1

2 Social-economic drivers overwhelm climate in underlying the COVID-19 early

3 growth rate

4

5 Zhenghua Liu^{1,*}

6

7 ¹School of Minerals Processing and Bioengineering, Central South University,

8 Changsha 410006, China

9 *Corresponding authors:

10 Zhenghua Liu, liuzhenghua2017csu@163.com

11

12 **Running title:** Socio-economic and climate factors driving COVID-19 spread

13

14

15 **Key findings**

16 1) GDP per capita is most important in driving the spread of COVID-19, which

17 shows negative relationships within China, while positive in the other countries.

18 2) Socio-economic and climate factors are key in driving the early growth rate of

19 COVID-19, while the former is more important.

20 3) Socio-economic and climate features explain more variations of early growth

21 rates in China due to the decreased stochasticity of transmission processes upon strict

22 quarantine measures.

23

24 Abstract

25 Identifying the drivers underlying the spatial occurrence and spreading rate of
26 COVID-19 can provide valuable information for their preventions and controls. Here,
27 we examine how socio-economic and climate drivers affect the early growth rates of
28 COVID-19 in China and the other countries, the former of which have consistently
29 stricter quarantine during early epidemic and thus are used to enquire the influences
30 of human interventions on transimissions. We find that the early growth rates of
31 COVID-19 are higher in China than the other countries, which is consistent with
32 previous reports. The global spread is mainly driven by the socio-economic factors
33 such as GDP per capita, human movement and population density rather than climate.
34 Among socio-economic factors, GDP per capita is most important showing negative
35 relationships in China, while positive in the other countries. However, the
36 predicability of early growth rates by socio-economic and climate variables is at least
37 1.6 times higher in China's provinces than the other countries, which is further
38 supported by metapopulation network model. These findings collectively indicate that
39 the stochasticity of transimission processes decrease upon strict quarantine measures
40 such as travel restrictions.

41

42 The coronavirus disease 2019 (i.e., COVID-19) has developed into a pandemic
43 with an unprecedented speed since its outbreak in 2019 and exerts a great threat for
44 public health at the global scale [1]. Geographical and climate factors are frequently
45 reported to be important for the spread of COVID-19, such as latitude, temperature,
46 humidity [2-5], and seasonality [5]. This is not surprising as climate is usually related
47 the plague infection activity, such as SARS [6], MERS [7], Ebola [8], Zika [9] and
48 influenza [10].

49 Socio-economic developments, however, could also profoundly influence
50 epidemic transmission dynamic [10-12]. This is indicated by the historical
51 co-occurrence of increasing frequency of epidemics with recent human population (Fig.
52 S1), which is largely because the socio-economic status relevant to rapid population
53 growth [13], human movements [14, 15] and uncontrolled urbanisation [10] could
54 provide conditions and environments conducive to vector proliferation and plague
55 transmissions [16]. For instance, at regional scale, COVID-19 spread is promoted by
56 city size in USA [17] and air pollutants of nitrogen dioxide [18] and PM2.5 [19], and
57 is constrained by control measures, such as travel restrictions [20]. At global scale,
58 COVID-19 cases could be seeded globally [15, 20, 21] via the well-developed traffic
59 networks composed of railways, highways, waterways and airways across cities,
60 provinces, states and countries. With the increasing movement of infected humans,
61 virus spread and persistence are improved and the spatial dimensions of transmission
62 are extended from small household foci to large community clusters [22, 23], thus
63 leading to the recent expansions of COVID-19. Identifying socio-economic factors

64 and their interactions with climate in underlying the worldwide COVID-19 spread
65 will be helpful to understand its spatial patterns and provide useful information
66 regarding its prevention, mitigation and control.

