

2. Estimating the detection delay distribution

Reported data on daily detected infections for COVID-19 do not reflect the true infection rate on a given day, rather, it lags the true infections due to both the incubation period (during which patients are asymptomatic and less likely to be tested), and the delay between onset of symptoms, testing, and incorporation of test results into official data. We need estimates for the true infection rates for each day to calculate the daily reproduction number (i.e., $\hat{R}(t)$), therefore identifying the lag structure between measured infection (I_M) and true infection (I_N), which we call "Detection Delay" is key to back tracking from measured infection to estimates of true infection rate.

Prior research has provided several estimates for subsets of overall Detection Delay. Incubation period, the time between infection to onset of symptoms, has been estimated by several teams. Li and colleagues [1], using data from 10 early patients in China, find the mean incubation period to be 5.2 days, and the delay from onset to first medical visit to be 5.8 days for those infected before January 1st and 4.6 days for the later cases. Lauer and colleagues [2] use data from 181 cases to estimate incubation period with mean of 5.5 and median of 5.1, and offer fitted distributions using Lognormal, Gamma, Weibull, and Erlang specifications. In a supplementary graph, they also provide a figure that includes the lags from the onset of symptom to official case detection. Guan et al. [3] use data from 291 patients and estimate median incubation period of four days with interquartile range of 2 to 7 days. Linton and colleagues [4] use data from 158 cases to estimate the incubation period with a mean (standard deviation) of 5.6 (2.8) days. This delay goes down to 5 (3) when excluding Wuhan patients. They also report onset to hospital admission delay of 3.9 (3) days for living patients (155 cases). They provide their full data in an online appendix, where we calculated the onset to case report lag with mean of 5.6 days, median of 5, and standard deviation of 3.8 days. A *New York Times* article [5] reports that the Center for Disease Control estimates the lag between onset of symptoms to case detection to be four days. Finally, a Bayesian estimation of the detection delay using abrupt changes in national and state policies by Wibbens and colleagues find the mode of the delay to be 11 days and ranging between 5 and 20 days [6].

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Overall, these findings are consistent and point to an incubation period of about 5 days and an onset to detection lag of about the same length. We use Lauer et al. estimates for a Lognormal incubation period with parameters 1.62 and 0.418 (leading to mean (standard deviation) of 5.51 (2.4) days), and another Lognormal distribution with parameters 1.47 and 0.52 (resulting in 5 (2.8) days) for onset to detection delay. Combining these two distributions using 10 million Monte-Carlo simulations, we generate the following Detection Delay lag structure that is used in the analysis. The code calculating this distribution is found at <https://github.com/marichig/weather-conditions-COVID19/>.

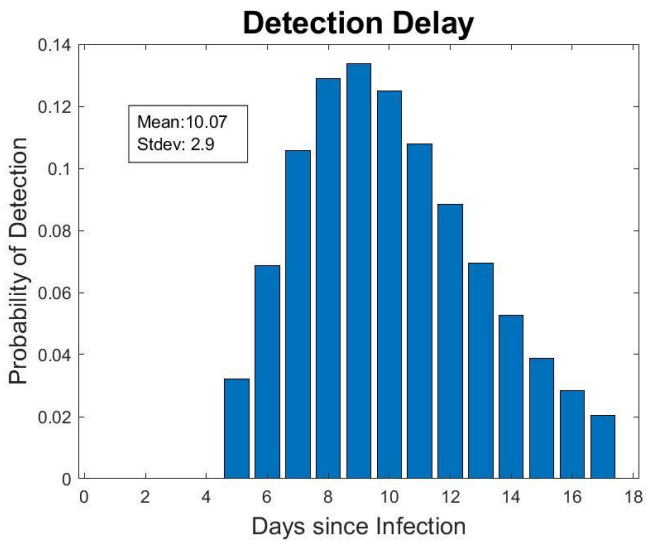


Figure S4: Distribution of Detection Delay

3. Algorithmic estimation of true infection rate

Here we develop an algorithm that provides a more accurate estimation of true exposure than a fixed shift in reported data or averaging data over a time period. We later compare our algorithm's performance with the simpler, more common, methods. We find that accurate estimation of effects of weather variables hinges directly on accurately estimating true infections, making the algorithm in this section key to overall estimation.

Using the delay structure specified in the previous section, one can estimate true infection rates using various methods. The most common solution is to just shift the official infections based on the average, median, or mode of the Detection Delay (9 to 11 days). This approximation may be fine in steady state, but becomes more inaccurate when estimating time series with exponential growth: the detected infections today are more likely to be from (the many more) recent infections than (the fewer) 10 days ago.

The main objective of our algorithm is to find better estimates for the true infection. We first calculate the expected number of daily detected cases, given a series of actual infections unknown in the real world. Call the actual infection on day t, $X(t)$, and the detected infections on day t, $I(t)$. The following equation would relate the two constructs:

$$E(I(t)) = \sum_{d=1}^{L} X(t-d)p(d)$$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Where $E(.)$ takes the expectation on detected infections, $p(.)$ is the probability distribution for the detection delay estimated in section 2 of appendix, and the index d ranges between 1 to $L = 17$ days to account for different delay lengths. This equation does not account for test coverage, but as discussed below test coverage cancels out of the final reproduction number calculations, and as such only impacts variability of outcomes and otherwise has limited impact on results. Note that this equation is under-specified: for one value of the known measure I, one has to find up to L values of the unknown X (in our case, no detection is expected in the first 4 days, so L-4=13 values of X contribute to a value of I (Figure S4)). However, given the overlap on X's determining subsequent I values, the system of equations connecting I and X values for I time series extending over T days would include T known values (for I) and $T + L$ unknown X values. Different approaches could then be pursued to find approximate solutions for this system of equations.

Using exact Maximum Likelihood suffers from intractability of specifying the Likelihood for highly correlated Poisson distributions (Poisson is a natural alternative in this case). We compared two alternatives, one using Normally distributed approximations for I as a function of X, and another using a direct minimization of the gap between I values and their expectation. The latter proves both simpler conceptually and more accurate in synthetic data, so we picked that for the main analysis:

$$X = Argmin\left(\sum_{t=1}^{T}\left(I(t) - \sum_{d=1}^{L} X(t-d)p(d)\right)^2\right)$$

Given the under-specification of original system of equations, this optimization will include many solutions. To identify a more realistic solution from that set, we add a regularization term that penalizes the gap between subsequent values for X, specifically, we use the following optimization:

$$X = Argmin\left(\sum_{t=1}^{T}\left(I(t) - \sum_{d=1}^{L} X(t-d)p(d)\right)^2 + \lambda \sum_{t=-L+1}^{T}\left(X(t-1) - X(t)\right)^2\right)$$
$$s.t. X(t) \geq 0$$

The solution to this optimization can be found using standard quadratic programming methods, allowing for fast and scalable solutions. We conducted sensitivity analysis to find the regularization parameter, λ, offering the best overall ability of the algorithm to find true infections in synthetic data. The algorithm that works well is with λ values in the 0.1 to 0.5 range and not very sensitive to exact value; we used a value of 0.2 in our analysis. An implementation of this code in Matlab is available from <https://github.com/marichig/weather-conditions-COVID19/>.

For each location in our dataset, we used this algorithm to estimate the true infections ($\hat{I}_N(t) = X(t)$), on a daily basis, starting from 17 days before the first detected infection, and stopping 5 days before the last day with data (because only infections from 5 days or further back could be found in current measures of infection; see the Detection Delay distribution (Figure S4)). These values were then used to create the dependent variable, $\hat{R}(t)$, as discussed in the body of the article:

$$\hat{R}(t) = \frac{\hat{I}_N(t)\tau}{\sum_{s=t-\tau-1}^{s=t-1} \hat{I}_N(s)}$$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

We recognize that not all infections are reported, and a large fraction may remain unknown. Assuming that only a fraction *f* (0≤*f*≤1) of actual infections are reported, I_M would be *f* of total infections that could have been detected on a given day, and estimated $\hat{I}_N(t)$ will be the fraction *f* of true infections as a result. While these under-estimations are likely very significant if we cared about absolute values of $\hat{I}_N(t)$, note that $\hat{I}_N(t)$ values show up both in the numerator and denominator of $\hat{R}(t)$ equation. Therefore, multiplying both by a fixed constant makes no difference in the estimated $\hat{R}(t)$.

We also recognize that, early on during the infection, *f* may increase with expanding test capacity until reaching a steady state value. Therefore, as later discussed, we drop the first few data points for each region and check the sensitivity of the result to dropping fewer or more days. Finally, our synthetic analysis (section 5.2.2.3, Experiment 10) shows results are robust to various trajectories for *f* over the course of epidemic.

4. Statistical sensitivity analyses and robustness checks

We conducted six different set of sensitivity tests to assess the robustness of our findings to various assumptions and boundary conditions[1]. Here is a summary of the results, before we go into the details: (1) In our main specification we used a delay of τ=20 days from exposure to resolution to calculate reproduction number R_0. Here we tested the robustness of our results over a spectrum of reasonable durations of delay from 15 to 25 days, finding no major impact on the results. (2) We tested outcomes under three additional exclusion criteria and specifications: exclusion of the last few days of data (for which true infection estimates may be less reliable), exclusions of top 1% R_0 of our sample (which may be generated due to reporting issues), and inclusion of location fixed effects. Results are qualitatively robust under all these specifications. (3) In our main specification we dropped the first 20 days since new infection exceeds 1 for each location to account for early-on changes in test coverage. Here we tested the robustness of our results to other exclusion periods, ranging from first 10 days to 30 days, finding no major impact on the outcomes (4) To exclude the possibility that our results are driven by mechanical features of our variable construction and model specification, we used a set of placebo weather variables, which are randomly shifted across locations and over a specific number of days, and re-estimated our main models using these placebo weather variables. We found few significant effects under these placebo tests. (5) In our main specification we chose a linear spline effect of mean temperature with a knot at 25 degrees. Here we tested how our results are sensitive to different choices of knots, finding the 25 degree provides the best balance. (6) Finally, we reported analysis that includes moon illumination as an additional independent variable to our main specification. We found consistent and significant negative effect of moon illumination on reproduction number. However, lacking any theoretical justification for this effect, we did not include this factor in our preferred model specification.

[1] Same as in our main specification all models tested here used $\log(R_0)$ as the outcome and included location specific linear-trends. All models excluded days with new infections <1 and first 20 days since new infection exceeds 1 for the first time in each location unless specified otherwise.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

4.1. Sensitivity to duration of disease (15, 20, 25)

Table S2 presents the results using our main specification with R_0 calculated from 15, 20 and 25 days of delay respectively. The coefficients and significance level for each weather variable are largely unchanged and consistent across different durations, especially for air pressure, mean temperature, and interaction effect between humidity and mean temperature.

