



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original Research

Human mobility and coronavirus disease 2019 (COVID-19): a negative binomial regression analysis

L.I. Oztig^{a,*}, O.E. Askin^b^a Department of Political Science and International Relations, Faculty of Economics and Administrative Sciences, Yildiz Technical University, Istanbul, Turkey^b Department of Statistics, Yildiz Technical University, Davutpasa, Esenler, 34220, Istanbul, Turkey

ARTICLE INFO

Article history:

Received 29 April 2020

Received in revised form

2 July 2020

Accepted 2 July 2020

Available online 10 July 2020

Keywords:

Human mobility

COVID-19

Airports

Schengen countries

ABSTRACT

Objectives: This study aimed to examine the link between human mobility and the number of coronavirus disease 2019 (COVID-19)–infected people in countries.

Study design: Our data set covers 144 countries for which complete data are available. To analyze the link between human mobility and COVID-19–infected people, our study focused on the volume of air travel, the number of airports, and the Schengen system.

Methods: To analyze the variation in COVID-19–infected people in countries, we used negative binomial regression analysis.

Results: Our findings suggest a positive relationship between higher volume of airline passenger traffic carried in a country and higher numbers of patients with COVID-19. We further found that countries which have a higher number of airports are associated with higher number of COVID-19 cases. Schengen countries, countries which have higher population density, and higher percentage of elderly population are also found to be more likely to have more COVID-19 cases than other countries.

Conclusions: The article brings a novel insight into the COVID-19 pandemic from a human mobility perspective. Future research should assess the impacts of the scale of sea/bus/car travel on the epidemic. The findings of this article are relevant for public health authorities, community and health service providers, as well as policy-makers.

© 2020 The Royal Society for Public Health. Published by Elsevier Ltd. All rights reserved.

Introduction

The globalized world, in which the scale of the movement of people is at unprecedented scale¹ is susceptible to the spread of diseases on a global scale. With sophisticated transport networks that have increased reach, the speed of travel, and the volume of passengers, ‘pathogens and their vectors can now move further faster and in greater numbers than ever before’.² The global spread of coronavirus disease 2019 (COVID-19) that has led to the infection, and deaths, of thousands of people at a rapid scale, is indicative of how infectious diseases can become a global health problem that have the ability to reach more people, and at a faster rate, in an increasingly globalized world.

Human mobility and the spread of diseases

Throughout history, in addition to human migration, trade caravans, religious pilgrimages, and military maneuvers played a central role in the spread of diseases.³ During the Middle Ages, trade routes between Europe and Asia were instrumental in the spread of the plague into Europe.⁴ In the 1500s, the population of the New World suffered from infectious diseases brought by European explorers. The second voyage of Christopher Columbus to the Caribbean in 1493 brought smallpox to the region. In the 1518 smallpox epidemic, thousands of indigenous inhabitants of the Caribbean region died.⁵ In the 16th and 17th centuries, ships from Africa not only brought slaves but also smallpox and fever-carrying mosquitoes to the New World.⁴ Overall, the pathogens carried by migrants had devastating consequences for native Americans who had no immunity for them.

The confluence of American troops with European and African troops in France, and the development of new virus strains, created a permissive environment for the 1918 influenza pandemic that resulted in the deaths of approximately 40 million people in a year.⁶

* Corresponding author.

E-mail addresses: loztig@yildiz.edu.tr (L.I. Oztig), oykumesra@gmail.com (O.E. Askin).

The 1957 pandemic that erupted in China spread to the world within six months.² In 1972, a smallpox epidemic erupted in the autonomous province of Kosovo of (the then named) Yugoslavia on April 11th 1972. Epidemiologic and serologic investigations revealed that smallpox was imported to Yugoslavia from a hajj pilgrim who had visited Mecca and returned to the country by bus via Iraq (where smallpox cases were reported at the time). As a result of the smallpox outbreak, 175 people were infected, among which 35 people died.⁷

Global travel, given the unprecedented volume, speed, and reach, is an important factor in the rapid spread of current diseases.⁸ The study by Olsen et al.⁹ indicates that many severe acute respiratory syndrome (SARS)-infected people traveled on commercial aircraft. The study further revealed that after one flight carrying 120 people (among which one person was symptomatic), SARS developed in 16 people. Illness in passengers was related to the physical proximity to the symptomatic person.