67 Here, we estimated and compared the early growth rates of COVID-19 between
68 China and the other countries because the former could be regarded as a treatment
69 group in respective of its stricter measures in social distancing and travel restrictions
70 especially during early epidemic. For each region at the province level of China or the
71 state level of the other countries, we calculated a series of growth rates based on
72 exponential model with extensible windows starting from the first confirmed case of
73 COVID-19 and regarded the maximum value as early growth rate (Fig. S2, S3). The
74 maximum growth rate is expected to locate on the inflection point of the slope in
75 logistic growth curve during the exponential phase, where the growth of COVID-19 is
76 free to control measures and natural resource limitations such as susceptible
77 population, and thus is mostly closed to the theoretical growth characteristics of virus
78 [24]. Different from the other counties, the growth rates in China were estimated at
79 the province level, since strict quarantine was consistently applied across the country
80 although there is a high spatial heterogeneity in climate and socio-economic status.
81 Further, we explored the underlying drivers of COVID-19 early growth rates with the
82 compiled 30 socio-economic and 16 climate factors (Table S1) based on 33 China's
83 provinces and 29 countries outside of China after. More details of growth rate
84 estimate and environmental variables are shown in supplementary materials and
85 methods.

86 As of 26th Mar 2020 when the daily COVID-19 cases data were collected, there
87 were more than 82,034 and 447,557 infected individuals confirmed in China and the
88 other counties, respectively (Fig. S4a). The early growth rates in China were $0.495 \pm$
89 0.189 (mean $\pm$ SD), and significantly higher than the other countires (0.335 ± 0.167 ;
90 two-side T-test, $P < 0.001$; Fig. S4b). This result agrees with the previous studies
91 showing the higher basic reproduction number in China (3.38 ± 1.41) than the other
92 countries (2.92 ± 1.02 ; Table S2) and the lower doubling time in China (2.16 days)
93 than USA (2.46 days) and European countries (2.63 days) [25]. Outside of China,
94 there was a significantly negative correlation between early growth rate and
95 temperature or UV radiation (all $P < 0.05$; Table S3; Fig. S5), which is consistent with
96 a significant geographic pattern of the increasing early growth rate towards high
97 latitude (Pearson's $r = 0.383$, $P = 0.04$). However, the above relationship between
98 early growth rate and temperature or latitude was not observed in China ($P > 0.05$;
99 Fig. S5; Table S4). Additionally, it was worth noting that there were significantly
100 unimodal relationships between growth rates and precipitations ($P < 0.05$; Fig. S5)
101 across the other countries, but not in China. Such non-significant effects of climate
102 variables in China are consistent with recent studies that the COVID-19 spread in
103 China was independent on temperature and humidity [26, 27]. But it is contrasting
104 with other reports showing that climate factors such as temperature and precipitation
105 determine the spatial and temporal variations in the incidences and transimissions of
106 most epidemics [28], such as malaria [16], cholera [29], and Ebola [30].

107 Compared to climate, socio-economic factors played more important roles in
108 driving the spread of COVID-19 (Fig. 1, S4b, S6). Among them, GDP per capita was
109 most important (Fig. 1a) and explained 19.5% and 8.9% of the variations of early
110 growth rate in China and the other countries (Fig. 1b), respectively. In China, GDP
111 per capita negatively correlated with early growth rate (Fig. S6; Table S4), while
112 positively for the other countries (Fig. S6; Table S3). Such contrasting effects of GDP
113 per capita may be resulted from the control measures such as travel restrictions,
114 which disabled the human mobility and COVID-19 transmissions [20]. This is further
115 supported by the commuting rate (Fig. S6), which showed no relationships with the
116 growth rates in China, but significantly increased the rate in the other countries ($r =$
117 0.543 , $P = 0.002$; Fig. S6; Table S3). This result indicates that in the regions with high
118 GDP per capita, COVID-19 spread is promoted by attracting more immigrations [31]
119 and increasing commuting rate (Fig S7), but COVID-19 transmission could be also
120 effectively prevented by the high investments on human interventions such as travel
121 restrictions and isolation of suspected cases, confirmed cases and contacts.