Table S2. Regression results with various duration of delay to calculate R_0

	15 Days Main Effects	15 Days Interaction Effects	20 Days Main Effects	20 Days Interaction Effects	25 Days Main Effects	25 Days Interaction Effects
Population density	0.188***	0.191***	0.176***	0.179***	0.149***	0.152***
	(0.0104)	(0.0105)	(0.0104)	(0.0104)	(0.0103)	(0.0103)
Wind speed	0.0889***	0.0877***	0.0790***	0.0782***	0.0698***	0.0696***
	(0.0135)	(0.0136)	(0.0134)	(0.0135)	(0.0133)	(0.0134)
Snowfall	0.0257	0.0385*	0.0273	0.0395**	0.0264	0.0369*
	(0.0148)	(0.0154)	(0.0147)	(0.0153)	(0.0146)	(0.0152)
Precipitation	0.0297***	0.0292***	0.0278***	0.0273***	0.0254***	0.0249***
	(0.00633)	(0.00633)	(0.00629)	(0.00629)	(0.00625)	(0.00624)
Air pressure	-0.00231	-0.00225	-0.00263	-0.00255	-0.00415	-0.00404
	(0.00280)	(0.00280)	(0.00279)	(0.00278)	(0.00276)	(0.00276)
Air pressure^2	0.000288**	0.000275**	0.000344***	0.000331***	0.000375***	0.000363***
	(8.80e-05)	(8.80e-05)	(8.74e-05)	(8.74e-05)	(8.67e-05)	(8.68e-05)
Humidity	-0.000485	-0.00204**	2.47e-05	-0.00143	0.000411	-0.000795
	(0.000528)	(0.000790)	(0.000525)	(0.000785)	(0.000521)	(0.000779)
Temperature before 25	-0.00460***	-0.0154***	-0.00533***	-0.0156***	-0.00573***	-0.0145***
	(0.00126)	(0.00390)	(0.00125)	(0.00388)	(0.00124)	(0.00385)
Humidity*Temperature before 25		0.000152**		0.000145**		0.000126*
		(5.11e-05)		(5.08e-05)		(5.04e-05)
Temperature after 25	-0.0323***	0.0199	-0.0318***	0.0247	-0.0280**	0.0331*
	(0.00869)	(0.0164)	(0.00864)	(0.0163)	(0.00857)	(0.0162)
Humidity*Temperature after 25		-0.00107***		-0.00116***		-0.00127***
		(0.000296)		(0.000294)		(0.000291)
Sun hour	-0.000821	-0.000260	0.00160	0.00224	0.00473	0.00546
	(0.00306)	(0.00307)	(0.00304)	(0.00305)	(0.00302)	(0.00303)
Diurnal temperature	0.00638***	0.00682***	0.00495**	0.00537**	0.00385*	0.00421*
	(0.00186)	(0.00187)	(0.00185)	(0.00185)	(0.00183)	(0.00184)
Observations	19,141	19,141	19,140	19,140	19,135	19,135
R-squared	0.464	0.465	0.533	0.534	0.582	0.582

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4.2. Sensitivity to excluding outliers (extreme large R0 and last couple of days) and including fixed effects

Table S3 presents the results when we exclude top 1% R_0, last 4 days of our data (19% of the total sample in our main specification), or when we include location fixed effects. The effect of wind speed, precipitation, air pressure, mean temperature and interaction between humidity and mean temperature are all robust to both exclusion criteria, while the positive effects of snow fall and diurnal temperature are no longer significant when we exclude the last 4 days of data. With the inclusion of location fixed effects, all of the weather effects become weaker, confirming our

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

intuition (from the synthetic experiments) that weather effects might be underestimated with the inclusion of location fixed effects. The main reason is that within region variation of temperature is less than cross-regional variation, and with fixed effect regressions, one loses the opportunity to fully utilize temperature variation in the data. However, even with fixed effect, the effect of wind speed, precipitation, mean temperature and interaction between humidity and temperature are still consistent and significant.

Table S3. Regression results with various exclusion criteria and specifications

	Exclude extreme R_0 Main effects	Exclude extreme R_0 Interaction effects	Exclude last 4 days of data Main effects	Exclude last 4 days of data Interaction effects	Include location fixed effects Main effects	Include location fixed effects Interaction effects
Population density	0.156***	0.158***	0.118***	0.120***		
	(0.0102)	(0.0102)	(0.0119)	(0.0119)		
Wind speed	0.0805***	0.0793***	0.0861***	0.0830***	0.0364**	0.0323**
	(0.0132)	(0.0132)	(0.0141)	(0.0142)	(0.0114)	(0.0115)
Snowfall	0.0240	0.0337*	0.0172	0.0340	-0.00242	0.0139
	(0.0144)	(0.0150)	(0.0183)	(0.0189)	(0.0126)	(0.0132)
Precipitation	0.0265***	0.0261***	0.0384***	0.0385***	0.0269***	0.0267***
	(0.00616)	(0.00616)	(0.00665)	(0.00665)	(0.00521)	(0.00522)
Air pressure	-0.000422	-0.000400	-0.00774**	-0.00746**	0.00351	0.00356
	(0.00273)	(0.00273)	(0.00288)	(0.00288)	(0.00233)	(0.00233)
Air pressure^2	0.000271**	0.000261**	0.000365***	0.000340***	-4.55e-05	-6.30e-05
	(8.60e-05)	(8.60e-05)	(9.06e-05)	(9.07e-05)	(7.29e-05)	(7.31e-05)
Humidity	9.00e-05	-0.00110	-0.000101	-0.00236**	-0.000633	-0.00284***
	(0.000513)	(0.000767)	(0.000583)	(0.000903)	(0.000471)	(0.000720)
Temperature before 25	-0.00502***	-0.0131***	-0.00481***	-0.0190***	-0.00418***	-0.0181***
	(0.00122)	(0.00379)	(0.00134)	(0.00427)	(0.00121)	(0.00345)
Humidity*Temperature before 25		0.000114*		0.000200***		0.000198***
		(4.97e-05)		(5.62e-05)		(4.56e-05)
Temperature after 25	-0.0340***	0.000807	-0.0289**	0.0236	-0.0169*	0.0280
	(0.00849)	(0.0163)	(0.00916)	(0.0175)	(0.00770)	(0.0150)
Humidity*Temperature after 25		-0.000701*		-0.00106***		-0.000843**
		(0.000293)		(0.000316)		(0.000263)
Sun hour	0.00253	0.00288	0.0104**	0.0108***	0.00209	0.00236
	(0.00298)	(0.00299)	(0.00327)	(0.00327)	(0.00253)	(0.00253)
Diurnal temperature	0.00529**	0.00560**	0.00185	0.00223	0.00281	0.00313*
	(0.00181)	(0.00181)	(0.00200)	(0.00200)	(0.00154)	(0.00154)
Observations	18,949	18,949	15,519	15,519	19,160	19,140
R-squared	0.528	0.528	0.577	0.578	0.739	0.739

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4.3. Sensitivity to shifting the data start date after first exposure

Table S4 and Table S5 present the results when we exclude first 10, 15, 20, 25, and 30 days since new infection exceeds 1 for the first time in each location. Overall our main results are consistent and insensitive to different exclusion criteria, except when we exclude first 30 days of data for each location, where we would lose more than half of our sample compared with the main specification (i.e., exclude first 20 days), the effect of wind speed, precipitation, temperature become weaker and are no longer significant. It is possible that by constraining our

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

estimation on later periods, we are focusing on periods when lockdown and social distancing are fully in effect and thus there are few variations left in R_0 that can be explained by weather. However, the coefficients for these aforementioned effects are still consistent and have the same sign. For example, from column 5 we estimated that with one degree increase in mean temperature after 25 degrees, the estimated R_0 will still decrease by ~2%.

Table S4. Regression results with various exclusion criteria for initial periods – Main effects

Main Effects	Exclude first 10 days	Exclude first 15 days	Exclude first 20 days	Exclude first 25 days	Exclude first 30 days
Population density	0.0871***	0.100***	0.176***	0.223***	0.254***
	(0.00528)	(0.00718)	(0.0104)	(0.0154)	(0.0235)
Wind speed	0.110***	0.109***	0.0790***	0.0517**	0.0165
	(0.0100)	(0.0115)	(0.0134)	(0.0159)	(0.0192)
Snowfall	0.0369**	0.0290*	0.0273	0.00834	0.00891
	(0.0117)	(0.0130)	(0.0147)	(0.0167)	(0.0197)
Precipitation	0.0253***	0.0308***	0.0278***	0.0235**	0.0129
	(0.00468)	(0.00538)	(0.00629)	(0.00747)	(0.00917)
Air pressure	-0.00337	-0.00170	-0.00263	-0.00128	-0.00252
	(0.00215)	(0.00244)	(0.00279)	(0.00318)	(0.00381)
Air pressure^2	0.000547***	0.000472***	0.000344***	0.000267**	0.000307**
	(6.37e-05)	(7.48e-05)	(8.74e-05)	(0.000100)	(0.000118)
Humidity	0.00125**	0.000963*	2.47e-05	0.000505	0.000784
	(0.000393)	(0.000447)	(0.000525)	(0.000626)	(0.000765)
Temperature before 25	-0.000903	-0.00127	-0.00533***	-0.00506***	-0.00393*
	(0.000842)	(0.00102)	(0.00125)	(0.00152)	(0.00189)
Temperature after 25	-0.0215**	-0.0236**	-0.0318***	-0.0290**	-0.0200
	(0.00676)	(0.00774)	(0.00864)	(0.00939)	(0.0104)
Sun hour	0.00150	0.00349	0.00160	-0.00333	-0.00503
	(0.00231)	(0.00262)	(0.00304)	(0.00359)	(0.00442)
Diurnal temperature	0.00584***	0.00467**	0.00495**	0.00459*	-7.63e-05
	(0.00139)	(0.00159)	(0.00185)	(0.00216)	(0.00264)
Observations	32,041	25,787	19,140	13,614	9,041
R-squared	0.546	0.532	0.533	0.569	0.622

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S5. Regression results with various exclusion criteria for initial periods – Interaction effects

Interaction Effects	Exclude first 10 days	Exclude first 15 days	Exclude first 20 days	Exclude first 25 days	Exclude first 30 days
Population density	0.0907***	0.103***	0.179***	0.227***	0.258***
	(0.00529)	(0.00719)	(0.0104)	(0.0154)	(0.0236)
Wind speed	0.104***	0.105***	0.0782***	0.0517**	0.0182
	(0.0100)	(0.0115)	(0.0135)	(0.0159)	(0.0193)
Snowfall	0.0572***	0.0476***	0.0395**	0.0218	0.0139
	(0.0121)	(0.0134)	(0.0153)	(0.0176)	(0.0210)
Precipitation	0.0244***	0.0300***	0.0273***	0.0228**	0.0124
	(0.00468)	(0.00537)	(0.00629)	(0.00747)	(0.00917)
Air pressure	-0.00324	-0.00159	-0.00255	-0.00125	-0.00239
	(0.00215)	(0.00244)	(0.00278)	(0.00318)	(0.00381)
Air pressure^2	0.000513***	0.000445***	0.000331***	0.000258*	0.000303*
	(6.38e-05)	(7.49e-05)	(8.74e-05)	(0.000100)	(0.000118)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Interaction Effects	Exclude first 10 days	Exclude first 15 days	Exclude first 20 days	Exclude first 25 days	Exclude first 30 days
Humidity	-0.00158**	-0.00155*	-0.00143	-0.000948	0.000413
	(0.000596)	(0.000675)	(0.000785)	(0.000932)	(0.00113)
Temperature before 25	-0.0207***	-0.0186***	-0.0156***	-0.0151***	-0.00710
	(0.00300)	(0.00338)	(0.00388)	(0.00450)	(0.00544)
Humidity*Temperature before 25	0.000266***	0.000238***	0.000145**	0.000147*	4.97e-05
	(3.85e-05)	(4.37e-05)	(5.08e-05)	(5.99e-05)	(7.37e-05)
Temperature after 25	0.0624***	0.0510***	0.0247	0.0230	0.0156
	(0.0131)	(0.0147)	(0.0163)	(0.0174)	(0.0192)
Humidity*Temperature after 25	-0.00170***	-0.00152***	-0.00116***	-0.00109***	-0.000787*
	(0.000243)	(0.000269)	(0.000294)	(0.000316)	(0.000359)
Sun hour	0.00201	0.00397	0.00224	-0.00259	-0.00445
	(0.00231)	(0.00262)	(0.00305)	(0.00359)	(0.00443)
Diurnal temperature	0.00642***	0.00527***	0.00537**	0.00508*	0.000157
	(0.00140)	(0.00159)	(0.00185)	(0.00217)	(0.00266)
Observations	32,041	25,787	19,140	13,614	9,041
R-squared	0.548	0.533	0.534	0.569	0.622

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4.4. Placebo tests (random shifts of weather data)

We first randomly permuted weather variables across locations in our data, and then shifted all weather variables in each location to earlier periods by a specific number of days, where the number is randomly drawn from a uniform distribution U(0,300). We then performed the statistical analysis using these "placebo" weather variables. As shown in Table S6, most of the weather effects are completely gone, especially in our main specification where first 20 days are excluded. The only exception is when we observe a positive and significant effect of humidity and temperature after 25 degrees when first 30 days are excluded, which have opposite signs from the results in our main conclusion. The results are likely to be purely driven by chance, but nevertheless we conclude that caution should be exercised when interpreting results from specifications dropping a large number of initial periods.

