Are Covid-19 Cases Independent of the City Sizes?

Aydogan, Yigit

Kirklareli University

9 April 2020

Are Covid-19 Cases Independent of the City Sizes?

Abstract

In this paper it is shown that more crowded cities did not exacerbate the Covid-19 pandemic. Reports from Turkey constitutes evidence that the Gibrat's Law holds and the infection grows among population in proportion to the city sizes. Covid-19 cases are lognormally distributed throughout the country. While the 0-19 age group of the society is associated with a negative impact on the reported cases, 40-59 group has the most additive effect. Distribution of the reported deaths from Covid-19 does not grow in proportion to the city size, and may well be approximated by both exponential and normal distributions.

Keywords: Covid-19, Gibrat's Law, Law of the Proportionate Effect, City Size Distribution.
JEL Codes: D01, C21, I18.

1. Introduction

After the spread of the Sars-Cov-2 virus throughout the world in the early 2020, medical research papers have been quickly published. Yet, there is little work on the novel Covid-19 pandemic in the field of economics. This paper is intended to spark a quick start to demographic and economic analysis of Covid-19 using the toolkit of microeconomics. It is important to show that the classical approaches have a lot to say on this contemporary problem.

Distributional properties are important points of view for understanding the underlying behavior of both social and biological systems. Size dependence of growth and properties of size distributions are widely examined relationships in variety of disciplines, i.e. photosynthesis, atomic particles, cities and firms being among the most popular. Yet, it is somehow overlooked for pandemics.

The Law of the Proportionate Effect, i.e. Gibrat's Law, states that unit (firms) growth is proportionate to (independent of) its size (Gibrat, 1931). There are numerous ways to take the picture of an infected group or humans. In this quest, the toolkit of microeconomic analysis might be useful to understand the novel coronavirus Covid-19 pandemic and help developing policy responses.

The literature on the city size aspect of pandemics is of limited supply. Eggo, Cauchemez and Ferguson (2010) investigated the spatial properties of 1918 pandemic influenza for England, Wales and the US. They provide a detailed and complex analysis including power law relationship between infectivity and mortality, transmission trees and connectivity models. It was shown that spatial structures of the countries are important, and the susceptibility of cities increases slowly with their population. They conclude as stating that analysis under contemporary human mobility rates might differ. Wood et al. (2007) assert that infection speed increases (unaffected) if originating city is smaller (larger) for the 2003 SARS pandemic. It was also shown for 1918 pandemic influenza that reproductive number R is not correlated with city size (Davis and Lappin, 1923).

In this paper, the fruitful analytic power of distributional analysis of microeconomics literature is utilized. In the second section, the Covid-19 infection data of Turkey by April 3, 2020 is presented. Next, regressions using city sizes and age categorizations are to be tabulated. In the fourth section, distributional plots of Covid-19 in Turkey are depicted. Last section provides conclusions and policy discussion.

2. Data

Covid-19 was originated in Wuhan, China during late December 2019, and classified as a pandemic on March 10, 2020 (World Health Organization, n.d.). Turkey had experienced a moderately later infection outbreak than other highly populated nations. The first case in Turkey was announced on March 10, 2020. By the time this work has been done, the only official infection data for cities has been revealed by the Ministry of Health on April 1 and 3, a total of two datasets for cases, and one for deaths from Covid-19 on April 3, 2020. The robustness of these data sets is a major concern for econometric analysis. The questions about the data is to be addressed during the next sections, accordingly.

Table 1. Covid-19 Data in Turkey by April 1 and 3, 2020

#	City	Population	April 1	April 3	Growth	Death	Case per 100k	Death per 100k
1	Istanbul	15,519,267	8,852	12,231	38%	210	78.81	1.35
2	Izmir	4,367,251	853	1,105	30%	27	25.30	0.62
3	Ankara	5,639,076	712	860	21%	11	15.25	0.20
4	Konya	2,232,374	584	601	3%	11	26.92	0.49
5	Kocaeli	1,953,035	410	500	22%	14	25.60	0.72
6	Sakarya	1,029,650	207	337	63%	4	32.73	0.39
7	Isparta	444,914	268	289	8%	2	64.96	0.45
8	Bursa	3,056,120	135	259	92%	8	8.47	0.26
9	Adana	2,237,940	197	241	22%	4	10.77	0.18
10	Zonguldak	596,053	112	197	76%	12	33.05	2.01
11	Samsun	1,348,542	112	167	49%	4	12.38	0.30
12	Kayseri	1,407,409	109	130	19%	4	9.24	0.28
12	**Subtotal**	**39,831,631**	**12,551**	**16,917**	**37%**	**311**	**28.62**	**0.60**
81	**Total**	**83,154,997**	**14,681**	**19,576**	**33%**	**414**	**23.54**	**0.50**