122 Overall, climate and socio-economic factors jointly explained more spatial
123 variations in early growth rate for China and the other countries, that is 54.3% and
124 32.8%, respectively (Fig. 1b). It means that the predictive certainty of COVID-19
125 transmissions under strict interventions in China was 1.6 times higher than the other
126 countries with less interventions. Such a finding shows that the aspects of
127 containment measures, public health policies and immigration monitoring may inhibit
128 the local transmission dynamics of COVID-19 [2], which are consistent with the

findings in the spread of influenza [10] and Zika [9]. In addition, the low explained variations in early growth rate for the other countries may be caused by high mobility which makes local transmission dynamics to be homogeneous [32].

To further test the effects of strict quarantine on the spatial variations of COVID-19 growth rate, we implemented the Global Epidemic and Mobility Model (GLEAM), an stochastic and individual-based model using a metapopulation network to simulate epidemic spread [15, 33, 34]. In the metapopulation networks, nodes or subpopulations were connected by traveller fluxes, and characterized by random socio-economic and climate factors which shape the exponential growth rate of COVID-19 (Fig. 2a). As expected, GLEAM results show that the explained variations of COVID-19 growth rate decreased towards the elevated quarantine level (ANOVA, $P < 0.001$; Fig 2b). This is consistent with empirical results and indicates strict quarantine to enhance the predictability of COVID-19 transmission. In general, high mobility would disturb the heterogeneous structure of metapopulation network [35], cause the increase in synchronization among subpopulations due to the strong spatial coupling in epidemic transmissions, like the spread of influenza in European [36], and thus decrease the predictability in epidemic transmissions [37, 38].

Although we here show a clear picture on how COVID-19 growth rate was shaped by climate and socio-economic factors and constrained by strict quarantine, two major caveats could be acknowledged regarding the analytical framework. First, COVID-19 growth rate could be estimated by various methods, such as exponential [4, 17, 27] or logistic model [39, 40], and basic reproductive number R_0 [17, 26, 41].

151 Here, we chose exponential model with extensible windows and determined
152 COVID-19 growth rate as maximum growth rate in logistic growth curve during the
153 exponential phase. In this phase, COVID-19 spread was majorly dependent on
154 intrinsic growth rate of virus, regardless of susceptible population resources
155 associated with external control measures, such as quarantine and vaccination, and
156 thus the maximum growth rate is more accurate to explore the nature potential of
157 virus. We did not consider logistic model as it failed in estimating COVID-19 growth
158 rate during early epidemic. For other models, such as exponential model with moving
159 windows and R_0 estimated by maximum likelihood method, we obtained similar
160 conclusions as follows. (i) Socio-economic factors played more important role in
161 COVID-19 spread than climate (Fig. S8) and (ii) the variations in early growth rate
162 were explained more in China than the other countries (Fig. S8c).

163 Second, the growth rate may be affected by the data source of reported cases. For
164 example, not all infected individuals could be tested due to the insufficient capacity for
165 testing and tracing or asymptomatic cases, which would cause the reported cases less
166 than the number of infected individuals. Further, the imported cases count a
167 considerable proportion of reported cases at the early stage, which may disturb the
168 characteristics of local transmission dynamics in COVID-19 spread. Besides, the
169 protocols for reporting COVID-19 cases may vary across countries, which may result
170 in unmeasured spatial variations resulted from data collection, processing and
171 summary.

172

173 **Methods and materials**

174 **Data sources**

175 We analyzed the global time series data on infections from January 22th, 2020 to
 176 March 26th, 2020 in Ding Xiang Yuan [42]. We only used the regions with at least
 177 one reported case of COVID-19 and five data points to remain in the time series after
 178 truncation in order to proceed with model fitting, which limited our analysis to 155
 179 regions (that is, 121 countries and 34 China's provinces). We further collected 16 and
 180 30 environmental and social-economic variables, respectively, as explanatory factors
 181 for COVID-19 spread (Table S1). Climate data at country-level and China's
 182 province-level, such as the monthly average, highest and lowest temperatures, and the
 183 monthly average precipitation, were collected from Weatherbase [43]. The variables
 184 related to area, elevation, population size and GDP of China's provinces were
 185 available from China's National Bureau of Statistics [44]. For the other countries, the
 186 social-economic variables such as population, GDP, body mass index (BMI), air
 187 quality and climate variables of UV level, were collected from World Health
 188 Organization [45]. We estimated the commuting rate of China's provinces based
 189 human migration follows from January 1st, 2020 to January 23th, 2020 available from
 190 Github repository of <https://github.com/2020-nCoV/RawData>. The commuting rates
 191 of other countries were eastimated by open-source data of air traveller flow in
 192 February 2010, and are available from
 193 <https://sites.google.com/site/liangmaoufl/data-portal>.