Human mobility and COVID-19

While SARS spread to 37 countries (8000 cases) and Middle East respiratory syndrome (MERS) spread to 27 countries (2494 cases),¹⁰ COVID-19 has spread to more than 200 countries and infected more than a million people in the world, initiating an unprecedented global crisis. Wuhan, the epicenter of the pandemic, is central China's major air and train transportation hub. As of 2019, in Wuhan, international outbound air travel constituted 13.5% of all outbound air travel, whereas the top 40 domestic outbound air routes constituted 81.3%. High air and train traffic across China due to the lunar new year Spring Festival, that started on January 10th 2020, appeared to have played a facilitating role in the spread of COVID-19 throughout the country and abroad.¹⁰

The first COVID-19 case outside China (a traveler from Wuhan) was reported to the WHO by the Thai government on January 13th 2020.¹¹ Three days later, the Japanese government informed the WHO of its first confirmed infection in a traveler from Wuhan.¹² Strikingly, owing to China's lockdown of the coronavirus-hit Hubei province on January 23rd, many people left Wuhan, which has resulted in the spreading of the diseases in and outside China.¹³ Soon afterward, India, Philippines, Russia, Spain, Sweden, and the UK confirmed their first cases.

Based on the findings of the previous literature and the current trends in the spread of COVID-19, we hypothesize that in countries in which there is a high mobility of people, the number of COVID-19–infected people are correspondingly higher. We also hypothesize that there is a positive association between high numbers of airports in a country and high numbers of COVID-19–infected patients, and that Schengen countries are more likely to have higher numbers of COVID-19–infected patients than non-Schengen countries.

Methods

The dependent variable

The dependent variable of this study is the number of COVID-19–infected people. The data on COVID-19 cases are extracted from the official site of WHO published as of April 3rd 2020.¹⁴ We analyzed 144 countries for which the complete data on independent and control variables are available.

It should be noted that our dependent variable consists of cases that are reported to the WHO. Depending on the late development of/lack of testing equipment and the numbers of tests administered to individuals, the actual number of COVID-19–infected people in countries might be much higher. Lack of adequate testing, or some

cases, any testing, in many countries might be affecting the availability and accuracy of data. For instance, the full impact of COVID-19 on India (the world's second most populous nation), Indonesia (the world's fourth most populous nation), African nations, and various smaller countries remains unknown. This constitutes a limitation to our study. As COVID-19 is reported to have emerged in China and then spread to other countries, we do not include China in our analysis.

Independent variables

We operationalize human mobility by looking at the number of airline passengers carried into the countries. The data are extracted from World Development Indicators.¹⁵ Airline passengers include both domestic and international aircraft passengers of air carriers registered in the country. We note that the most recent data on the airline passengers is from 2018. Although the data does not correspond to actual human mobility as of 2020, we assume that the pattern of air travel is unchanged until the start of the pandemic. We measure airport numbers and the Schengen system as factors that facilitate human mobility. The data on airport numbers are extracted from the World Factbook of the Central Intelligence Agency.¹⁶ We code Schengen countries as 1 and 0 otherwise.

Control variables

We control for population density in our analysis. In countries with high population density, people have contact with large numbers of people which facilitate person-to-person spread of many infectious diseases.¹⁷ Furthermore, the elderly people are more susceptible to infections 'because of waning cell-mediated immunity and impaired host defenses but also because of chronic diseases and use of drugs and treatments that may be immunosuppressive'.¹⁷ By bearing in mind the fact that there is no scientifically established relationship between immunity and the risks of contracting COVID-19 disease, we control for the percentage of elderly people in population our analysis. The data on population density and the percentage of elderly people (65 years and older) are extracted from World Development Indicators.¹⁸

Model specification

To analyze the variation in COVID-19–infected people across countries, we use negative binomial regression (NBR) model. NBR is based on the Poisson-gamma mixture distribution. It is useful for predicting count-based data. We choose this method because our dependent variable (the number of COVID-19–infected people) consists of only non-negative integer values and the variance of the dependent variable is greater than the mean.

The dependent variable is substantially positively skewed and kurtotic (skewness = 6.002, kurtosis = 43.308). Owing to the non-normal, highly-skewed and non-continuous nature of the dependent variable, standard linear regression techniques (such as ordinary least squares regression) are not suitable for this data set. Therefore, Poisson regression (PR) is the first-choice modeling technique and can be defined as follows:

$$P(Y_i = y_i) = \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!} \quad (1)$$

where $P(\cdot)$ shows the probability of Y infected people observed in the i^{th} country over a specified time period (until April 3rd). y_i can take the values 0,1,2,... and μ_i denotes the expected COVID-19–infected frequency for country i . In PR model, the expected

COVID-19–infected frequency (μ_i parameter) is estimated as a function of the vector of exploratory variables x_i such that:

$$\ln(\mu_i) = x_i^T \beta \quad (2)$$

where β is a vector of estimated coefficients of exploratory variables including the percentage of elderly people in population, the logarithm of the population density, the number of airline passengers, the number of airport, and the Schengen system. The vector of coefficients is then estimated by maximizing the logarithm of the likelihood function given below.

$$\ln L(\beta) = \sum_i \left[-\exp(x_i^T \beta) + (x_i^T \beta) y_i - \ln y_i! \right] \quad (3)$$