Table 1 presents the diagnosis data for top 12 cities (from a total of 81), sorted by the first reporting. The decision on the list of 12, rather than e.g. 10, is based on the fact that the list remains full with the second reporting of the cases with a few changes of rank. On the other hand, most of the cases and deaths have found in those cities. One can easily notice that the cities are not sorted by their populations when it comes to the infected individuals. Growth of the spread from April 1 to April 3, 2020 also varies significantly among cities. Data on observed deaths per 100,000 people reveals that with 2.01 the city of Zonguldak has the highest death per inhabitants. Isparta, alongside with Zonguldak, has more cases than some of the largest cities in Turkey by April 1.

Table 2. Age Distribution of the Population in Turkish Cities

#	City	0-19	20-39	40-59	0-29	30-59	60+
1	Istanbul	29%	34%	26%	45%	44%	11%
2	Izmir	25%	31%	28%	39%	44%	17%
3	Ankara	28%	32%	27%	44%	43%	13%
4	Konya	32%	30%	24%	48%	38%	14%
5	Kocaeli	31%	33%	25%	46%	43%	11%
6	Sakarya	29%	31%	26%	45%	41%	14%
7	Isparta	26%	30%	25%	43%	39%	18%
8	Bursa	28%	31%	27%	43%	43%	14%
9	Adana	33%	30%	25%	47%	40%	13%
10	Zonguldak	24%	28%	29%	38%	43%	19%
11	Samsun	28%	29%	27%	42%	41%	17%
12	Kayseri	32%	30%	24%	47%	40%	13%
12	**Subtotal**	**29%**	**31%**	**26%**	**44%**	**42%**	**14%**
81	**Total**	**31%**	**30%**	**24%**	**47%**	**38%**	**15%**

Research on Covid-19 shows that the demography of infected groups maintains important information about infection spreading and mortality rates (Dowd et al., 2020). Table 2 presents age groups' share in top 12 city populations, categorized according to clinical studies, e.g. Wu et al. (2020). It can be seen that Turkey has a relatively young population, and the chosen subgroup reflects a close picture of the country. Above cited medical papers suggest, death rate is expected to be higher among the age group of 60+. Since the data does not cover the average age of infected individuals, high mortality cities of Zonguldak and Istanbul does not fit to the picture. Zonguldak has slightly higher elderly population, yet the deviation is not as striking as its death rate. While being out of scope of this work, one explanation of Zonguldak being exceptionally high death rate could be the effect of concentrated coal mining activities and presence of 7 coal-powered thermal power stations on the respiratory system of the locals which is targeted by Covid-19 (Shi et al., 2020).

3. Does the City Size Matter?

In order to understand a contemporary pandemic, it is important to know its *emergent* properties. One can build up a complex network system to prove the transmission path of a virus among interrelated humans. There is a much simple way of knowing if the present large clusters of residential areas make the society more vulnerable against easily human-to-human transmitted diseases.

The Law of The Proportionate Effect for the spread of Covid-19 can be formally stated as:

$$\log(cases_i) = \alpha_1 + \beta_1 \log(pop_i) + \epsilon_1 \qquad (1)$$

where *cases* and *pop* are Covid-19 cases observed in a city and its population, respectively. If $\beta = 1$ then one can say that the distribution of Covid-19 cases among Turkish cities is independent of the initial size of the city. Alternatively, a *robustness-challenge* equation can be stated as:

$$\log(cases_i) = \alpha_2 + \beta_3 \log(pop_i) + \beta_4 (ratio_i) + \epsilon_3 \qquad (2)$$

where *ratio* denotes shares of age groups among the population, namely 0-19, 20-39, 40-59, 60+, 0-29 and 30-59. As an extension, *death* statistics are also used in the regression for a more complete analysis.

$$\log(deaths_i) = \gamma + \beta_2 \log(pop_i) + \epsilon_2 \qquad (3)$$

where *deaths* denotes deceased Covid-19 patients in Turkey by April 3,2020.