194

195 **Statistical inferences in the early growth rate of COVID-19**

196 During early epidemic of COVID-19, the cumulative cases are usually
197 characterized by exponential growth (refs). Thus, here, the early growth rate r of
198 COVID-19 in each region was estimated by the slope of $\ln(cases) = \ln(a) + rt$ line
199 via ordinary least squares linear regression, where $cases$ is the cumulative reported
200 cases of COVID-19 at time t and a is a constant. We defined the extensible
201 windows as the various durations (with a minimum value of five) of time series data
202 starting from the day of the first record, calculated a series of growth rate r for each
203 window and determined the growth rate of COVID-19 as the maximum and
204 statistically significant value ($P < 0.05$). To identify the robustness of estimations on
205 our conclusions, we also determined the growth rates with another four approaches as
206 follows. (1) In the exponential model with extensible windows, we also determine the
207 growth rate as the second, third and fourth maximum r values and the average of top
208 four maximum values. (2) We fitted the exponential model with moving windows that
209 have a fixed size ranged from 7 to 14 and determined early growth rate as the
210 maximum estimation for each moving window. (3) The basic reproduction number R_0
211 was calculated by maximum likelihood estimation [46]. The serial interval of
212 COVID-19 in each region was equal to that in Wuhan, China, with a mean of 7.5 days
213 and a standard deviation of 3.4 days [47] and the maximum reproduction number was
214 assumed to be eight. (4) we also fitted the logistic model with extensible and moving
215 windows, but it was failed to estimated early growth rate according to P value and
216 determinant coefficients R^2 , and thus it was not further considered.

Quarantine effects on the predictability in the early epidemic growth rates

To examine the effects of travel restrictions on the predictability in epidemic growth rate, we stochastically simulated the COVID-19 spread with the Global Epidemic and Mobility Model (GLEAM). GLEAM model is an individual-based, stochastic, and spatial epidemic model [15, 33]. Metapopulation networks of GLEAM were aggregated by N subpopulations, in which there is 100,000 susceptible people and the specifically exponential growth rate r driven by local climate (C) and socio-economic (S) factor, as followed:

$$r = aC + bS + \epsilon,$$

where a and b are constant and ϵ is random error generated by the normal distribution of $N(0, 0.01)$. Both of climate and socio-economic factors were generated by the uniform distribution of $U(0.5, 1)$. Subpopulations were connected by the traveller fluxes. The flux rate c_{ij} , that is travellers from subpopulation i to subpopulation j account for the population size of subpopulation i , was generated from the logarithmic normal distribution of $LN(-5, 1)$.

Human-to-human transmission of epidemic within a subpopulation was modeled by SEIR model that include four compartments including susceptible, latent, infectious and recovered [15]. Individuals in susceptible compartments can acquire the virus through contacts with individuals in infectious compartment and were transferred into latent compartment with a rate R_0/D_I , where R_0 is basic reproduction number and D_I is mean infectious period. Latent individuals progress to the infectious stage with a rate $1/D_E$, where D_E is mean latent period, and infectious

individuals progress into the recovery stage with a rate $1/D_I$. Notably, R_0 was obtained through the growth rate r with the function of $R_0 = (1 + rD_E)(1 + rD_I)$ [48]. According to previous study [21], the latent period D_E was assumed to be 6.1 days, while infectious period was 1.5 days. Model simulations were performed in Python 3.6.3 [49], while analyses were in R 3.6.1 [50]. Code files are available on https://github.com/zhenghualiu/COVID-19/tree/master/metapopulation_model.