Table S6. Regression results using placebo weather with various exclusion criteria for initial periods

	(1) Exclude first 10 days	(2) Exclude first 15 days	(3) Exclude first 20 days	(4) Exclude first 25 days	(5) Exclude first 30 days
Population density	0.0922***	0.112***	0.191***	0.237***	0.259***
	(0.00523)	(0.00713)	(0.0103)	(0.0152)	(0.0233)
Wind speed	0.00164	-0.00188	-0.0102	-0.0161	-0.0237
	(0.00980)	(0.0112)	(0.0131)	(0.0152)	(0.0186)
Snowfall	-0.0275*	-0.0195	-0.0167	-0.0191	-0.0181
	(0.0118)	(0.0132)	(0.0153)	(0.0176)	(0.0208)
Precipitation	-0.00585	-0.00484	-0.00524	-0.0113	-0.00510
	(0.00448)	(0.00507)	(0.00588)	(0.00684)	(0.00822)
Air pressure	-0.000632	-0.00113	0.000193	-0.00115	-0.00157
	(0.00156)	(0.00186)	(0.00228)	(0.00269)	(0.00334)
Air pressure^2	0.000103*	0.000116*	7.55e-05	0.000102	0.000115
	(4.16e-05)	(4.95e-05)	(6.05e-05)	(7.17e-05)	(8.85e-05)
Humidity	7.06e-05	0.000163	0.000215	0.00128*	0.00227**

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

	(1) Exclude first 10 days	(2) Exclude first 15 days	(3) Exclude first 20 days	(4) Exclude first 25 days	(5) Exclude first 30 days
	(0.000401)	(0.000463)	(0.000544)	(0.000646)	(0.000794)
Temperature before 25	-0.000969	-0.000413	0.00148	0.000845	-0.000121
	(0.000844)	(0.00102)	(0.00125)	(0.00153)	(0.00190)
Temperature after 25	-0.00316	0.000447	0.00705	0.0161*	0.0308***
	(0.00424)	(0.00495)	(0.00597)	(0.00703)	(0.00876)
Sun hour	-0.00531*	-0.00439	-0.00245	-0.00144	0.00369
	(0.00214)	(0.00245)	(0.00286)	(0.00335)	(0.00404)
Diurnal temperature	0.000694	0.000534	0.000855	0.000109	-0.00221
	(0.00139)	(0.00157)	(0.00182)	(0.00213)	(0.00259)
Observations	32,041	25,787	19,140	13,614	9,041
R-squared	0.537	0.524	0.528	0.565	0.619

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4.5. Different knot for linear spline effect of mean temperature

In the main specification we used a linear spline effect of mean temperature with a knot at 25 degrees as it provides better fit than linear or quadratic effect of temperature. Here we test the sensitivity of our results to the choice of knots over a wide range of mean temperature from -15 degrees to 30 degrees. As shown in Table S7, the temperature effect after the knot is statistically significant and much larger at 25 degrees (-.0318, p<.001) than knots at other degrees. Hence, we chose the knot at 25 degrees as our main specification.

Table S7. Regression results with different knots for linear spline effect of mean temperature

VARIABLES	(1) Knot at -15	(2) Knot at -10	(3) Knot at -5	(4) Knot at 0	(5) Knot at 5
Population density	0.175***	0.175***	0.175***	0.175***	0.175***
	(0.0104)	(0.0104)	(0.0104)	(0.0104)	(0.0104)
Wind speed	0.0785***	0.0785***	0.0785***	0.0784***	0.0787***
	(0.0134)	(0.0134)	(0.0134)	(0.0134)	(0.0134)
Snowfall	0.0238	0.0237	0.0239	0.0256	0.0265
	(0.0146)	(0.0146)	(0.0147)	(0.0148)	(0.0155)
Precipitation	0.0282***	0.0282***	0.0282***	0.0282***	0.0282***
	(0.00629)	(0.00629)	(0.00629)	(0.00629)	(0.00629)
Air pressure	-0.00231	-0.00236	-0.00234	-0.00246	-0.00240
	(0.00278)	(0.00279)	(0.00279)	(0.00279)	(0.00279)
Air pressure^2	0.000334***	0.000336***	0.000335***	0.000339***	0.000338***
	(8.74e-05)	(8.74e-05)	(8.75e-05)	(8.76e-05)	(8.77e-05)
humidity	9.80e-05	0.000102	0.000101	0.000123	0.000109
	(0.000525)	(0.000525)	(0.000525)	(0.000526)	(0.000525)
Temperature before 25	0.0116	0.00485	-0.00354	-0.00220	-0.00440
	(0.0755)	(0.0200)	(0.00938)	(0.00536)	(0.00336)
Temperature after 25	-0.00608***	-0.00616***	-0.00613***	-0.00633***	-0.00637***
	(0.00123)	(0.00124)	(0.00126)	(0.00128)	(0.00136)
Sun hour	0.00174	0.00175	0.00175	0.00177	0.00171
	(0.00305)	(0.00305)	(0.00305)	(0.00305)	(0.00305)
Diurnal temperature	0.00482**	0.00485**	0.00484**	0.00492**	0.00485**
	(0.00185)	(0.00185)	(0.00185)	(0.00186)	(0.00185)
Observations	19,140	19,140	19,140	19,140	19,140

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

VARIABLES	(1) Knot at -15	(2) Knot at -10	(3) Knot at -5	(4) Knot at 0	(5) Knot at 5
R-squared	0.533	0.533	0.533	0.533	0.533

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

Table S6 (Cont'd). Regression results with different knots for linear spline effect of mean temperature

VARIABLES	(1) Knot at 10	(2) Knot at 15	(3) Knot at 20	(4) Knot at 25	(5) Knot at 30
Population density	0.175***	0.175***	0.175***	0.176***	0.175***
	(0.0104)	(0.0104)	(0.0104)	(0.0104)	(0.0104)
Wind speed	0.0804***	0.0781***	0.0774***	0.0790***	0.0782***
	(0.0135)	(0.0135)	(0.0135)	(0.0134)	(0.0134)
Snowfall	0.0303*	0.0230	0.0220	0.0273	0.0243
	(0.0153)	(0.0150)	(0.0148)	(0.0147)	(0.0147)
Precipitation	0.0280***	0.0283***	0.0285***	0.0278***	0.0281***
	(0.00629)	(0.00630)	(0.00630)	(0.00629)	(0.00629)
Air pressure	-0.00252	-0.00227	-0.00222	-0.00263	-0.00241
	(0.00279)	(0.00279)	(0.00279)	(0.00279)	(0.00279)
Air pressure^2	0.000345***	0.000332***	0.000330***	0.000344***	0.000337***
	(8.77e-05)	(8.76e-05)	(8.75e-05)	(8.74e-05)	(8.74e-05)
Humidity	0.000117	9.40e-05	9.59e-05	2.47e-05	6.80e-05
	(0.000525)	(0.000525)	(0.000525)	(0.000525)	(0.000525)
Temperature before 25	-0.00353	-0.00633***	-0.00657***	-0.00533***	-0.00599***
	(0.00217)	(0.00164)	(0.00138)	(0.00125)	(0.00123)
Temperature after 25	-0.00755***	-0.00558*	-0.00312	-0.0318***	-0.0323
	(0.00162)	(0.00224)	(0.00382)	(0.00864)	(0.0225)
Sun hour	0.00161	0.00175	0.00176	0.00160	0.00160
	(0.00305)	(0.00305)	(0.00305)	(0.00304)	(0.00305)
Diurnal temperature	0.00476**	0.00483**	0.00483**	0.00495**	0.00481**
	(0.00185)	(0.00185)	(0.00185)	(0.00185)	(0.00185)
Observations	19,140	19,140	19,140	19,140	19,140
R-squared	0.533	0.533	0.533	0.533	0.533

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

4.6 Inclusion of moon illumination

Finally, as an exploratory analysis we included moon illumination (percentage; range: 0-100) as an additional independent variable in our model. While all of our previous conclusions regarding other weather variables remain unchanged, we observed a significant and unneglectable negative effect of moon illumination: 1 percent increase in moon illumination is associated with 0.25% decrease in estimated reproduction number. This effect is robust to looking at different subsets of locations with different start dates of epidemic, so it cannot be explained based on mechanical artifacts of the timing of majority of locations. Overall, we lack a causal theory for why such an effect may exist. Moreover, as moon illumination is constant across locations in each day and exhibit cyclic behavior over each lunar month, it is possible that the moon illumination is confounded with other cyclic behavioral patterns occurring globally in our data, and thus we decided not to include it in the main analysis. Nevertheless, we reported it here to show other weather effects are robust to this additional inclusion and point to possible future avenues of investigation.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table S7. Regression results including moon illumination effect

VARIABLES	(1) Exclude first 10 days	(2) Exclude first 15 days	(3) Exclude first 20 days	(4) Exclude first 25 days	(5) Exclude first 30 days
Population density	0.0895***	0.0936***	0.153***	0.173***	0.187***
	(0.00526)	(0.00715)	(0.0104)	(0.0158)	(0.0243)
Wind speed	0.0869***	0.0755***	0.0512***	0.0261	-0.00224
	(0.0101)	(0.0116)	(0.0135)	(0.0159)	(0.0192)
Snowfall	0.0308**	0.0194	0.0190	0.000102	0.00479
	(0.0117)	(0.0129)	(0.0146)	(0.0166)	(0.0196)
Precipitation	0.0216***	0.0247***	0.0214***	0.0170*	0.00710
	(0.00468)	(0.00536)	(0.00627)	(0.00744)	(0.00913)
Air pressure	-0.00652**	-0.00615*	-0.00650*	-0.00491	-0.00599
	(0.00215)	(0.00244)	(0.00278)	(0.00318)	(0.00380)
Air pressure^2	0.000570***	0.000501***	0.000374***	0.000299**	0.000350**
	(6.35e-05)	(7.44e-05)	(8.69e-05)	(9.96e-05)	(0.000117)
Humidity	0.000817*	0.000534	-0.000163	0.000463	0.000736
	(0.000393)	(0.000445)	(0.000522)	(0.000622)	(0.000760)
Temperature before 25	-9.14e-05	0.000422	-0.00254*	-0.00109	-0.000135
	(0.000841)	(0.00102)	(0.00126)	(0.00154)	(0.00192)
Temperature after 25	-0.0227***	-0.0250**	-0.0320***	-0.0252**	-0.0121
	(0.00673)	(0.00770)	(0.00859)	(0.00933)	(0.0104)
Sun hour	0.000597	0.00208	0.000867	-0.00306	-0.00518
	(0.00230)	(0.00260)	(0.00303)	(0.00356)	(0.00439)
Diurnal temperature	0.00523***	0.00372*	0.00379*	0.00305	-0.00107
	(0.00139)	(0.00158)	(0.00184)	(0.00215)	(0.00262)
Moon illumination	-0.00179***	-0.00251***	-0.00246***	-0.00280***	-0.00273***
	(0.000127)	(0.000145)	(0.000173)	(0.000214)	(0.000265)
Observations	32,041	25,787	19,140	13,614	9,041
R-squared	0.549	0.537	0.538	0.575	0.626

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

5. Specification and validation using synthetic data

We conducted extensive synthetic analyses to inform the selection of a reliable statistical method and to build confidence in our final estimation method. These analyses could be divided into those focused on specifying the estimation method (5.2), and those designed to validate our method in a study where analyst was blinded to the true specification (5.3). After providing an overview in 5.1 we explain these two sets of analyses. Detailed codes are available at <https://github.com/marichig/weather-conditions-COVID19/>.

