

A MICROECONOMIC ANALYSIS OF THE COVID-19 DISTRIBUTION IN TURKEY

Türkiye’de COVID-19 Dağılımının Mikroiktisadi Araçlarla Analizi

Yiğit AYDOĞAN*

Geliş: 28.06.2020 / *Kabul:* 01.12.2020

DOI: 10.33399/biibfad.759410

Abstract

Larger cities do not amplify the COVID-19 pandemic in Turkey. Reports from Turkish cities provide evidence that the Gibrat’s Law holds and the infection grows among population in proportion to the city sizes. Growth of the pandemic is not faster in larger cities. 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: R11; D39; C21; I19.

Öz

Daha büyük olan şehirler Türkiye’de COVID-19 pandemisinin etkisini artırmamaktadır. Türkiye’deki şehirler üzerine yapılan çalışmaya göre Gibrat Yasası geçerlidir ve salgın bireyler arasında şehrin büyüklüğüyle orantılı olarak yayılmaktadır. Pandeminin yayılma hızı şehir büyüklüğüyle birlikte artmamaktadır. COVID-19 vakaları ülkede log-normal dağılım göstermektedir. Şehirlerdeki 0-19 yaş aralığındakilerin nüfusa oranı raporlanan vaka sayıları üzerine negatif etkiye sahiptir, 40-59 yaş grubu en fazla pozitif etkiye sahiptir. COVID-19 kaynaklı

*Arş. Gör. Dr., Kırklareli Üniversitesi, İktisat Bölümü, yigitaydogan@gmail.com
ORCID: <https://orcid.org/0000-0002-1823-0352>.

ölümlerin dağılımı da şehir büyüklüğüyle orantılı olup üstel ve normal dağılımlarla temsil edilebilmektedir.

Anahtar Kelimeler: COVID-19, Gibrat yasası, orantılı etki yasası, şehir büyüklük dağılımı.

JEL Kodu: R11; D39; C21; I19.

1. Introduction

The novel coronavirus pandemic (COVID-19) is the most acute issue to be tackled by all scientists in all related disciplines. The world has suspended most of the earthly activities involving contact with another human being. Therefore, the situation at hand surpasses medical solutions and demands also the understanding of statistical approaches. There has been a large number of preprints/publications made available already, regarding the modeling of the spread of COVID-19. This paper aims to contribute to the demographic and statistical analysis of COVID-19, investigating the scaling laws which may relate to the observed cases and deaths in Turkey.

Since the pioneering work of Bernoulli (1760) mathematical modeling of epidemics has evolved around two main pillars. The first strand in the literature is about temporal investigations of certain perimeters and the second is spatial modeling the spread of infections. COVID-19 is not an exemption (e.g. Barlow and Weinstein, 2020; Lu et al., 2020; Peng et al., 2020 for the former, and Zhang et al., 2020; Blasius, 2020 for the latter). Moreover, there are spatiotemporal models which combine both approaches (Bekiros and Kouloumpou, 2020). A thorough review of recent literature on COVID-19 can be found in Park et al., (2020).

Understanding the behavior of social and biological systems, such as infected individuals and spread of a virus, involves specifying the scaling laws. 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. The pre-COVID-19 literature on the city size aspect of pandemics is of limited supply. Eggo et al., (2011) 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 Lippin, 1923).

Stylized facts and scaling laws have been important tools in economic theory starting as early as Pareto (1897). The idea has been developed slowly in the field of economics and related sciences, e.g. Gibrat (1931), Zipf (1949) and Simon and Bonini (1958). In the last few decades, scaling has also been increasingly popular among researchers from biology (see Prothero (1986) and West, Brown and Enquist (2000) for a literature survey). Early literature on the scaling laws related to COVID-19 statistics yields mixed results. Preliminary findings of Heroy (2020) indicates that US county-level data does not show relation between city sizes and exponential growth rate of COVID-19, but the exponential spread may start earlier in larger cities. Blasius (2020) reports a truncated power-law distribution of COVID-19 cases and deaths in both global and US counties, assuming an exponential growth of the variables. On the other hand, power-law relationship between growth rate of COVID-19 and US city sizes have been reported in Stier et al., (2020), indicating faster spread of the virus in larger cities.

This paper investigates the daily infection and death statistics of all cities in Turkey with a microeconomic perspective. The primary aim of the analysis is to determine whether larger cities increase the spread of the infection or not. It would be important to determine if the contemporary lifestyle of the humankind in complex urban networks posit extra vulnerability against pandemics. The importance of the research question mandates that if such agglomerations of people were

to increase the effects of deadly infections, future of the urbanization might be re-designed with caution.