Statistical analyses

We considered 21 China's provinces and 26 countries outside of China for downstream analyses. The reasons for selecting these 26 countries are as followed. (i) In these regions, the exponential growth periods of COVID-19 outbreaks were in March and the growth rate could be characterized by average temperature and precipitation of that month. (ii) These countries show the similar area size ranges to the provinces of China (Fig. S5) and could minimize the spatial scale effects on growth rates. (iii) The shared socio-economic variables among these countries were obtained thus ensuring data integrity. (iv) The similar number of sample points between China and the other countries could reduce the sample size heterogeneity in statistical analyses for solid comparison.

The relationships between COVID-19 growth rate and environmental or socio-economic factors were examined by Pearson's correlations. To select appropriate predictors from 46 potential variables for determining a minimal adequate model, we filtered the independent variables with high collinearity by showing the absolute

values of Pearson's correlation greater than 0.8. By using random forest model with 500 classification trees, we defined the unimportant variable as a predictor with a relative importance less than 0.5% and excluded the most unimportant variable from exploratory variables. This process was repeated until there are no unimportant variables. Then, the random forest model [51] with the final set of exploratory variables were used to estimate the relative importance of variables to COVID-19 growth rate based on 500 classification trees with 10 cross validations.

The variation in early growth rate explained by variables were estimated by redundancy analysis with the hierarchy algorithm of Chevan and Sutherland [52, 53]. To evaluate the influences of climate and socio-economic factors on COVID-19 spread, the variance in COVID-19 growth rate was further partitioned by two explanatory matrices composed of climate and socio-economic factors with distance-based redundancy analysis [54]. Further, we constructed the structural equation model (SEM) with maximum-likelihood estimation procedures [55] to quantify the influences of climate and socio-economic latent variables on COVID-19 growth rate. The exploratory variables were shared by both province and state level, including monthly average temperature and precipitation, GDP per capita, population size, population density and commuting rate. The latent variable of climate was measured by monthly average highest temperature and precipitation, with the loading on monthly average highest temperature constrained to be 1. The socio-economic latent variable was measured by GDP per capita, population, population density and commuting rate, with the latent variable variance fixed to 1 and the loading on GDP

283 per capita constrained to be 1. To acquire minimal adequate model, only the
284 covariances among variables with statistical significance ($P < 0.1$) were kept for SEM
285 fitting (Table S5, S6). The goodnesss of fitting were estimated by *Chi-square* value,
286 the comparative fit index (*CFI*) and the root mean square error of approximation
287 (*RMSEA*).

288 All above analyses were conducted in R 3.6.1 environment [50]. R codes are
289 available on <https://github.com/zhenghualiu/COVID-19/tree/master/Rcodes>.