5.1. Summary of our approach and findings

Before going into the details of the analysis, we review the main objectives and approach of this test and the main findings. Then in sections 5.2 and 5.3 we provide more details about the analysis.

- **Approach:** To build and validate our method and examine its sensitivity to different assumptions, we created synthetic data from simulated epidemics with several different assumed temperature effects on infection. The true exposure, the exact detection delays, and the temperature functions were hidden from our estimation method to assess the method's success objectively. Our objectives were two fold and we created a task allocation among researchers to meet those objectives: first researcher HR used the iterations of this process to improve our statistical estimation method and ensure that it was able to find temperature functions under various assumptions (Section 5.2.2); second, we used a more realistic individual-level model of infection (stochastic agent-based model) built by two other investigators (NG and MG) not involved in the first synthetic data analysis (used for method design) to assess if our statistician (RX) who was unaware of true functional forms or the new model structure could identify correct effects in this different simulation environment (section 5.3). This design addressed the risk that a method fine tuned on synthetic data may perform well under the assumed simulation setting but fail in other environments.

- **Results:** The key finding from these experiments which are elaborated in details include: 1) The model specification we use can accurately identify correct weather impacts if true infection was observable; 2) In the absence of data on true exposure (which is the case in COVID-19 due to testing delays), however, estimation of weather impact becomes complicated, and many intuitive specifications, often used in other studies, fail to recover true impacts. This may be a serious challenge afflicting many attempts to identify the link between weather and COVID-19 transmission; 3) Our algorithm for uncovering the true exposure, and the specification we selected offer a potentially conservative but qualitatively informative view of the true underlying impacts; 4) Our preferred specification is robust to a few key uncertainties that may vary between simulated numbers and the actual epidemic; and 5) A statistician blinded to true data generating process was able to use this method to identify true weather effects from synthetic data generated from a different, more detailed, agent-based model of COVID-19 epidemics. Given these results from the analysis of the synthetic data, we can have more confidence in the analysis using the actual data.

5.2. Statistical specification using synthetic data

Our approach consists of building a simulation model of epidemic to generate synthetic data (with known weather impact functions) followed by estimating various statistical specification to assess their ability in identifying the true functional forms.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

5.2.1. Simulation model

We used a simple SIR-based simulation model to generate synthetic epidemics. This model was applied across various locations (with different vectors of weather variables ($W(t)$) impacting epidemic curve based on $g(W(t))$) to generate the raw data going into alternative statistical methods to identify the function g below. Equations of the simulation model are presented in Table S8.

Table S8. Equations of the SIR-based stochastic simulation model of epidemics

Iterations completed on a daily basis until epidemic ends or until t=50 is reached.

$I_N(t) = Poisson\left(\frac{bS(t)C_I(t)}{S(0)} g(W(t))\right)$	New infections assumed Poisson based on susceptible stock (S), Infectious stock (C_I), force of infection (b), and weather effect (g(W(t))), primarily focusing on mean temperature.
$S(t+1) = S(t) - I_N(t)$	Updating susceptible stock for next period
$R(t+1{:}t+21) += Multinomial(I_N(t), D_{Rcv})$	Adding to future daily recovery rates (R) to incorporate the recovery of all those infected today using a multinomial distribution. D_{Rcv} is the recovery delay distribution, assumed Poisson with mean of 20 days. The "+=" operator adds to the existing vector on the left-hand side values on the right hand.
$I_M(t+1{:}t+18) += Multinomial(round(I_N(t) * f), D_{Dtc})$	Adding to future measurement of detected cases (I_M) to incorporate fraction *f* of those infected today in future detection data. A multinomial distribution is used following the Detection Delay lag structure (D_{Dtc}; see section 2 of Appendix). In our baseline model we use $f = 0.1$, and test more complex functional forms where f increases over time from zero to a maximum of 0.3 in response to measured infections in 3 different scenarios.
$C_I(t+1) = C_I(t) + I_N(t) - R(t)$	Updating next period stock of infectious based on recovery and new infections. Initial infectious population of 3 is assumed so that few epidemics die out due to stochasticity.
Initialization and functional assumptions	
$S(0) = Max(1000, Normal(5e5, 2e5))$	For different simulated locations we initialize the susceptible population to vary a wide range. In lieu of explicit response mechanisms, these different sizes represent different slow down trajectories in the new infections over time.
$b = \frac{2.5}{20} Normal(1, 0.25) h(W)$	For different simulated locations the baseline force of infection is drawn from a normal distribution, set to have an expected basic reproduction number of 3, and to be potentially correlated with weather variables (using $h(W)$ term) to ensure robustness of results to such correlations.
$g(W)$	We use different *g* functions, including quadratic and linear forms. We focus on temperature as the primary W variable that is read from data and input into the synthetic data.
$h(W)$	In our main synthetic analysis, we use:

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

	$h(W) = \frac{40 + Mean\big(W(1:50)\big)}{50}$ This creates a positive correlation of 0.47 between temperature that affects transmission rates and the reproduction number for a specific location.
W	We use mean daily temperature data from January 1st 2020 for t=1 for all locations, and continue accordingly.

Using these specifications, we then conducted multiple experiments to identify a viable specification. In each experiment we simulated the model for 20 iterations (with different random realizations) for a sample of 1,000 locations randomly drawn from our 3,739 locations with their actual temperature data feeding into W.

We summarize the results from each experiment using a graph of the shape of the estimated relationship between temperature and natural logarithm of reproduction number, i.e., $ln(g(w))$ and compare that with the true relationship (the thick dashed line in figures below). The success measure for our method is to have the estimated relationship from our method close to the thick dashed line. Effects falling between true curve and a horizontal line would be conservative, and those falling outside this range may be misleading. Our actual temperature data in the simulation period is bounded to smaller ranges (90% of data falls between -10 and 20 degree Celsius) than reported in these figures. Therefore, extrapolations outside this range are not necessarily indicated by the data, rather, emerge from the estimated functional forms. Nevertheless, we graph a much wider temperature range (-30 to 50) to highlight the risks of such extrapolation. We also report the means of estimated parameters for $ln(g(w))$, 95% confidence interval, Coefficient of determination (r^2), and sample size for each experiment.

5.2.2. Synthetic experiments

In the rest of this section, unless specified in the top row of a figure, we focus on quadratic equations for $ln(g(w))$. The main specification uses $Ln\big(g(w)\big) = -0.01w - 0.002w^2$. We also conduct sensitivity analysis to other functional forms for $Ln\big(g(w)\big)$. We summarize the results of these synthetic analyses under 10 experiments, which could be categorized in three subsets. Experiments 1 to 3 (sections E1-E3 below) introduce the main challenges in correctly associating weather with reproduction number, concluding that while not an impossible task, the best one can expect from similar efforts may be to find estimates that are conservative but not misleading. Next, in experiments E4 to E6 we show support for the chosen statistical specification against other plausible alternatives. Finally, experiments E7 to E10 show the robustness of the preferred specification to a variety of assumptions.

5.2.2.1. The complexity of inferring the true impact of weather conditions

E1) Endowed with true, deterministic infections, the method finds the correct impact of weather. Relaxing either assumption deteriorates results.

In Figure S5, we compare three different scenarios. In the first two (A and B), the estimation method is provided with the actual true infections (rather than those estimated using the method

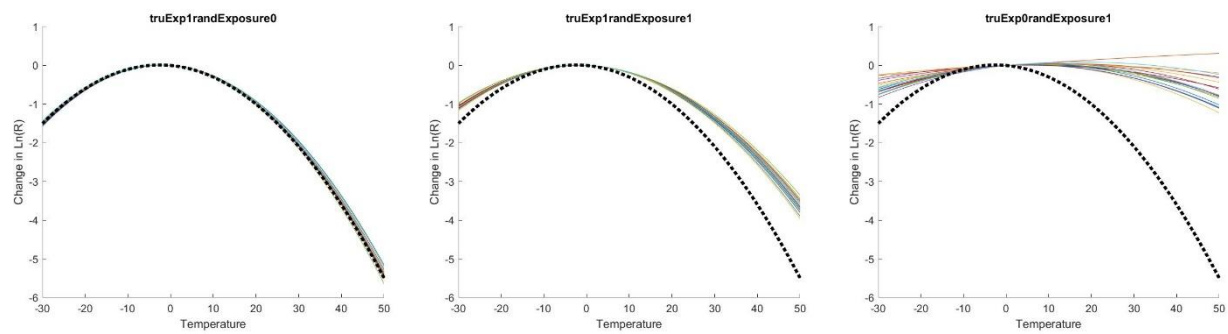
medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

discussed above). Moreover, the first experiment (A) also assumed deterministic infection rate (that is, $I_N(t) = E(I_N(t))$ in Table S). Plot C shows results using our baseline specification: estimating true infections using quadratic programming, including location-specific trend lines but not any fixed effects, dropping days with estimated exposure below 1, as well as the first 20 days after the estimated exposure first reaches 1, and excluding the outlier estimated reproduction numbers (those above 95%). The two assumptions on using deterministic infection and true infections in the first two experiments are not realistic. Instead, they inform the challenges to unbiased estimation of reproduction number due to stochasticity of infections (comparing plots A and B) and estimation of true infections from reported data (comparing plots B and C).

Inspection of these results reveals two major challenges to estimating reproduction number: i) Randomness in infection rate leads to weaker identified effects. ii) The imperfect identification of true infections from reported cases significantly reduces the magnitude of estimated effects. Both of these effects generate a bias towards null estimated effects, even when true effects are very significant. As the experiments reported in the following sections show, our baseline model, despite its conservative estimates, might be among the best available options to find estimates for the impact of weather on transmission rates.

Note that the true linear and square terms in $Ln(g(w))$ are reported in the title for each panel in the following figures.