One of the important properties of the Poisson distribution is that the mean and the variance are equal to the μ_i parameter. However, the assumption of identical mean and variance was not satisfied for the data used in this study ($\mu = 5,498.25$ and $\sigma = 21,577.85$). The greater ratio of variance to mean leads to overdispersion frequently caused by heterogeneity among observations. Thus, we apply NBR to overcome this problem of overdispersion. A gamma-distributed error term is added to Eq (2) to relax the PR assumption by including additional randomness.

$$\ln(\mu_i) = x_i^T \beta + \varepsilon_i \quad (4)$$

where ε_i follows gamma distribution with mean 1 and variance α . The NBR distribution has a mean μ_i and variance $\mu_i + \alpha \mu_i^2$, where α is the overdispersion parameter used as a measure of dispersion. In analyzing the variation in the dependent variable, the following model is considered:

$$\begin{aligned} \ln(\text{Num.of.COVID19 infected}) \\ = \beta_0 + \beta_1 * (\text{old}) + \beta_2 * (\log.\text{popdensity}) + \beta_3 * (\log.\text{airtransfer}) \\ + \beta_4 * (\log.\text{airportnumber}) + \beta_5 * (\text{Schengen}) \end{aligned}$$

Results

Table 1 shows the estimates of model parameters ($\hat{\beta}$), standard errors (Std Err of $\hat{\beta}$), 95% confidence interval (CI) for the $\hat{\beta}$ by profiling the likelihood function, incident rate ratios (IRRs) and goodness-of-fit statistics such as Cragg–Uhler pseudo- R^2 , logarithmic likelihood, and Akaike information criteria (AIC).

The estimated coefficients of all variables used in this study are statistically significant (at least 95% confidence level) and in the

Table 1
Estimation results of negative binomial regression.

Variables	Estimate	Std err of $\hat{\beta}$	Conf. Int. of $\hat{\beta}$	IRR
Intercept	−5.203***	0.763	(−6.649, −3.699)	0.005
Air transfer	1.151***	0.138	(0.901, 1.394)	3.161
Airport number	0.982***	0.204	(0.549, 1.401)	2.670
Schengen	1.503***	0.431	(0.622, 2.383)	4.498
Population density	0.876***	0.216	(0.308, 1.422)	2.403
Old	0.066**	0.025	(0.017, 0.119)	1.068
Dispersion parameter	0.567	0.058		
Pseudo- R^2	0.76			
Log.Lik.	−1049.602			
AIC	2113.20			

^a Schengen = 0 is taken as reference category; IRRs, incident rate ratios; AIC, Akaike information criteria.

* $P < 0.1$, ** $P < 0.05$, *** $P < 0.01$.

positive direction. Countries that have higher volume of airline passengers (IRR = 3.161, $P < 0.01$); higher number of airports (IRR = 2.607, $P < 0.01$), higher population density (IRR = 2.403, $P < 0.01$), higher percentage of elderly population (IRR = 1.068, $P < 0.05$), and Schengen countries (IRR = 4.498, $P < 0.01$) are found to be more likely to have higher numbers of COVID–19 infected cases than other countries (see also Figure 1 in supplementary material).

Discussion

This study answers the question of why some countries have higher numbers of COVID-19–infected people compared with others. Analysis of the data suggests a link between the scale of human mobility and the number of COVID-19 patients in countries. Our results indicate a positive association between the magnitude of airline travel and high numbers of COVID-19–infected patients. Furthermore, we find that countries which have higher number of airports, Schengen countries, countries which have higher population density and higher percentage of elderly population are found to be more likely to have more COVID-19 cases than other countries.

The quick spread of COVID-19 appears to be propelled by ‘superspreading’. Superspreading refers to heightened transmission of the disease to at least eight contacts and has been observed for several infectious diseases including SARS, MERS, and influenza.¹⁹ Our study suggests that better connected areas are more likely to be infected first and have more infections initially (but it is still too early to report the potential consequences on less well-connected areas that may become infected in due course).

There are a number of limitations in this article. Although we measured human mobility by looking at the volume of air travel, future studies can provide a comprehensive analysis on the impact of sea/bus/train/car travels on the spread of COVID-19. Patients zero and their travel history will provide important insights into cross-country comparisons. In addition, when a virus arrives in a country, its contagion and spread hinges on local transmission pathways and public health provision.¹⁹ Efforts and (relative) successes of countries in handling the COVID-19 crisis should be analyzed in a comparative manner. Furthermore, we note that certain emerging trends might influence general applicability of the findings as we move into the future. For example, in addition to the reduced volume of travel, increased testing and future development of vaccines might also affect the applicability of the findings with passage of time.