290

291 References

- 292 1. Anderson, R.M., et al., *How will country-based mitigation measures influence the course of*
293 *the COVID-19 epidemic?* The Lancet, 2020. **395**(10228): p. 931-934.
- 294 2. Oliveiros, B., et al., *Role of temperature and humidity in the modulation of the doubling time*
295 *of COVID-19 cases.* medRxiv, 2020.
- 296 3. Ficitola, G.F. and D. Rubolini, *Climate affects global patterns of COVID-19 early outbreak*
297 *dynamics.* medRxiv, 2020.
- 298 4. Notari, A., *Temperature dependence of COVID-19 transmission.* arXiv preprint
299 arXiv:2003.12417, 2020.
- 300 5. Sajadi, M.M., et al., *Temperature and latitude analysis to predict potential spread and*
301 *seasonality for COVID-19.* Available at SSRN 3550308, 2020.
- 302 6. Chan, K., et al., *The effects of temperature and relative humidity on the viability of the SARS*
303 *coronavirus.* Advances in virology, 2011. **2011**.
- 304 7. Van Doremalen, N., T. Bushmaker, and V. Munster, *Stability of Middle East respiratory*
305 *syndrome coronavirus (MERS-CoV) under different environmental conditions.*
306 *Eurosurveillance*, 2013. **18**(38): p. 20590.
- 307 8. Ali, H., et al., *The social and political dimensions of the Ebola response: Global inequality,*
308 *climate change, and infectious disease,* in *Climate change and health.* 2016, Springer. p.
309 151-169.
- 310 9. Harris, M., J.M. Caldwell, and E.A. Mordecai, *Climate drives spatial variation in Zika*
311 *epidemics in Latin America.* Proc Biol Sci, 2019. **286**(1909): p. 20191578.
- 312 10. Dalziel, B.D., et al., *Urbanization and humidity shape the intensity of influenza epidemics in*
313 *US cities.* Science, 2018. **362**(6410): p. 75-79.
- 314 11. Risse, G.B., *Epidemics and history: ecological perspectives and social responses.* AIDS: The
315 *burdens of history*, 1988. **33**.
- 316 12. Tian, H., et al., *Urbanization prolongs hantavirus epidemics in cities.* Proc Natl Acad Sci U S
317 A, 2018. **115**(18): p. 4707-4712.
- 318 13. Teurlai, M., et al., *Socio-economic and Climate Factors Associated with Dengue Fever Spatial*
319 *Heterogeneity: A Worked Example in New Caledonia.* PLoS Negl Trop Dis, 2015. **9**(12): p.
320 e0004211.
- 321 14. Dalziel, B.D., B. Pourbohloul, and S.P. Ellner, *Human mobility patterns predict divergent*
322 *epidemic dynamics among cities.* Proc Biol Sci, 2013. **280**(1766): p. 20130763.
- 323 15. Chinazzi, M., et al., *The effect of travel restrictions on the spread of the 2019 novel*
324 *coronavirus (COVID-19) outbreak.* Science, 2020: p. eaba9757.
- 325 16. Tompkins, A.M., *Relative importance of climatic, geographic and socio-economic*
326 *determinants of malaria in Malawi.* Malaria Journal, 2013. **12**(1): p. 416.
- 327 17. Stier, A., M. Berman, and L. Bettencourt, *COVID-19 attack rate increases with city size.*
328 *Mansueto Institute for Urban Innovation Research Paper Forthcoming*, 2020.
- 329 18. Yao, Y., et al., *Ambient nitrogen dioxide pollution and spread ability of COVID-19 in Chinese*
330 *cities.* medRxiv, 2020: p. 2020.03.31.20048595.
- 331 19. Wu, X., et al., *Exposure to air pollution and COVID-19 mortality in the United States: A*
332 *nationwide cross-sectional study.* medRxiv, 2020: p. 2020.04.05.20054502.
- 333 20. Kraemer, M.U., et al., *The effect of human mobility and control measures on the COVID-19*
334 *epidemic in China.* Science, 2020.