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$ **(C)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.00941	-0.00907	-0.00873
Quadratic	-0.002	-0.00197	-0.00194
Mean N	5164.4	Mean r^2	0.913588

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.00557	-0.00477	-0.00397
Quadratic	-0.00143	-0.00137	-0.00131
Mean N	21885.7	Mean r^2	0.374341

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.004341	0.00712	0.0099
Quadratic	-0.00064	-0.00042	-0.0002
Mean N	8565.7	Mean r^2	0.477039

Figure S5: Impact of stochasticity in infections and imperfect estimation of true infections on quality of estimated parameters.

E2- Using true infections offers reliable, and slightly conservative, estimates for a range of functions.

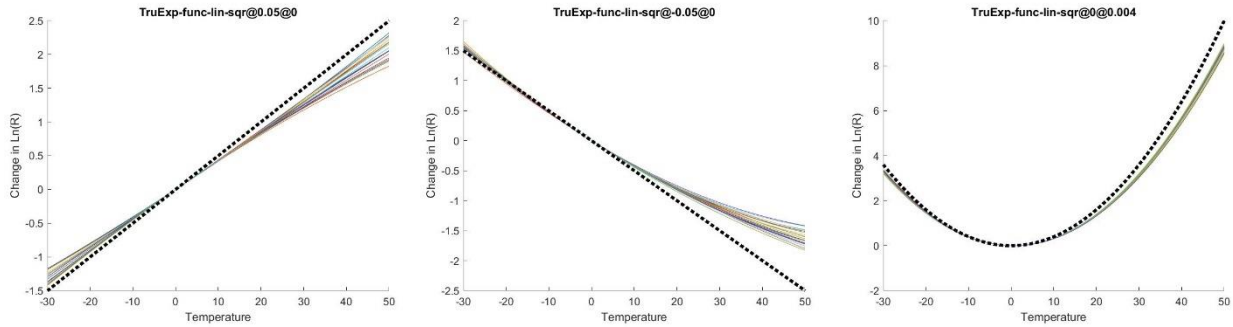
Before focusing on the main specification, we report another set of experiments that show the performance of estimation method with true infections under three other functional forms (Figure S6). These experiments differ from the first set (E1) as they introduce location fixed effects—including location fixed effects, as expected, would add value if true infections were available.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

This is in contrast with the situation when we use estimated infections, which will be reported under section 5.2.2.2 below.

The main observations in this set of experiments are: i) True infection rates, even including randomness in infection, would offer close estimates for the underlying impacts of temperature on reproduction number; ii) The effects are potentially closer to true values because of the use of fixed effects (this can be seen in comparing panel C in these experiments with panel B in Figure S5); and iii) The effects remain slightly conservative, i.e., weaker than true effects (shown in the title for each panel).

(A) $Ln(g(w)) = 0.05T$ **(B)** $Ln(g(w)) = -0.05T$ **(C)** $Ln(g(w)) = 0.004T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.04142	0.043089	0.044759
Quadratic	-0.00012	-4E-05	4.33E-05
Mean N	28854.25	Mean r^2	0.783404

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.0472	-0.04615	-0.0451
Quadratic	0.000194	0.000253	0.000311
Mean N	21259.7	Mean r^2	0.71709

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.00451	-0.00316	-0.00181
Quadratic	0.003495	0.003565	0.003636
Mean N	29404.6	Mean r^2	0.785278

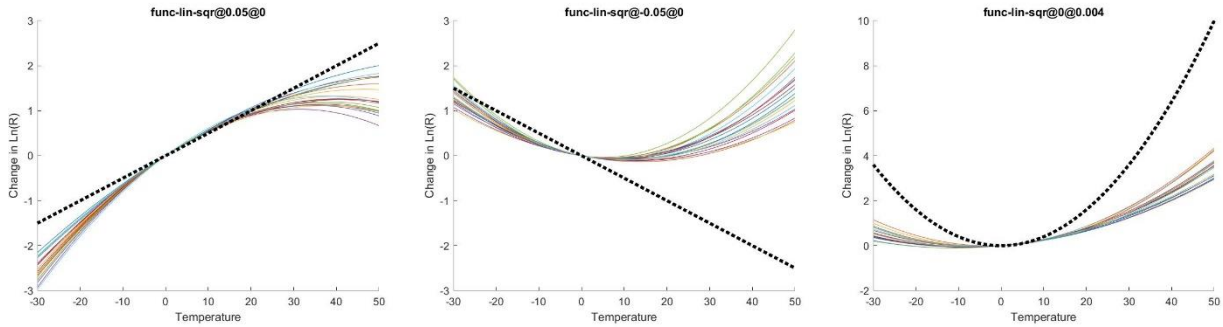
Figure S6: Performance of estimation method, provided with true infections, in identifying different functional relationships between weather and reproduction number.

E3- Using proposed estimation method provides conservative, but largely consistent, estimates.

In the next set of experiments, we test the main estimation method, which uses estimated exposures, to find the effect of temperature under the different functional forms simulated above. In these experiments, we continue to use only a location-specific trend, but no fixed effects, and the same inclusion criteria used in experiment 1. Overall, estimated effects, as shown in Figure S7, are qualitatively consistent with the true functional forms, but also show important deviations: i) The results include some biases in estimated parameters, including quadratic terms that are not in the main function; ii) Results are somewhat conservative (pointing towards null effects) in the regions of the temperature actually covered by W data. Based on these observations, the use of the estimation method should include appropriate caution: the results may under-value the true magnitudes of temperature effects, but may also include unknown biases depending on the true shape of the function.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

(A) $Ln(g(w)) = 0.05T$ **(B)** $Ln(g(w)) = -0.05T$ **(C)** $Ln(g(w)) = 0.004T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.058741	0.06328	0.067819
Quadratic	-0.00109	-0.00089	-0.00068
Mean N	18666.3	Mean r^2	0.589189

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.01823	-0.01585	-0.01347
Quadratic	0.000732	0.000917	0.001102
Mean N	9746	Mean r^2	0.532407

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.012535	0.014942	0.01735
Quadratic	0.000753	0.000897	0.00104
Mean N	20113.6	Mean r^2	0.55263

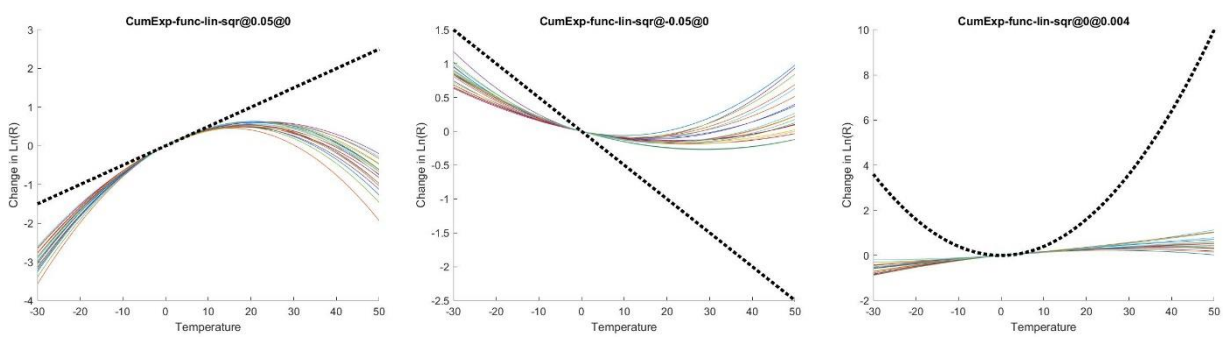
Figure S7: Performance of preferred estimation method with realistically available data in identifying different functional relationships between weather and reproduction number.

5.2.2.2. Comparing with a few alternative specifications

E4- Results with simple shifting of infections

Our preferred specification uses the estimated infections based on the quadratic programming method discussed in section S3. Here we compare those results against a simpler specification that shifts back detected infections each day by 10 days to infer the true infection rate on each day. Results are shown in Figure S8. We assess this alternative under the same functional forms discussed in experiment 3 (E3) and thus results are directly comparable with that experiment. In short, the simple shift method offers results that are largely worse than those in the preferred specification and may include problematic biases. These results motivated the search for alternative estimation methods to more precisely identify the true infections.

(A) $Ln(g(w)) = 0.05T$ **(B)** $Ln(g(w)) = -0.05T$ **(C)** $Ln(g(w)) = 0.004T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)

	Mean (2.5%)	Mean Estimate	Mean (97.5%)

	Mean (2.5%)	Mean Estimate	Mean (97.5%)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Linear	0.05421	0.058296	0.062382
Quadratic	-0.00186	-0.00168	-0.0015
Mean N	21028.65	Mean r^2	0.640187

Linear	-0.02073	-0.01837	-0.01601
Quadratic	0.000141	0.00032	0.000499
Mean N	10164.45	Mean r^2	0.623598

Linear	0.015384	0.017511	0.019638
Quadratic	-0.00019	-6.7E-05	5.21E-05
Mean N	23002.75	Mean r^2	0.63561

Figure S8: Impact of using a simple shift of measured infections on performance of estimation under various functional forms.

E5- Using fixed effects deteriorates results in synthetic data

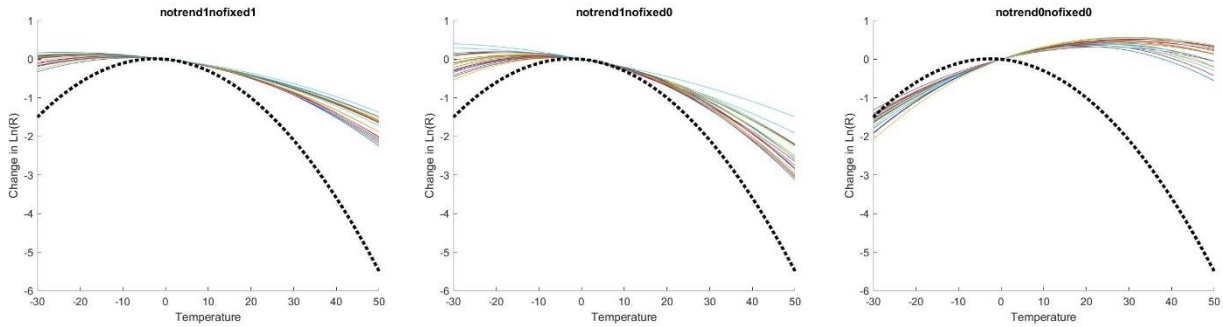
In this experiment, we compare three settings in which we exclude both fixed effects and trends (Figure S9, panel A), exclude trends but include fixed effects (panel B), and include both fixed effects and trends (panel C). These experiments are directly comparable with our preferred specification where trends are included but fixed effects are not (Panel C in Figure S5).

The overall performance is far from perfect in all cases. Moreover, whereas fixed effects typically help with reducing the bias in regression, the combination of both fixed effects and location specific trends offers the most biased results (Panel C). In that setting, two different location-specific parameters absorb much of the variations in weather between and within locations, and combined with errors in the identification of true infections from reported data, the estimates for weather function become unreliable. Removing either of the fixed effects improves the correspondence between true functional forms and actual estimates, leading to conservative but not misleading results. This leaves us with a choice between three specifications that exclude at least one of the fixed effects and/or location specific trends.