Table 3. Estimation Results of Equations (1)-(3)

	(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**
Population	1.03** (1.05**) [0.72**]	1.10** (1.11**) [0.80**]	1.03** (1.04**) [0.69]	0.99** (1.01**) [0.70**]	1.14** (1.17**) [0.86**]	1.05** (1.07**) [0.76**]	0.93** (0.95**) [0.63**]
% of 0-19		-5.55** (-5.60**) [-5.70**]					
% of 20-39			-4.50** (-6.03) [1.42]				
% of 40-59				8.99** (9.28**) [8.72**]			
% of 60+					7.35** (7.62**) [6.98*]		
% of 0-29						-4.32** (-4.44**) [-4.33**]	
% of 30-59							8.38** (8.56**) [8.32**]
R^2	0.45 (0.51) [0.41]	0.53 (0.59) [0.52]	0.46 (0.52) [0.41]	0.52 (0.58) [0.50]	0.51 (0.57) [0.48]	0.53 (0.58) [0.50]	0.52 (0.58) [0.50]
y = Cases by April 1, 2020, (Cases by April 3, 2020), [Deaths by April 3, 2020]							

***95%, **99%**

Table 3 presents estimation results of Equations (1), (2) and (3). As noted, results from the top line are from the first dataset of April 1, 2020. Parentheses on the second line denote the data reported on April 3, 2020 and brackets denote death statistics which were also reported on the same date. Percentages on the left side refer to the age groups among the population of the cities. R^2 is the coefficient of determination.

Results from panel (a) states that Covid-19 cases observed in Turkey are independent of the city sizes, with a coefficient close to "1". When this statement is challenged with demographics, and a second data set it can be seen that the Law still holds, except for panel (b) and (e). Both panels incorporate significant groups related to the novel Covid-19 pandemic. 0-19 age group is reported to show almost no symptoms and 60+ age group has been reported as the highest vulnerability against the virus (Wu et al., 2020). Hence, the expected intrinsic value of these variables is expected to be relatively high and they affect the present query when addressed separately without adequate theoretical and functional background. On the other hand, 20-39 age group returns statistically insignificant coefficients, as expected building upon the previous assessment. Moreover, the share of people 30+ years old in a city seems to be positively related to the Covid-19 cases, while younger portions of the society are on the other side.

Equation (3) also provides interesting results about deaths from Covid-19 in Turkey. According to the Table 3, with a coefficient smaller than 1, as the city size grows the death toll does not increase proportionally. This fact lays constant among all the panels in Table (3).

4. Distributional Properties

In order to support such regularities as presented in the previous section, plotting the distribution and growth of Covid-19 cases should be informative. If the Gibrat's Law holds, spreading of the Covid-19 among Turkish cities follows a random multiplicative process, and expected to be distributed as lognormal.