- 335 21. Wu, J.T., K. Leung, and G.M. Leung, *Nowcasting and forecasting the potential domestic and*
336 *international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling*
337 *study*. The Lancet, 2020. **395**(10225): p. 689-697.
- 338 22. Lieberman, K.A., et al., *Spatial Dimensions of Dengue Virus Transmission across Interepidemic*
339 *and Epidemic Periods in Iquitos, Peru (1999–2003)*. PLOS Neglected Tropical Diseases, 2012.
340 **6**(2).
- 341 23. Stoddard, S.T., et al., *House-to-house human movement drives dengue virus transmission*.
342 *Proceedings of the National Academy of Sciences of the United States of America*, 2013.
343 **110**(3): p. 994-999.
- 344 24. Ma, J., *Estimating epidemic exponential growth rate and basic reproduction number*. Infect
345 Dis Model, 2020. **5**: p. 129-141.
- 346 25. Kochanzyk, M., F. Grabowski, and T. Lipniacki, *Accounting for super-spreading gives the*
347 *basic reproduction number R_0 of COVID-19 that is higher than initially estimated*. medRxiv,
348 2020: p. 2020.04.26.20080788.
- 349 26. Luo, W., et al., *The role of absolute humidity on transmission rates of the COVID-19 outbreak*.
350 2020.
- 351 27. Jamil, T., et al., *No Evidence for Temperature-Dependence of the COVID-19 Epidemic*.
352 medRxiv, 2020.
- 353 28. Parham, P.E., et al., *Climate, environmental and socio-economic change: weighing up the*
354 *balance in vector-borne disease transmission*. Philosophical Transactions of the Royal Society
355 B: Biological Sciences, 2015. **370**(1665): p. 20130551.
- 356 29. Olago, D., et al., *Climatic, Socio-economic, and Health Factors Affecting Human*
357 *Vulnerability to Cholera in the Lake Victoria Basin, East Africa*. AMBIO: A Journal of the
358 Human Environment, 2007. **36**(4): p. 350-358.
- 359 30. Redding, D.W., et al., *Impacts of environmental and socio-economic factors on emergence*
360 *and epidemic potential of Ebola in Africa*. Nat Commun, 2019. **10**(1): p. 4531.
- 361 31. Morley, B., *Causality between economic growth and immigration: An ARDL bounds testing*
362 *approach*. Economics Letters, 2006. **90**(1): p. 72-76.
- 363 32. Ruan, Z., et al., *Epidemic spreading between two coupled subpopulations with inner*
364 *structures*. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2017. **27**(10): p. 103104.
- 365 33. Balcan, D., et al., *Modeling the spatial spread of infectious diseases: The GLObal Epidemic*
366 *and Mobility computational model*. Journal of computational science, 2010. **1**(3): p. 132-145.
- 367 34. Balcan, D., et al., *Multiscale mobility networks and the spatial spreading of infectious*
368 *diseases*. 2009. **106**(51): p. 21484-21489.
- 369 35. Grady, D., *On the Structure of Complex Networks: Universality in Transportation, Borders in*
370 *Human Mobility, and Robust Skeletons*. Dissertations & Theses Gradworks, 2012.
- 371 36. Merler, S. and M. Ajelli, *The role of population heterogeneity and human mobility in the*
372 *spread of pandemic influenza*. Proc Biol Sci, 2010. **277**(1681): p. 557-65.
- 373 37. Colizza, V., et al., *The role of the airline transportation network in the prediction and*
374 *predictability of global epidemics*. Proceedings of the National Academy of Sciences, 2006.
375 **103**(7): p. 2015-2020.
- 376 38. Hufnagel, L., D. Brockmann, and T. Geisel, *Forecast and control of epidemics in a globalized*
377 *world*. Proceedings of the National Academy of Sciences, 2004. **101**(42): p. 15124-15129.
- 378 39. Zhou, F., et al., *Clinical course and risk factors for mortality of adult inpatients with*

379 *COVID-19 in Wuhan, China: a retrospective cohort study*. The Lancet, 2020.

380 40. Batista, M., *Estimation of the final size of the COVID-19 epidemic*. Preprint.] medRxiv, 2020.

381 41. Poirier, C., et al., *The Role of Environmental Factors on Transmission Rates of the COVID-19*

382 *Outbreak: An Initial Assessment in Two Spatial Scales*. Available at SSRN 3552677, 2020.

383 42. Ding Xiang Yuan <https://3g.dxy.cn/newh5/view/pneumonia>.

384 43. Weatherbase. <https://www.weatherbase.com/weather/countryall.php3>.

385 44. China's National Bureau of Statistics. <http://www.stats.gov.cn/>.

386 45. World Health Organization. <https://www.who.int/>.

387 46. Jombart, T., A. Cori, and P. Nouvellet, *earlyR: Estimation of transmissibility in the early*

388 *stages of a disease outbreak*. 2017.

389 47. Zhang, S., et al., *Estimation of the reproductive number of novel coronavirus (COVID-19) and*

390 *the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis*. Int J

391 Infect Dis, 2020. **93**: p. 201-204.

392 48. Wallinga, J. and M. Lipsitch, *How generation intervals shape the relationship between growth*

393 *rates and reproductive numbers*. Proc Biol Sci, 2007. **274**(1609): p. 599-604.