We decided to go with inclusion of trends as our main specification. This choice was made with two considerations. First, having collected population density data, we could use those to provide some control for potential correlations between temperature and basic reproduction number due to variations in population density with climate. Second, while our *W* vector in simulations starts from early January, in practice the more positive trend in temperature during the spread of epidemic later in winter and early spring (when most of actual epidemic data is generated) is likely to correlate more strongly with behavioral and other responses that temper down the spread in each location. Thus, we would be worried that excluding location trends would lead regression results to pick up that spurious correlation and inflate the impact of temperature, creating illusory and misleading results. We therefore opted to focus on the more conservative estimates coming from our preferred specification. In practice, however, the inclusion or exclusion of fixed effects had limited impact on our empirical results (See Table S2 and section S4.2). This observation may suggest that our synthetic data generation process might have offered a more challenging setup for estimation than the empirical data generating process.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$ **(C)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.01604	-0.01299	-0.00993
Quadratic	-0.0007	-0.00046	-0.00022
Mean N	8691.8	Mean r^2	0.021457

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.02101	-0.01661	-0.01221
Quadratic	-0.00101	-0.00071	-0.00042
Mean N	8518.25	Mean r^2	0.103033

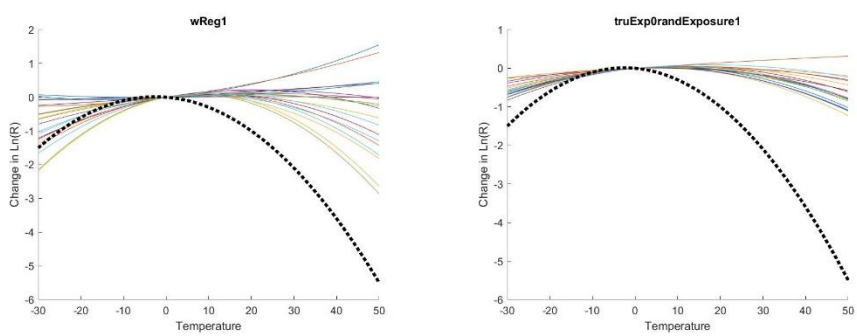
	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.03043	0.034002	0.037573
Quadratic	-0.00091	-0.00067	-0.00043
Mean N	8608.45	Mean r^2	0.582879

Figure S9: Comparing the use of no fixed or trend effects (A), only fixed effects without trends (B), and both location specific fixed and trend effects.

E6- Weighting data points in regression do not improve performance.

One potential issue in the current specification is that locations with larger outbreaks are weighted the same way as locations with smaller outbreaks. The data from larger outbreaks may well be more reliable, and the estimates of R calculated from that data thus more reliable. We assess if a correction for this issue can improve estimation results. To do so we use simulations to estimate how the variance in the dependent variable scales with the number of estimated daily true infections, and use that estimated variance to conduct weighted least square regressions. Results, reported in Figure S10-panel A (panel B showing baseline replicated from experiment 1), show more bias and more dispersion compared to the unweighted regressions. Hence, we do not pursue this correction in our main specification.

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.010824	0.013608	0.016391
Quadratic	-0.00073	-0.00049	-0.00026
Mean N	8602.05	Mean r^2	0.432425

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.004341	0.00712	0.0099
Quadratic	-0.00064	-0.00042	-0.0002
Mean N	8565.7	Mean r^2	0.477039

Figure S10: Impact of using weighted regression (A) vs. unweighted (B)

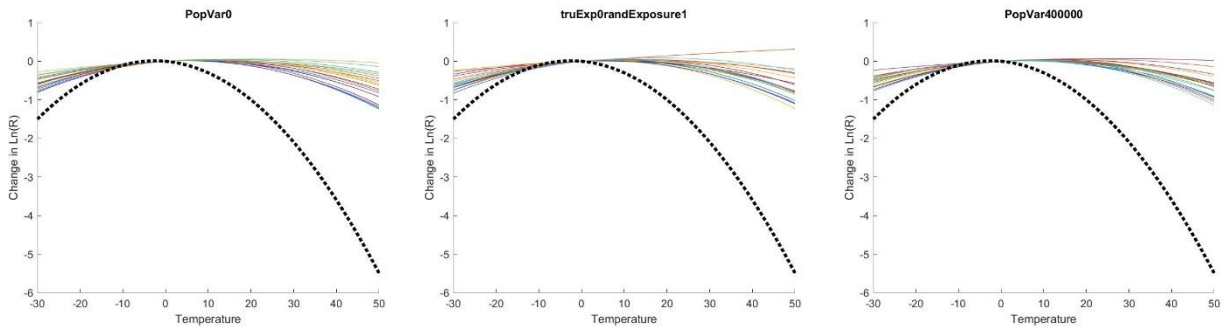
medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

5.2.2.3. Robustness to various features of data

E7- Results are robust to variance in susceptible population.

In the baseline simulations, the location-specific population size, which partially controls the speed of spread and how fast the slow down in reported cases will be realized, is normally distributed with mean of 500,000 and standard deviation of 200,000 (floored at 1,000 population). In Figure S11, we compare that baseline (panel B) against standard deviation of 0 and 400,000. The impacts are small and largely negligible, suggesting that the variance in speed by which the spread slows down does not impact the findings much.

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$ **(C)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.003937	0.006728	0.00952
Quadratic	-0.00064	-0.00041	-0.00019
Mean N	8636	Mean r^2	0.476701

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.004341	0.00712	0.0099
Quadratic	-0.00064	-0.00042	-0.0002
Mean N	8565.7	Mean r^2	0.477039

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.003532	0.006322	0.009112
Quadratic	-0.00061	-0.00039	-0.00016
Mean N	8309.35	Mean r^2	0.471357

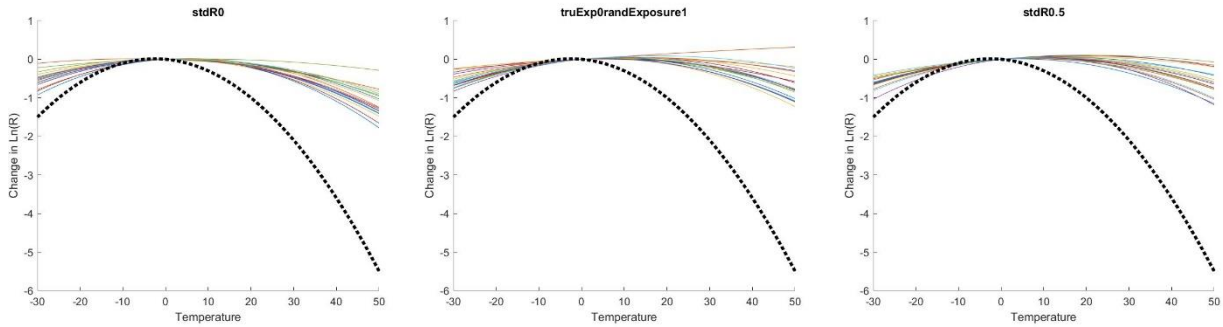
Figure S11: Impact of various variances in population size per location. (A) No variance (mean is 1e5). (B) Baseline (Standard deviation of 200,000). (C) Standard deviation of 400,000.

E8- Variations in basic reproduction do not impact the findings.

In the baseline simulations (E1), the location-specific basic reproduction number had a mean of 3 and variance of 0.9 (normally distributed and floored at 0; but also positively correlated with average temperature in the location). In Figure S12, we compare that baseline (reproduced in panel B) against standard deviations of 0 (Panel A) and 1.5 (Panel C) across locations. The impacts are small and largely negligible, suggesting that the variance in speed by which the spread slows down does not impact the findings much. Results suggest only minor sensitivity to this factor, with slightly better estimates in the absence of variance in basic reproduction number across locations. Given the wide range of basic reproduction numbers explored in this analysis compared to actual numbers, we do not think this issue would pose a robustness challenge to the main findings.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$ **(C)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.00011	0.002434	0.004976
Quadratic	-0.00073	-0.00052	-0.0003
Mean N	7577.25	Mean r^2	0.484981

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.004341	0.00712	0.0099
Quadratic	-0.00064	-0.00042	-0.0002
Mean N	8565.7	Mean r^2	0.477039

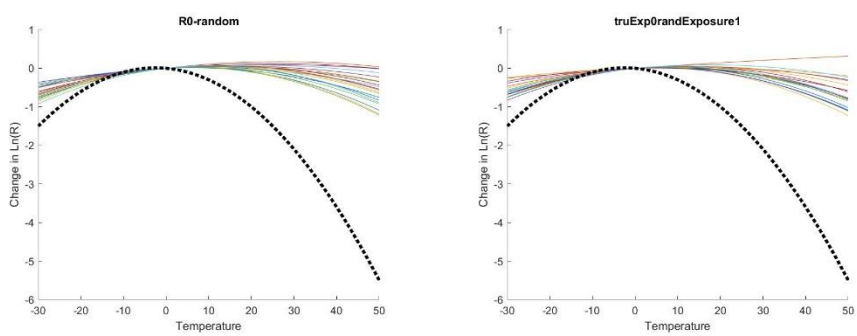
	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.005628	0.008417	0.011207
Quadratic	-0.00063	-0.00042	-0.00021
Mean N	9635.05	Mean r^2	0.455137

Figure S12: Impact of different variances in basic reproduction number. (A) Standard deviation of basic reproduction number is 0, with mean of 3. (B) Standard deviation of basic reproduction number is 0.9. (C) Standard deviation of basic reproduction number is 1.5.

E9-Correlation between basic reproduction rate and temperature has limited impact on results.

Another variant on distribution of basic reproduction number considers the correlation between that parameter and the temperatures informing the weather function. In the baseline specification and all the experiments so far, we used a correlated version of that relationship (with a correlation of 0.47 between basic reproduction number and average temperature; see simulation model specification; Section 5.2.1). Here we compare that setup (reproduced in Figure S13, Panel B) with the uncorrelated version where basic reproduction number is independently drawn for each location with mean of 3 and standard deviation of 0.9 (Panel A). Results are hardly sensitive to this potential correlation.

(A) $Ln(g(w)) = -0.01T - 0.002T^2$ **(B)** $Ln(g(w)) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.006253	0.008695	0.011137
Quadratic	-0.00059	-0.0004	-0.00021
Mean N	11443.95	Mean r^2	0.473288

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.004341	0.00712	0.0099
Quadratic	-0.00064	-0.00042	-0.0002
Mean N	8565.7	Mean r^2	0.477039

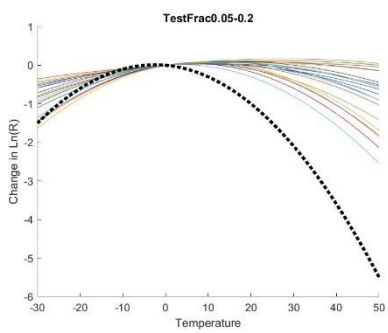
medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Figure S13: Impact of including correlation between basic reproduction number and average location-specific temperature (baseline; reproduced in Panel B) vs. having no correlation (panel A).

E10- Changes in test coverage do not change the results.

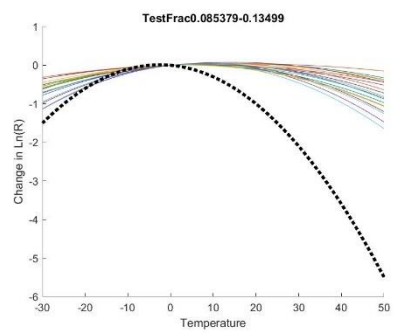
The simulation model so far assumed a constant test fraction of *f*=0.1, that is, in expectation only 10% of infections were detected. In practice this ratio may change over time as test capacity ramps up in response to infection measures. Here, we explore results under three such ramp up scenarios. In Figure S14-panels A-C, the following ramp up scenarios are assumed: A) $f = Min(0.2,0.001I_M)$; B) $f = Min\big(0.2,0.05\mathrm{Log}_{10}(I_M + 1)\big)$; and C) $f = Min(0.2,0.01\sqrt{I_M})$. Overall, results are rather insensitive to these very different test fraction numbers, suggesting robustness to this consideration.