Figure 1. Distribution of Infected and Dead Individuals

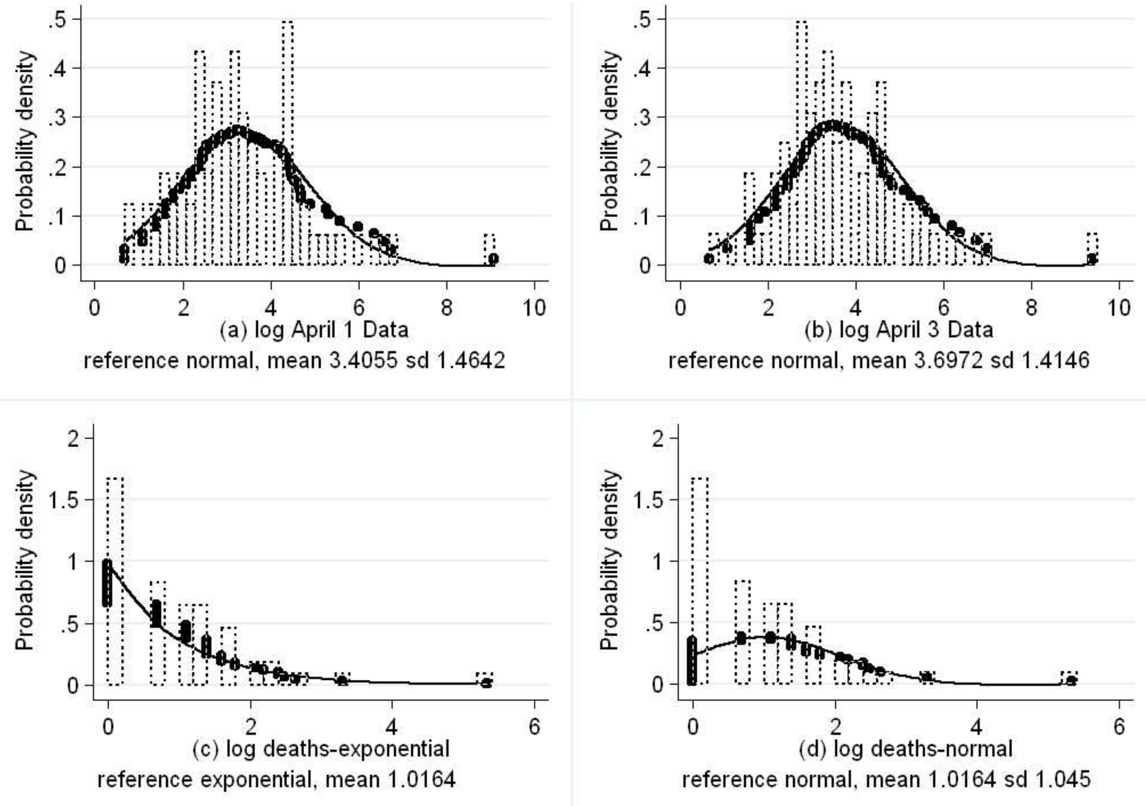
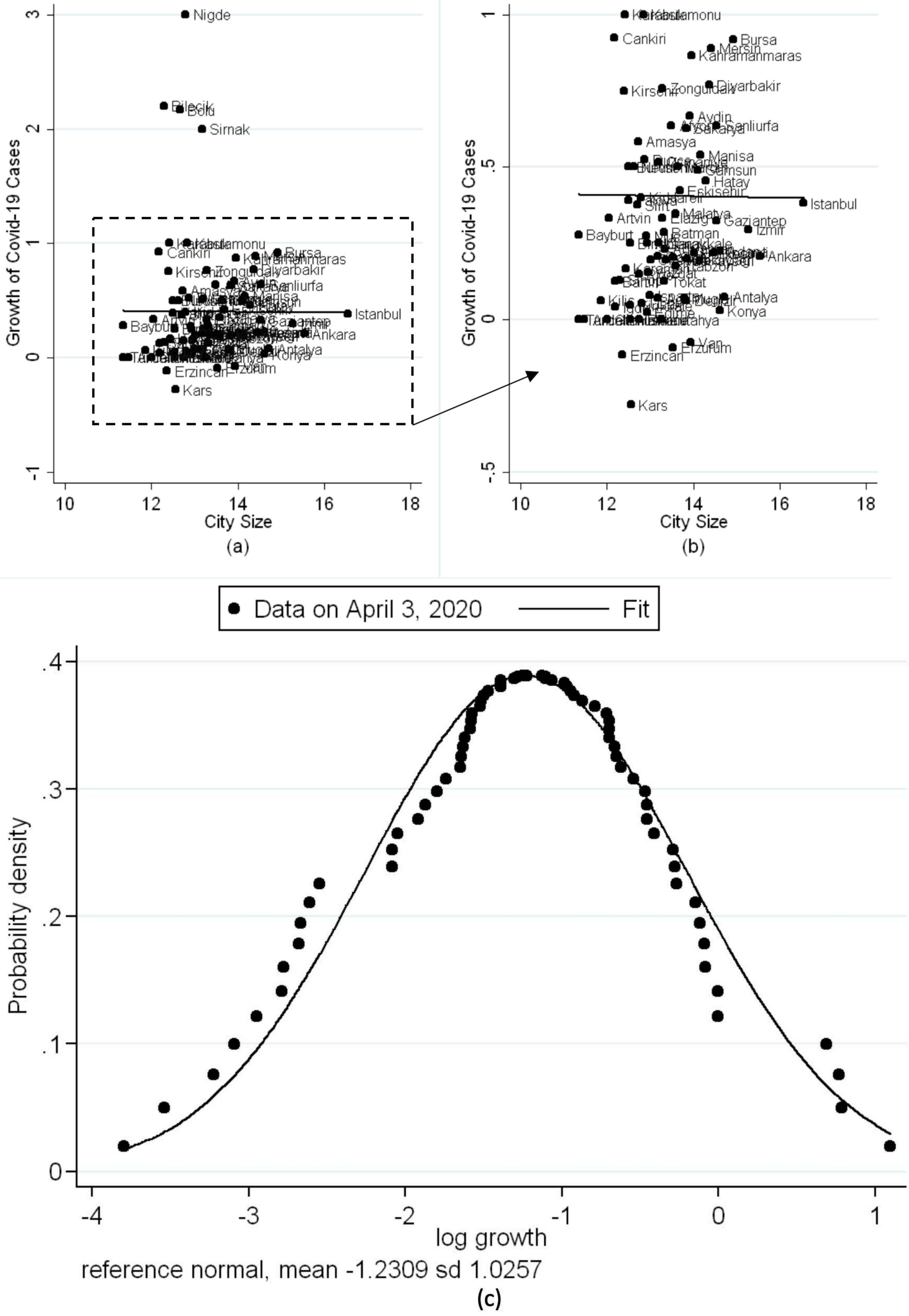