In this study, COVID-19 cases and death statistics are examined using publicly available, but hard to collect and combine, city-level pandemic data of Turkey. In the next section, the COVID-19 data of Turkey is presented. In this regard, it also serves to the scientific community by the presenting a genuine dataset. In the third section ordinary least squares regressions using city sizes and age categorizations are to be tabulated. In the fourth section, distributional plots of COVID-19 in Turkey are to be 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 (WHO, n.d.). Turkey had experienced a moderately later infection outbreak than other highly populated nations. The first infection case in Turkey was announced on March 10, 2020. By the time this work has been done, the only data for Turkish cities is from April 1 and 3, 2020, a total of two datasets for cases, and one dataset for COVID-19 related deaths from April 3, 2020. The case data on April 1, 2020 is the only one reported by the Turkish Ministry of Health (AA, n.d.) for convenience. Case and death numbers on April 3, 2020 is not conveniently reported but collected exclusively for this research, from several news agencies and the Minister's daily press-conferences. 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: Summary of COVID-19 Data in Turkey by April 1 and April 3, 2020 and Population Statistics, Reporting Top 12 Cases of 81 Cities in Total.

#	City	Population	April 1 Cases	April 3 Cases	Case Growth	April 3 Deaths	Case per 100k Residents	Death per 100k Residents
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

Source: AA (n.d.) and Author's Collection

The data of this study covers all 81 cities in Turkey. To keep tabulations manageable, Table 1 presents the diagnosis data for top 12 cities (from a total of 81), sorted by the first available date to emphasize the heterogeneity of growth. The decision on the list of 12, rather than conventional numbers such as top-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 been found in those cities. One can easily notice that the cities are not sorted by their populations when it comes to the infected count. Growth of the spread from April 1 to April 3, 2020 varies significantly among cities. Data on observed deaths per 100,000 inhabitants 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%

Source: Turkish Statistical Institute (2020)

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 scaling properties. 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).

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

$$\log(\text{cases}_i) = \alpha_1 + \beta_1 \log(\text{pop}_i) + \epsilon_1 \quad (1)$$

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:

$$\log(\text{cases}_i) = \alpha_2 + \beta_3 \log(\text{pop}_i) + \beta_4(\text{ratio}_i) + \epsilon_3 \quad (2)$$

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.

$$\log(\text{deaths}_i) = \gamma + \beta_2 \log(\text{pop}_i) + \epsilon_2 \quad (3)$$

where *deaths* denote deceased COVID-19 patients in Turkey by April 3,2020.

Table 3: Estimation Results of Equations (1)-(3). Explanatory Variables Change from Panels (a) to (g), and Various Brackets Differentiate Between the Dependent Variables (denoted in the last row).

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
<i>Population</i>	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**]
<i>% of 0-19</i>		-5.55** (-5.60**) [-5.70**]					
<i>% of 20-39</i>			-4.50** (-6.03) [1.42]				
<i>% of 40-59</i>				8.99** (9.28**) [8.72**]			
<i>% of 60+</i>					7.35** (7.62**) [6.98*]		
<i>% of 0-29</i>						-4.32** (-4.44**) [-4.33**]	
<i>% of 30-59</i>							8.38** (8.56**) [8.32**]
R²	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]
<i>Dependent variables are denoted as follows : Cases by April 1, 2020, (Cases by April 3, 2020), [Deaths by April 3, 2020]</i>							

Note: *95%, **99% statistical significance

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. The probability density function of the lognormal distribution can be stated as follows:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} \exp\left\{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right\} \quad (4)$$

where μ is the mean and σ is the standard deviation of observed variable x .

Exponential distribution has been a primary consideration in the literature of the epidemic growth and COVID-19 is no exception to that (see Lee et al., 2020 and references therein). Therefore, one can test that whether pandemic data in Turkey is better approximated with exponential distribution rather than lognormal. The probability density function of the exponential distribution is:

$$f(x) = e^{-x} \quad (5)$$

Figure 1: PDFs of Cases by April 1 and 3, 2020, and Deaths by April 3, 2020 in Turkish Cities