Previous studies found that airport screening measures failed in halting the spread of viruses.²⁰ In the context of superspreading of COVID-19, airports are more likely to be rearranged so as to minimize the risk of contagion. Researchers should contemplate on new techniques and methods at airports for the maximum safety of passengers and staff against pandemics diseases. There are also issues that urgently need to be further studied, such as the link between public health provision and COVID-19 mortality rates. Our study indicates a positive relationship between the percentage of elderly population and COVID-19 cases. Recent developments reveal that the virus has the potential to affect all age cohorts. Future studies can comparatively examine the spread and the mortality rate of COVID-19 in countries with younger population and those with aging populations. Psychological impacts of the COVID-19 pandemic also need to be systematically studied. The long-term implications of the COVID-19 pandemic on countries’ health systems and global health policymaking and management strategies will also provide interesting avenues of research for further researchers.

Author statements

Acknowledgements

We would like to thank the two anonymous reviewers and editors of Public Health for their constructive comments.

Ethical approval

None sought.

Funding

None sought.

Competing interests

None sought.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.puhe.2020.07.002>.

References

1. Wells GJ, Shuey R, Kiely R. *Globalization*. New York: Novinka Books; 2001.
2. Tatem AJ, Rogers DJ, Hay SI. Global transport networks and infectious disease spread. *Adv Parasitol* 2006;**62**:293–343.
3. Wilson ME. Geography of infectious diseases. In: Cohen J, Powderly WG, editors. *Infectious diseases*. 2nd ed. New York, NY: Mosby; 2004. a Hopkins DR. *Princes and peasants: smallpox in history*. Chicago: University of Chicago Press; 1983.
4. Hays JN. *The burdens of disease*. New Brunswick, NJ: Rutgers University Press; 2000.
5. Lane K. Spain and Spanish America. In: Francis JM, editor. *Iberia and the americas: culture, politics, and history: a multidisciplinary encyclopedia*, vol. 1. Santa Barbara: ABC Clio; 2006.
6. Tatem AJ, Rogers DJ, Hay SI. Global transport networks and infectious disease spread. *Adv Parasitol* 2006;**62**:293–343. a Oxford J. The great influenza: the epic story of the 1918 pandemic. *Nature* 2004;**429**:345–6.
7. WHO. *Epidemiologic aspects of small pox in Yugoslavia in 1972*. WHO/SE/73.57. 1973.
8. Wilson ME. Travel and the emergence of infectious diseases. *Emerg Infect Dis* 1995;**1**:39–46.
9. Olsen SJ, Chang HL, Cheung TYY, et al. Transmission of the severe acute respiratory syndrome on aircraft. *N Engl J Med* 2003;**349**:2416–22.
10. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 2020;**395**:689–97.
11. World Health Organization. *WHO statement on novel coronavirus in Thailand*. 13 January 2020. <https://www.who.int/news-room/detail/13-01-2020-who-statement-on-novel-coronavirus-in-thailand>.
12. World Health Organization. *Novel coronavirus – Japan (ex-China)*. 16 January 2020. <https://www.who.int/csr/don/16-january-2020-novel-coronavirus-japan-ex-china/en/>.
13. Al Hasan SM, Saulam J, Kanda K, et al. The novel coronavirus disease (COVID-19) outbreak trends in mainland China: a joinpoint regression analysis of the outbreak data from January 10 to February 11, 2020. *Bull World Health Organ* 17 February 2020. <https://doi.org/10.2471/BLT.20.253153> [Submitted]. E-pub. a Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* 2020;**395**:689–97.
14. World Health Organization. *Coronavirus disease (COVID-2019) situation reports*. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>. [Accessed 3 April 2020].
15. World Development Indicators. *Air transport, registered carrier departures worldwide*. <https://data.worldbank.org/indicator/IS.AIR.DPRT?view=chart>; 2018.
16. Central Intelligence Agency. *The world Factbook*. <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2053rank.html>.
17. Wilson ME. Population mobility and the geography of microbial threats. In: Apostolopoulos Yorghos, Sönmez Sevil, editors. *Population mobility and infectious disease*. New York: Springer; 2007. p. 32.
18. World Development Indicators. *Population density (people per sq. km. of land area)*. <https://data.worldbank.org/indicator/en.pop.dnst>; 2018. a World Development Indicators. *Population ages 65 and above(% of total population)*. <https://data.worldbank.org/indicator/sp.pop.65up.to.zs>. [Accessed 30 March 2020].
19. Lee A. Wuhan novel coronavirus (COVID-19): why global control is challenging? *Publ Health* 2020;**179**:A1–2.
20. John RK, King A, De Jong D, Bodie-Collins M, Squires SG, Tam TW. Border screening for SARS. *Emerg Infect Dis* 2005 Jan;**11**(1):6. a Selvey LA, Antao C, Hall R. Evaluation of border entry screening for infectious diseases in humans. *Emerg Infect Dis* 2015 Feb;**21**(2):197.