Multiscale mobility explains differential associations between the gross domestic product and COVID-19 transmission in Chinese cities

Authors: Xiao-Ke Xu[1,+], PhD, Lin Wang[2,+], PhD, Sen Pei[3,+,*], PhD

Affiliations:

[1]College of Information and Communication Engineering, Dalian Minzu University, Dalian 116600, China

[2]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK

[3]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA

[*] To whom correspondence should be addressed: Email: sp3449@cumc.columbia.edu

[+] These authors contributed equally to this study.

Keywords: COVID-19, human mobility, socioeconomic factors, transmission risk, GDP

Highlight/Teaser

In this letter, we find a Simpson's paradox in the association between GDP and COVID-19 transmission in Chinese cities stratified by location. The differential associations in cities within and outside Hubei province can be explained by different patterns of short-range and long-range multiscale mobility from Wuhan to other cities.

Main Text

The rapid dissemination of SARS-CoV-2 have caused a global pandemic of the novel coronavirus disease 2019 (COVID-19). Despite that early Non-Pharmaceutical Interventions (NPIs) have contained or mitigated the initial spread of COVID-19 in a number of countries, resurgence of cases has been reported in the northern hemisphere entering the winter of 2020. Due to the recurrent and endemic potential of COVID-19 outbreaks,[1] understanding the risk factors associated with COVID-19 transmission is essential for preparedness and response.

Recent studies suggest that human mobility[2-6] and geographic components[7] are major risk factors driving spatial spread of COVID-19. For instance, Jia et al.[2] and Tian et al.[3] both found that the cumulative number of confirmed cases was positively correlated with the cumulative number of passengers travelled from the epicenter (i.e., Wuhan, Hubei province) to each other city in mainland China. Here, using granular human mobility data in mainland China, we quantify the association between COVID-19 transmission and an important socioeconomic factor, the gross domestic product (GDP), in 296 Chinese cities. In China, cities with higher GDP are generally more populated and urbanized. We hypothesize that more frequent human interactions in urban settings could potentially increase the transmission risk of COVID-19.

The aforementioned seminal studies[2-7] mainly analyzed the daily number of confirmed COVID-19 cases documented by health authorities. However, documented case count data can substantially underestimate the transmission risk and epidemic size due to underreporting and reporting delay.[8] To more accurately describe the COVID-19 transmission, we estimated the number of true incident cases (including both documented and undocumented cases) infected each day in each of 296 Chinese cities before February 8, 2020. We conducted this estimation using a dynamic metapopulation model that accounts for unreported cases and adjusts for the lag from infection acquisition to confirmation (Supplementary Information).[8]

We computed the correlation between the GDP in 2018[2] and estimated cumulative infections over time (Fig. 1A, Table 1). A weak but stable positive correlation is observed across the whole study period, which agrees with our hypothesis. However, a simple stratification of all 296 cities according to whether a city is located within Hubei province unravels a much higher positive correlation for cities outside Hubei province and a moderate negative correlation for cities within Hubei province. Such a clear disparity in the city-stratified associations between GDP and COVID-19 transmission is reminiscent of the Simpson's paradox (i.e., distinct associations in certain subgroups disappear when subgroups are merged together).[9] Differential associations between GDP and the cumulative confirmed cases were also observed in Chinese cities stratified by location (Fig. S1). Similar patterns remained between cumulative infections per 10,000 people and GPD per capita (Figs. S2-3), as well as other three socioeconomic indictors for wealth and well-being, including average wage of employed staff and workers, public finance expenditure per capita and household saving deposits at year-end per capita (Fig. S4).

We find that the distinct association between GDP and COVID-19 transmission can be explained by the multiscale mobility from Wuhan to other cities. Before Wuhan's lockdown on January 23, 2020, transportation of several billion trips during the Spring Festival season (i.e., "chunyun") significantly accelerated the spatial spread of COVID-19 throughout China.[3,4,7] Using Baidu mobile phone data between January 1 and January 24, 2020, we found that 75.7% travelers leaving from Wuhan moved to cities within Hubei province through short-range travels. We then used data from Tencent during "chunyun" in 2018 to estimate the proportion of travelers transported via highway, railway, and airway (Fig. 1B). The within- and cross-province travels from Wuhan exhibit dramatic difference in mode of transport.