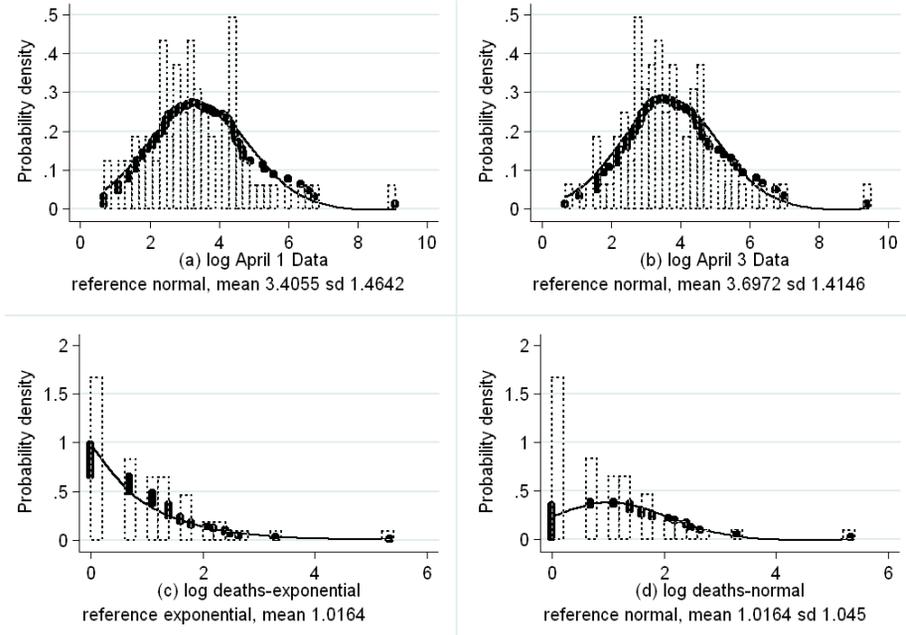
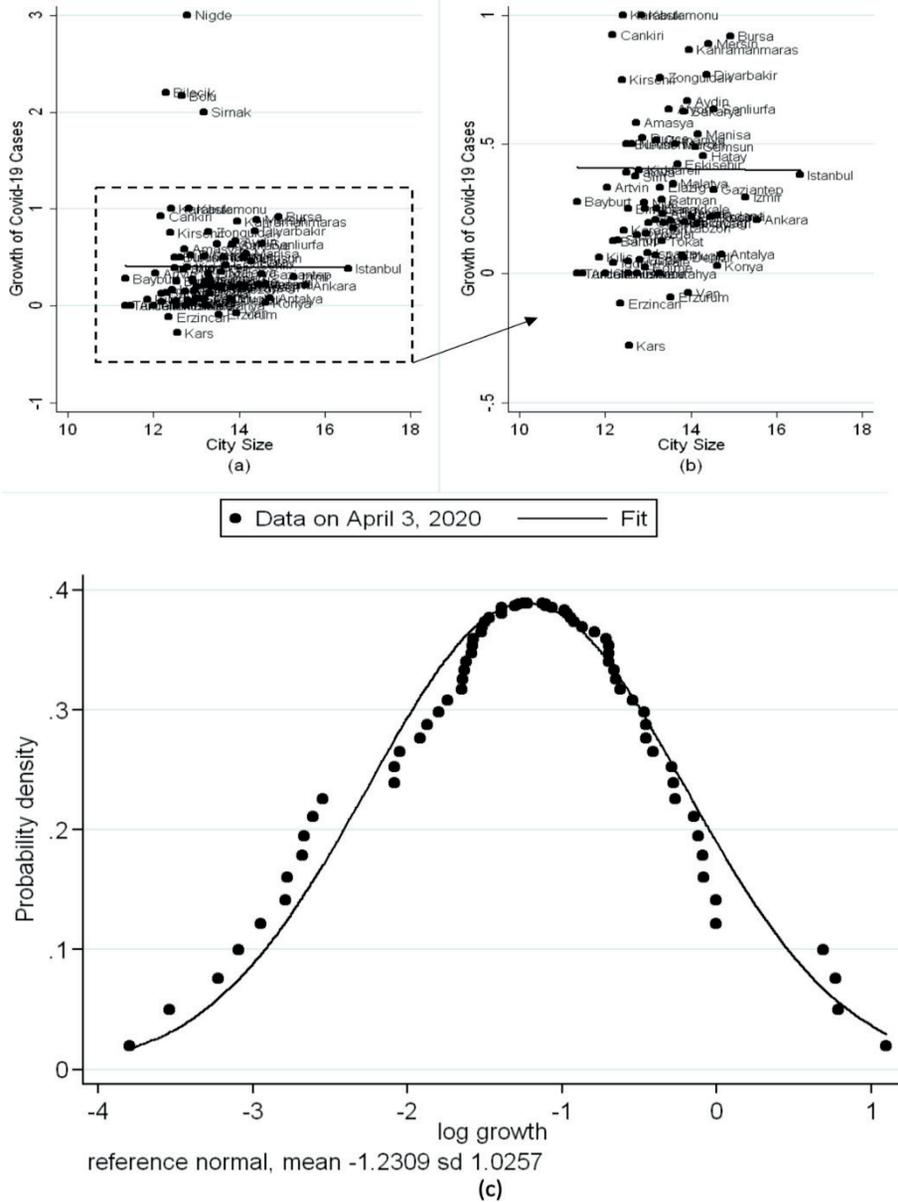