(A) $Ln\big(g(w)\big) = -0.01T - 0.002T^2$



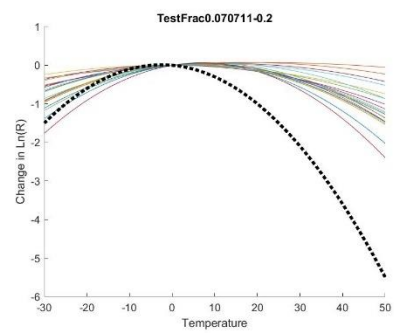
	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.005597	0.011911	0.018224
Quadratic	-0.00108	-0.0006	-0.00011
Mean N	3698	Mean r^2	0.656644

(B) $Ln\big(g(w)\big) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.003701	0.007745	0.011789
Quadratic	-0.00079	-0.00047	-0.00015
Mean N	6089.9	Mean r^2	0.537823

(C) $Ln\big(g(w)\big) = -0.01T - 0.002T^2$



	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.002677	0.007661	0.012644
Quadratic	-0.001	-0.00061	-0.00022
Mean N	4976.05	Mean r^2	0.57214

Figure S14: Impact of three different test coverage functions.

5.2.2.4. Summary of Synthetic Experiments

The key finding from these experiments include: 1) The model specification we use could identify correct weather impacts if true infection was observable; 2) In the absence of that data however estimation of weather impact becomes complicated, and many intuitive specifications fail to recover true impacts. This may be a general challenge afflicting any attempt to identify the link between weather and COVID-19 transmission; 3) The specification we selected may offer a conservative but qualitatively informative view of the true underlying impacts. For example, in most estimations the quadratic term is estimated at about 20% of its true value; and 4) This specification is largely robust to a few key uncertainties that may vary between simulated numbers and the actual epidemic.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

5.3. Blinded study specification and results

In this last step of synthetic data analysis, our objective was to use a more detailed individual-level model of infection (stochastic agent-based model) of interacting individuals, to create synthetic data of reported cases, distort the outcome with a delay function to represent test/report, and examine if our statistical methods are still capable of finding our weather functions. Synthetic data creation was done by two of our co-authors (NG and MG) who hided their assumed temperature effect functions from our statistician (RX) whose task was to discover the assumed temperature effect function.

To that end, we created an agent-based simulation model of infection, and simulated the model for 100 hypothetical towns of different populations, different R_0's (potentially due to different contact rates and population density), different start days of infection, and different temperatures. We started from the generic individual model of infection (available on the NetLogo's library) that is consistent with the basic SIR model at individual level. We modified the model using parameter values that are more consistent with COVID-19, and included several features needed to import and export data to the model. We modified the infection function to include the temperature effect on the probability of infection. We used three major scenarios for temperature effect (inverse u-shaped effect, linear increasing effect, and no-effect (placebo)). The scenarios included actual temperature values coming from a sample of 100 regions from the real-world data. The ABM model's output was generated using a detection delay with Poisson distribution with mean of 10 days. This data was used to estimate true infections with the method discussed in section 3. The model codes are available at https://github.com/marichig/weather-conditions-COVID19/. Figure S12 shows an example of creating synthetic data (scenario 1, explained in the following) with Panel A showing the true cumulative infections and panel B showing the reported values.

A B

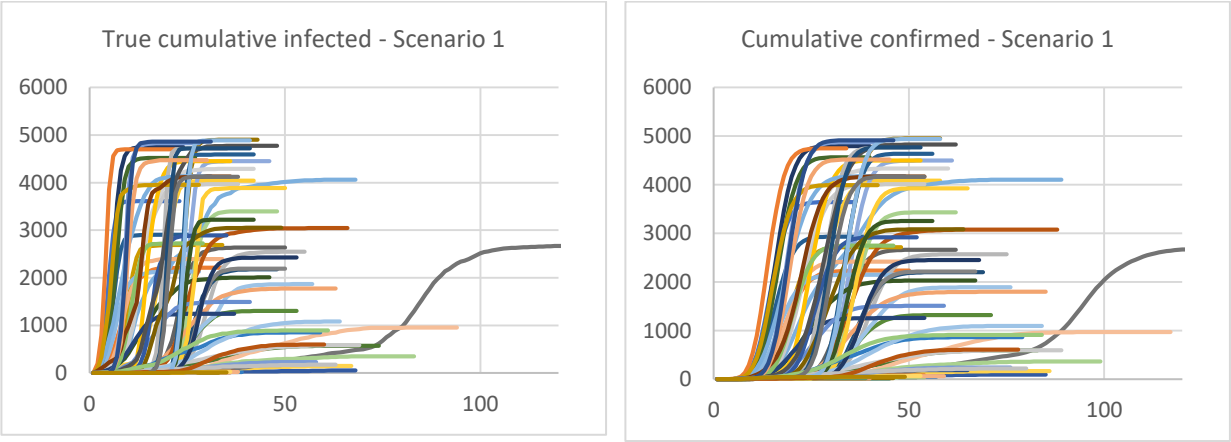


Figure **S12**. An example of the process of generating synthetic data with an assumed temperature hidden from our statistician. True cumulative cases (A) and cumulative number of confirmed cases (B).

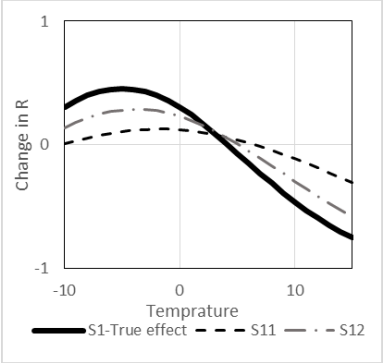
The tests (scenarios) included quadratic (S1), no-effect (S2), and positive linear effect (S3). For all scenarios we tested models both including fixed effects (Si1) and those without (Si2). In non-

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

fixed-effect tests, in order to make a control variable consistent with our main regression, we added one extra variable, a hypothetical variable of "population density", to represent variations in locations correlated with basic reproduction number. In this setup "population density" was correlated with the basic reproduction number excluding temperature-related factors ($\rho = .8$). Our statistician did not know the true temperature function, so they used both linear and quadratic terms to map the predicted temperature effect in all cases, even when the effect was linear.

Results are graphically summarized in Figure S13; for each scenario the results are compared with the "true" function of temperature (darker lines). Overall our statistician was able to correctly estimate the sign and magnitude of temperature effect in all cases, while the effect was generally underestimated further supporting the proposition that the method offers conservative estimates (e.g., in Figure S13, left panel, compare S11, and S12 curves with the true effect of "S1-true effect").

(A) $Ln\big(g(w)\big) = 1 - (\frac{T+5}{15})^2$



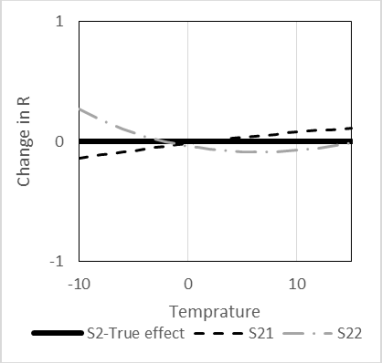
S11 - With fixed-effect

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	-0.0173	-0.0064	0.0045
Quadratic	-0.0022	-0.0017	-0.0011
N	1382	R^2	0.912

S12 - No fixed-effect

	(2.5%)	Estimate	(97.5%)
Linear	-0.0348	-0.0238	-0.0128
Quadratic	-0.0037	-0.0032	-0.0027
N	1382	R^2	0.683

(B) $Ln\big(g(w)\big) = 0$



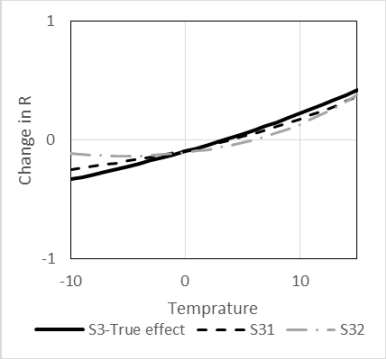
S21 - With fixed-effect

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.0028	0.0112	0.0195
Quadratic	-0.0005	-0.0002	0.0001
N	2496	R^2	0.893

S22- - No fixed-effect

	(2.5%)	Estimate	(97.5%)
Linear	-0.0239	-0.0158	-0.0078
Quadratic	0.0009	0.0012	0.0015
N	2496	R^2	0.766

(C) $Ln\big(g(w)\big) = 0.03T$



S31- - With fixed-effect

	Mean (2.5%)	Mean Estimate	Mean (97.5%)
Linear	0.0152	0.0223	0.0294
Quadratic	-0.00003	0.0003	0.0005
N	2206	R^2	0.881

S32- - No fixed-effect

	(2.5%)	Estimate	(97.5%)
Linear	0.0062	0.0121	0.018
Quadratic	0.0008	0.0011	0.0014
N	2206	R^2	0.73

Figure S13: A comparison of assumed functions in the agent-based model (dark lines) and the outcome of our regression analysis (dashed lines). Y-axis is change in R in comparison to average R. Note: S_{ij} represents results from Scenario *i* under conditions of fixed-effect (j=1), and no fixed effect regressions (j=2).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

6. Global projections over time

In this section we report graphs of projected Covid-19 Risk Due to Weather (CRW) at 4 different time periods in the coming year. These projections use a 15-day moving window to average different weather variables in the previous year (2019-2020), and use those averages as the predictor for the coming year (2020-2021). Daily projections year-round, both for these global locations and the largest 1072 cities across the world are available at: https://projects.iq.harvard.edu/covid19.