394 49. Oliphant, T.E., *Python for scientific computing*. Computing in Science & Engineering, 2007.

395 **9**(3): p. 10-20.

396 50. Team, R.C., *R language definition*. Vienna, Austria: R foundation for statistical computing,

397 2000.

398 51. Breiman, L., *Random Forests*. Machine Learning, 2001. **45**(1): p. 5-32.

399 52. Chevan, A. and M. Sutherland, *Hierarchical partitioning*. The American Statistician, 1991.

400 **45**(2): p. 90-96.

401 53. *rdaenvpart*. <http://blog.sciencenet.cn/home.php?mod=attachment&id=434918>.

402 54. Legendre, P. and M.J. Anderson, *Distance-based redundancy analysis: testing multispecies*

403 *responses in multifactorial ecological experiments*. Ecological monographs, 1999. **69**(1): p.

404 1-24.

405 55. Rosseel, Y., *Lavaan: An R package for structural equation modeling and more. Version 0.5–12*

406 *(BETA)*. Journal of statistical software, 2012. **48**(2): p. 1-36.

407

408 **Acknowledgements**

409 I appreciate the contributors for collecting COVID-19 case data around the world,
410 G.N. Wang for helping to collect data and Dr. Liang for providing the data of monthly
411 flows in global air travel passengers.

412

413 **Competing interests**

414 The author declare no competing interests.

415

416 **Figure legends**

417 **Figure 1.** The roles of climate and socio-economic variables playing in the early
418 growth rates of COVID-19. The relative importance of variables in early growth rate
419 were estimated by random forest (a), while the individual explanations were
420 calculated by redundancy analysis with the hierarchy algorithm (b). The total of
421 variation explained by this six factors were 54.3% and 32.8% at the province level of
422 China (black bars) and the state level of the other countries (white bars), respectively.
423 In structural equation model (SEM) of early growth rate driven by socio-economic
424 and climate factors at China's province level (c) and state level (d). The blue and red
425 windows represented the latent variables of socio-economic and climate, respectively.
426 The explanations of socio-economic (*S*), climate (*C*) and their interactions ($S \times C$) in
427 early growth rate were estimated by using adjusted R-squared in redundancy analysis
428 ordination. The goodness of fitting in SEM at China's province level was that
429 $Chi-square = 23.50$, $P = 0.13$, $CFI = 0.90$ and $RMSEA = 0.13$, while state level was
430 that $Chi-square = 22.65$, $P = 0.12$, $CFI = 0.94$ and $RMSEA = 0.13$. These variables
431 were shared by the province and state level, including both climate and
432 socio-economic factors. For climate variables, there were monthly average highest
433 temperature (Avg.HT), monthly average precipitation (Avg.Pre). For
434 socio-economic, there were GDP per capita (GDP.pc), GDP, population (Pop),
435 population density (Pop.density) and commuting rate (Commuting). † The fixed
436 indicators of latent variables. *** $P < 0.01$, ** $P < 0.05$.

Figure 2. The effects of quarantine regarding travel restrictions on the explainable variations in the early growth rates of COVID-19 driven by socio-economic (S) and climate (C) factors. **(a)** In metapopulation network model, the proportion of immigrants in subpopulation i from j was c_{ji} constrained by quarantine level q . For each subpopulation, COVID-19 spread was modeled by susceptible-exposed-infected-removed (SEIR) model, where DE and DI were the exposed and infected periods of COVID-19, respectively. R_0 is basic reproductive number, and is shaped by random socio-economic (S) and climate (C) factors with the constants of a and b and random error ε . **(b)** The total variations in early growth rate explained by socio-economic and climate factors with different quarantine level. The quarantine level of zero indicates that there were no travel restrictions among subpopulations, while the quarantine level of one indicate that travel flows among subpopulations are zero. In simulation process, the travel restrictions would be implemented after 20 days of epidemic outbreak and the growth rate were estimated based on accumulative cases during the period of travel restriction. The number of subpopulations in a metapopulation are denoted by the colors of bars.

Figure 1

Figure 2