Figure 1 depicts the distributions of infected individuals' data on April 1 and April 3, along with the distribution of death of COVID-19 patients by the Turkish cities. Reference distribution lines have been provided and estimated distributional parameters have been noted below each panel. Panels (a) and (b) clearly indicates that the distribution of COVID-19 cases can be approximated with the

lognormal. In terms of the logarithm of the death data, both exponential and normal distributions seem good fits.

Figure 2: Growth Plotting of COVID-19 Cases of Turkish Cities from April 1 to April 3, 2020. Panel (b) is a closer look at the selected area of Panel (a), and Panel (c) is the PDF.



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) to improve visualization. The horizontal regression line and the dispersed nature of the data are important features to be expected from the findings of the previous section. Panel (c) shows that 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

- Anatolian Agency. Retrieved on 04-08-2020.
<https://www.aa.com.tr/tr/koronavirus/turkiyenin-il-il-kovid-19-vaka-haritasi/1788776>
- Barlow, N. S., & Weinstein, S. J. (2020). Accurate closed-form solution of the SIR epidemic model. *Physica D: Nonlinear Phenomena*, 132540.
- Bekiros, S., & Kouloumpou, D. (2020). SBDiEM: A new mathematical model of infectious disease dynamics. *Chaos, Solitons & Fractals*, 109828.
- Bernoulli, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Histoire de l'Acad., Roy. Sci.(Paris) avec Mem*, 1-45.
- Blasius, B. (2020). Power-law distribution in the number of confirmed COVID-19 cases. *arXiv preprint arXiv:2004.00940*.
- 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.
- 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>.
- 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>
- Gibrat, R. (1931). *Les Inegalites Economiques*. Paris: Libraire du Recueil Sirey.
- Heroy, S. (2020). Metropolitan-scale COVID-19 outbreaks: How similar are they? *arXiv preprint arXiv:2004.01248*.

- Lee, K. B., Han, S., & Jeong, Y. (2020). COVID-19, flattening the curve, and Benford's law. *Physica A: Statistical Mechanics and its Applications*, 559, 125090.
- Lu, Z., Yu, Y., Chen, Y., Ren, G., Xu, C., Wang, S., & Yin, Z. (2020). A fractional-order SEIHDR model for COVID-19 with inter-city networked coupling effects. *arXiv preprint arXiv:2004.12308*.
- Pareto, V. (1897). *Cours d'Economie Politique*, vol. 2. Paris.
- Park, M., Cook, A. R., Lim, J. T., Sun, Y., & Dickens, B. L. (2020). A systematic review of COVID-19 epidemiology based on current evidence. *Journal of Clinical Medicine*, 9(4), 967.
- Peng, L., Yang, W., Zhang, D., Zhuge, C., & Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. *arXiv preprint arXiv:2002.06563*.
- Prothero, J. (1986). Methodological aspects of scaling in biology. *Journal of Theoretical Biology*, 118(3), 259-286.
- 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*.
- Simon, H. A., & Bonini, C. P. (1958). The size distribution of business firms. *The American Economic Review*, 607-617.
- 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>
- Stier, A., Berman, M., & Bettencourt, L. (2020). COVID-19 attack rate increases with city size. *Mansueto Institute for Urban Innovation Research Paper Forthcoming*.
- Turkish Statistical Institute. (2020). Address Based Population Registration System [Data set]. <https://biruni.tuik.gov.tr/medas/?kn=95&locale=tr>

- West, G. B., Brown, J. H., & Enquist, B. J. (2000). Scaling in biology: patterns and processes, causes and consequences. *Scaling in Biology*, 87-112.
- 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>
- World Health Organization. Retrieved on 04-08-2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
- Wu, J.T., Leung, K., Bushman, M. et al. (2020). 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>
- Zhang, Y., Tian, H., Zhang, Y., & Chen, Y. (2020). Is the epidemic spread related to GDP? Visualizing the distribution of COVID-19 in Chinese Mainland. *arXiv preprint arXiv:2004.04387*.
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Cambridge: Addison-Wesley Press