We observed different mobility patterns for short-range and long-range travels originating from Wuhan. As a regional economic center, Wuhan attracts a large number of migrant workers and new dwellers from nearby cities in Hubei province. To celebrate Spring Festival, many of them returned their birthplaces to visit the original family, traveling mainly by railway and highway. For cities within Hubei province, we found a negative correlation between the GDP and daily average population outflow from Wuhan (Fig. 1C), which reflects that people from less urbanized cities in Hubei province tend to migrate to Wuhan for seeking jobs or attending schools. Within Hubei province, Wuhan is strongly connected with regional cities with lower GDP via short-range mobility during the “chunyun” period. For cities outside Hubei province, we found a positive correlation between the GDP and daily average population outflow from Wuhan (Fig. 1D). Due to close economic ties and the hub-and-spoke structure of railway and flight networks, Wuhan is also strongly connected with other economic centers outside Hubei province via long-range railway and air travel. Given that the spatial spread of COVID-19 was largely driven by the population outflows leaving from Wuhan during “chunyun” period,[2] we expect that more infections occurred in low-GDP cities within Hubei province and high-GDP cities outside Hubei province. The stable correlation between GDP and cumulative infections for cities outside Hubei province indicates the success of NPIs to contain subsequent transmission from high-GDP cities to low-GDP cities in other provinces.[3,5]

The impact of multiscale mobility on the spatial spread of infectious diseases has been explored in previous mathematical modeling studies.[10] Numerical simulations of 2009 A/H1N1 pandemic influenza have suggested that the disease first rapidly spread from the Greater Mexico City to most hub populations of economic centers around the globe via long-range airline travels. Once a hub population established local transmission, the disease infected the surrounding areas via short-range ground travels. Here we provide a real-world case study of COVID-19 to demonstrate that structure and properties of multiscale mobility could shape the spatial transmission dynamics of SARS-CoV-2. As such, bulk analysis mixing up multiple geographical scales may lead to contradicting conclusions when evaluating the associations between the COVID-19 transmission and socioeconomic factors that differ for short-range and long-range mobility (e.g., urbanization). Similar issues may also occur during the early phase of an outbreak of an emerging infectious disease. A clear understanding of the mechanism underlying such differential patterns will be useful to the strategic response and planning against COVID-19.

Author Contributions

All authors contributed equally to the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China [61773091 and 61603073 to X. X., 11975025 to L. W.]; the LiaoNing Revitalization Talents Program [XLYC1807106 to X. X.]; and the National Science Foundation [DMS-2027369 to S. P.]. We acknowledge the resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk).

Conflict of Interest

No reported conflicts of interest.

Acknowledgment

The authors are in particular grateful to Dr. Ye Wu, Dr. Xiaofan Liu, Dr. Xin Lv and Dr. Zhanwei Du for comments and suggestions.

Reference

1. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* 2020; **368**:860–868.

2. Jia JS, Lu X, Yuan Y, Xu G, Jia J, Christakis NA. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* 2020; **582:**389–394.

3. Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 2020; **368:**638–642.

4. Zhong P, Guo S, Chen T. Correlation between travellers departing from Wuhan before the Spring Festival and subsequent spread of COVID-19 to all provinces in China. *J Travel Med* 2020; **27:**taaa036.

5. Kraemer MUG, Yang C-H, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020; **368:**493–497.

6. Du Z, Wang L, Cauchemez S, et al. Risk for Transportation of Coronavirus Disease from Wuhan to Other Cities in China. *Emerg Infect Dis* 2020; **26:**1049–1052.

7. Rader B, Scarpino SV, Nande A, et al. Crowding and the shape of COVID-19 epidemics. *Nat Med* 2020; https://doi.org/10.1038/s41591-020-1104-0.

8. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020; **368:**489–493.