Figure 1 depicts the distributions of infected individuals' data by April 1 and April 3, along with the distribution of death of Covid-19 patients by Turkish cities. Panels (a) and (b) clearly indicates that the distribution of Covid-19 cases can well be approximated with the lognormal. In terms of the logarithm of the death data, both exponential and normal distributions seem plausible.

Figure 2. Growth Plot of Covid-19 Cases by April 3, 2020



As a long-term regularity implied by the Gibrat's Law, the growth rates are also expected to be distributed lognormally (Stanley et al., 1996). Growth rate of the Covid-19 cases are plotted in Figure 2. Panel (a) depicts the growth rate of the data. Panel (b) is only a closer look into the boxed area in Panel (a). The horizontal regression line and the dispersed nature of the data are interesting. Panel (c) shows that logarithm of the growth rate of Covid-19 cases in Turkey from April 1 to April 3, 2020 is well approximated by a lognormal distribution. This further strengthens the robustness of the data and previous estimations on the Law of the Proportionate Effect.

5. Conclusions

Gibrat's Law, i.e. the Law of the Proportionate Effect, has simple but strong assumptions. If the regularity holds, it yields important implications about the underlying mechanisms of complex phenomena. In this study, it was shown that the Covid-19 cases reported in Turkey grows in proportion to the city sizes. Logarithm of the Covid-19 cases in Turkish cities follows a random walk. This fact might be helpful for policymaking on whether to plan for smaller cities to protect citizens from fast-track deadly infections, during the post-Covid-19 era. It could favor the economic benefits of contemporary large settlements against the claims that crowded cities would catalyze the transmission of deadly pandemics.

Despite the reservations, the data is shown to be robust. The weight of the younger population is negatively related with the Covid-19 cases in Turkey. On the other hand, it seems plausible to isolate 30-49 and 40-59 age groups in order to take the expansion under control.

Further studies after the pandemic ends may improve the model by incorporating demographics of infected persons from a more complete data set. Another study might be done to determine whether the distribution of deaths from Covid-19 could be approximated by the exponential distribution.

References

[1] Davis, W. H., & Lappin, R. C. (1923). *Mortality Rates 1910-1920 with Population of the Federal Censuses of 1910 and 1920 and Intercensal Estimates of Population*. US Government Printing Office.

[2] Dowd, J. B., Rotondi, V., Adriano, L., Brazel, D. M., Block, P., Ding, X., ... & Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *medRxiv*. https://doi.org/10.1101/2020.03.15.20036293.

[3] Eggo, R. M., Cauchemez, S., & Ferguson, N. M. (2011). Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. *Journal of the Royal Society Interface*, *8*(55), 233-243. http://doi.org/10.1098/rsif.2010.0216

[4] Gibrat, R. (1931). Les Inegalites Economiques. Paris: Libraire du Recueil Sirey.

[5] Shi, H., Han, X., Jiang, N., Cao, Y., Alwalid, O., Gu, J., ... & Zheng, C. (2020). Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet Infectious Diseases*.

[6] Stanley, M. H., Amaral, L. A., Buldyrev, S. V., Havlin, S., Leschhorn, H., Maass, P., ... & Stanley, H. E. (1996). Scaling behaviour in the growth of companies. *Nature*, *379*(6568), 804-806. https://doi.org/10.1038/379804a0

[7] Wood, J. G., Zamani, N., MacIntyre, C. R., & Becker, N. G. (2007). Effects of internal border control on spread of pandemic influenza. *Emerging infectious diseases*, *13*(7), 1038. https://doi.org/10.3201/eid1307.060740

[8] World Health Organization. (n.d.). Retrieved April 8, 2020, from WHO website: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen

[9] Wu, J.T., Leung, K., Bushman, M. *et al.* Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* (2020). https://doi.org/10.1038/s41591-020-0822-7