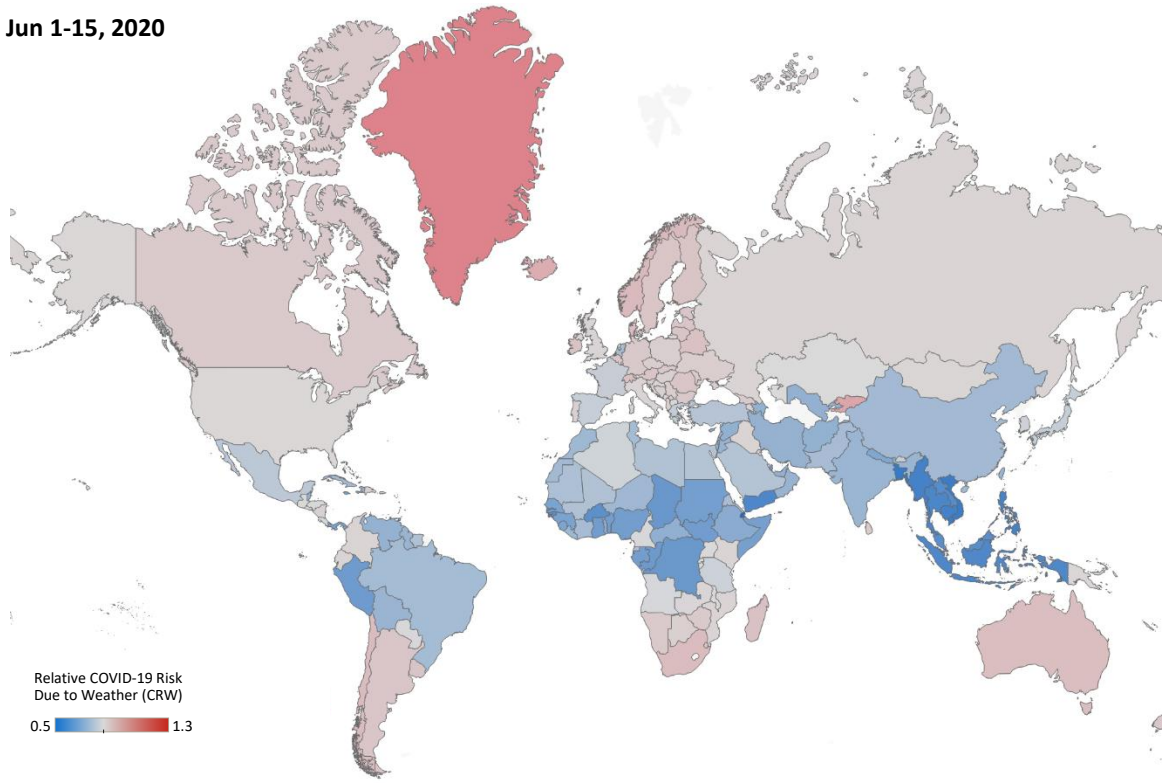


Figure S15: Relative COVID-19 Risk Due to Weather (CRW), averaged for the first half of June 2020

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

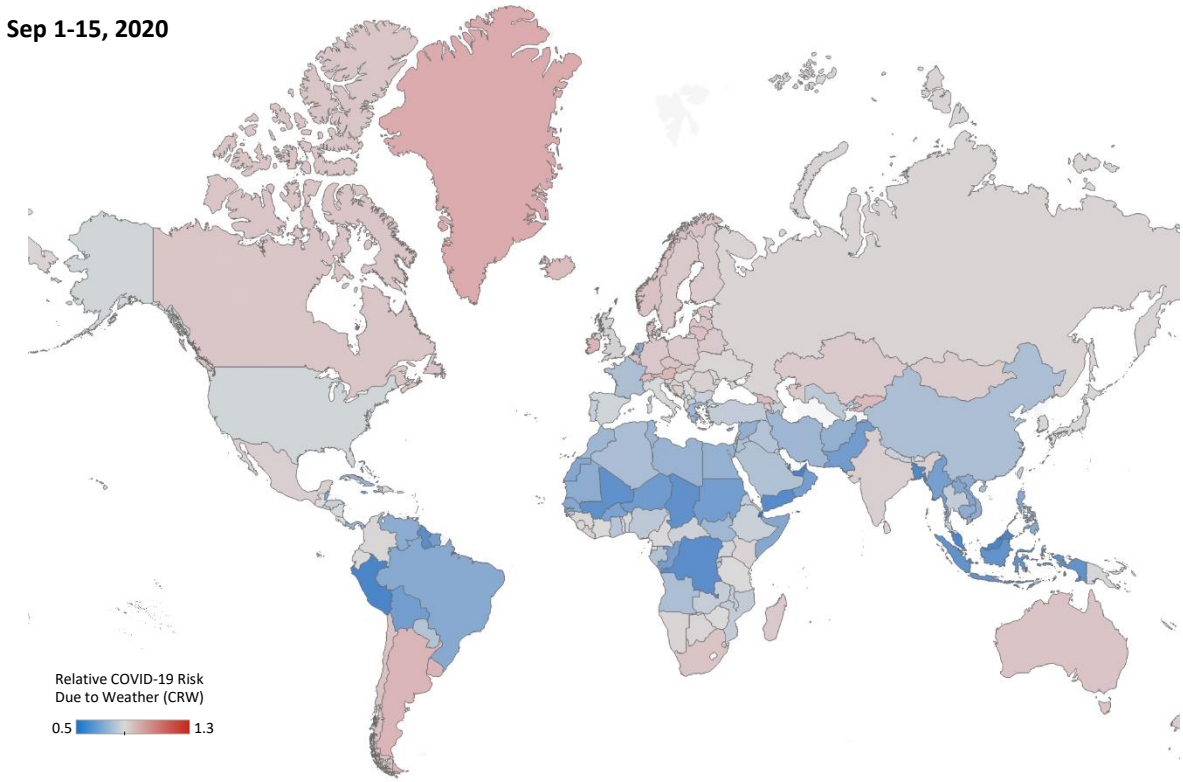


Figure S16: Projected Relative COVID-19 Risk Due to Weather (CRW), averaged for the first half of September 2020

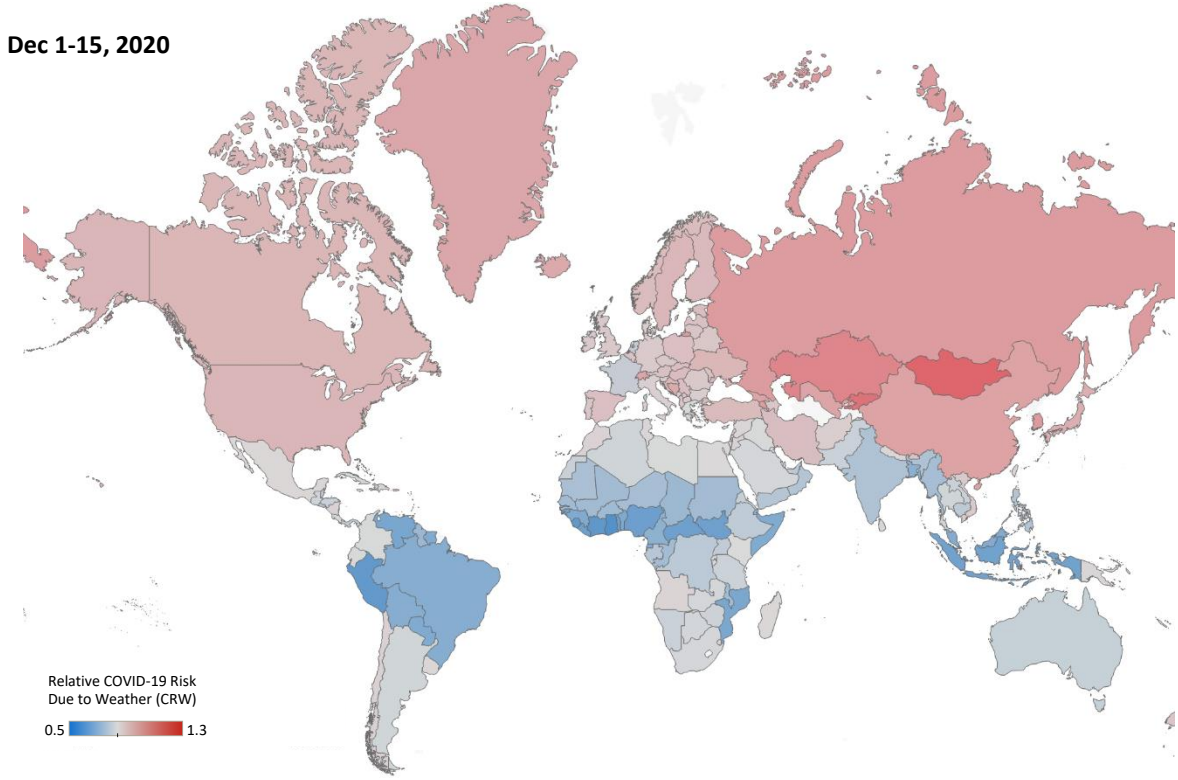


Figure S17: Projected Relative COVID-19 Risk Due to Weather (CRW), averaged for the first half of December 2020

medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20092627; this version posted May 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

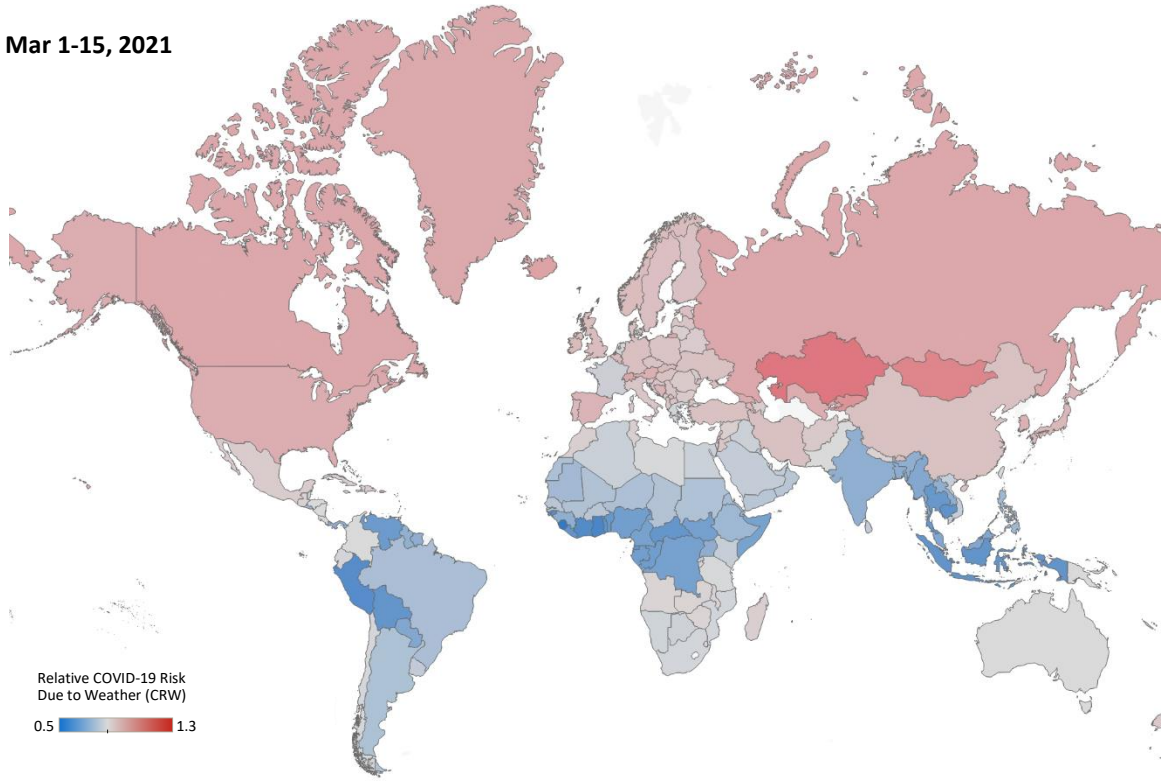


Figure S18: Projected Relative COVID-19 Risk Due to Weather (CRW), averaged for the first half of March 2021

References:

1. Li, Q., et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia.* New England Journal of Medicine, 2020.
2. Lauer, S.A., et al., *The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.* Ann Intern Med, 2020.
3. Guan, W.J., et al., *Clinical Characteristics of Coronavirus Disease 2019 in China.* N Engl J Med, 2020.
4. Linton, N.M., et al., *Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data.* J Clin Med, 2020. **9**(2).
5. Popovich, N., *How U.S. Coronavirus Diagnoses Are Lagging Behind the Outbreak*, in *The New York Times*. 2020: New York.
6. Wibbens, P.D., W. Koo, and A.M. McGahan, *Projected COVID Infections, Deaths, and Local Social-Distancing Restrictions.* Deaths, and Local Social-Distancing Restrictions (April 17, 2020), 2020.