9. Blyth CR. On Simpson's paradox and the sure-thing principle. *J Am Stat Assoc* 1972; **67:**364–366.

10. Balcan D, Colizza V, Goncalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA* 2009; **106:**21484–21489.

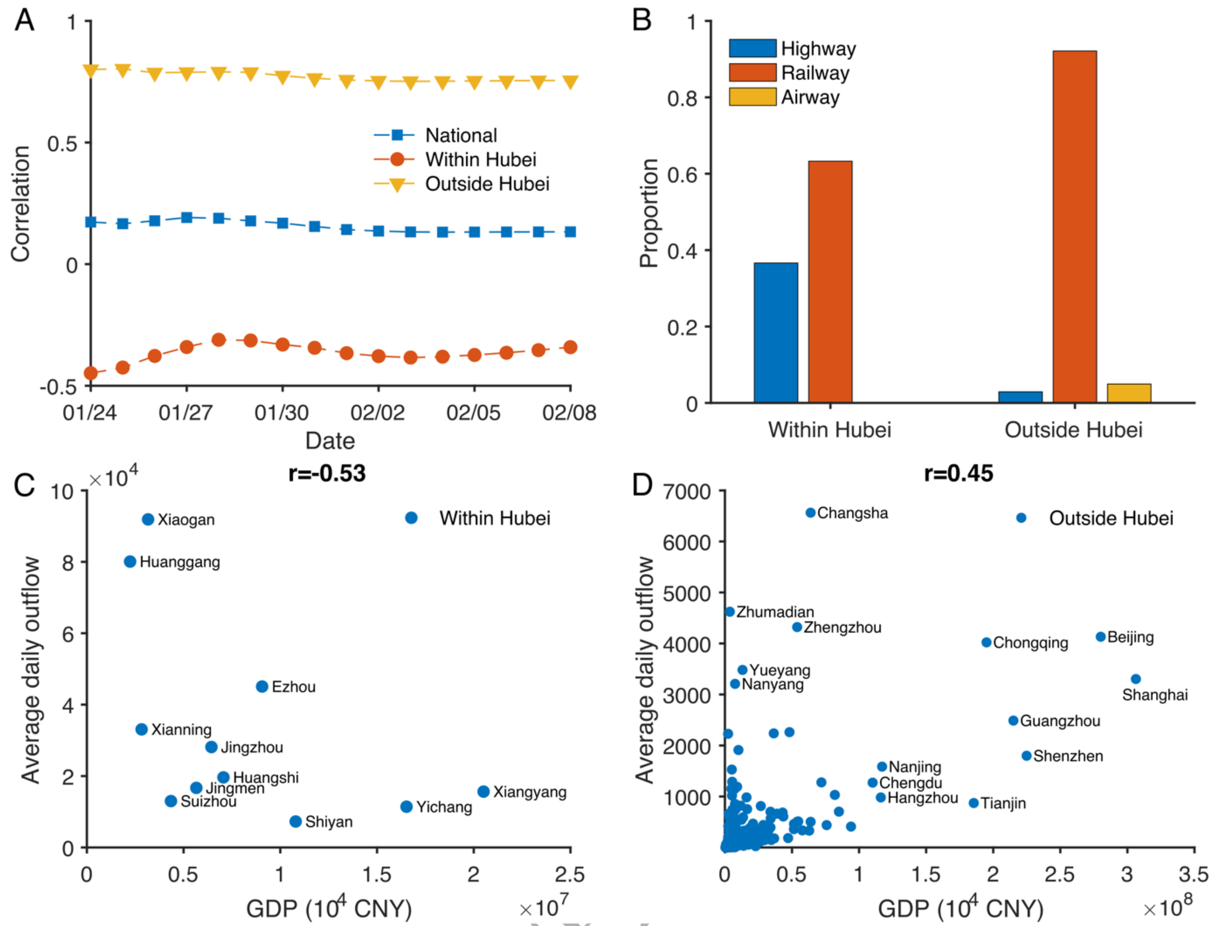

Figure 1. A, Pearson correlation between the estimated cumulative infections on each day from January 24 to February 8, 2020 and the GDP of cities within Hubei province (blue square), outside Hubei province (red circle) and all 296 cities (yellow triangular). B, Proportion of highway (blue bar), railway (red bar), and airway (yellow bar) travels among population outflow leaving from Wuhan to cities within (left) and outside (right) Hubei province. The short-range inter-city movements within Hubei province were mainly transported via railway (63.3%) and highway (36.6%), with negligible airway transportation (<0.1%). In contrast, the long-range movements from Wuhan to cities in other provinces were dominated by the railway transportation (92.1%), with airway and highway accounting for 5.0% and 2.9% total travels, respectively. C, Relationship between GDP and the average daily outflow from Wuhan to 11 cities within Hubei province between January 1 and January 24, 2020 (Pearson correlation coefficient $r = -0.53$). D, Relationship between GDP and the average daily outflow from Wuhan to 285 cities outside Hubei province (Pearson correlation coefficient $r = 0.45$).

Date	National	Within Hubei	Outside Hubei
24-Jan	0.17	-0.45	0.80
25-Jan	0.17	-0.43	0.80
26-Jan	0.18	-0.38	0.79
27-Jan	0.19	-0.34	0.79
28-Jan	0.19	-0.31	0.79
29-Jan	0.18	-0.31	0.79
30-Jan	0.17	-0.33	0.77
31-Jan	0.15	-0.34	0.76
1-Feb	0.14	-0.37	0.76
2-Feb	0.14	-0.38	0.75
3-Feb	0.13	-0.38	0.75
4-Feb	0.13	-0.38	0.75
5-Feb	0.13	-0.37	0.75
6-Feb	0.13	-0.36	0.75
7-Feb	0.13	-0.35	0.75
8-Feb	0.13	-0.34	0.76

Table 1. Pearson correlation between the estimated cumulative infections on each day from January 24 to February 8, 2020 and the GDP of cities within Hubei province, outside Hubei province and all 296 cities.