



Article

Revisiting the Effects of High-Speed Railway Transfers in the Early COVID-19 Cross-Province Transmission in Mainland China

Chun-Hsiang Chan and Tzai-Hung Wen *

Department of Geography, National Taiwan University, Taipei 10617, Taiwan; d04228002@ntu.edu.tw
* Correspondence: wenhung@ntu.edu.tw

Abstract: Coronavirus disease 2019 (COVID-19) is an ongoing pandemic that was reported at the end of 2019 in Wuhan, China, and was rapidly disseminated to all provinces in around one month. The study aims to assess the changes in intercity railway passenger transport on the early spatial transmission of COVID-19 in mainland China. Examining the role of railway transport properties in disease transmission could help quantify the spatial spillover effects of large-scale travel restriction interventions. This study used daily high-speed railway schedule data to compare the differences in city-level network properties (destination arrival and transfer service) before and after the Wuhan city lockdown in the early stages of the spatial transmission of COVID-19 in mainland China. Bayesian multivariate regression was used to examine the association between structural changes in the railway origin-destination network and the incidence of COVID-19 cases. Our results show that the provinces with rising transfer activities after the Wuhan city lockdown had more confirmed COVID-19 cases, but changes in destination arrival did not have significant effects. The regions with increasing transfer activities were located in provinces neighboring Hubei in the widthwise and longitudinal directions. These results indicate that transfer activities enhance interpersonal transmission probability and could be a crucial risk factor for increasing epidemic severity after the Wuhan city lockdown. The destinations of railway passengers might not be affected by the Wuhan city lockdown, but their itinerary routes could be changed due to the replacement of an important transfer hub (Wuhan city) in the Chinese railway transportation network. As a result, transfer services in the high-speed rail network could explain why the provinces surrounded by Hubei had a higher number of confirmed COVID-19 cases than other provinces.



Citation: Chan, C.-H.; Wen, T.-H. Revisiting the Effects of High-Speed Railway Transfers in the Early COVID-19 Cross-Province Transmission in Mainland China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 6394. <https://doi.org/10.3390/ijerph18126394>

Academic Editor: Paul B. Tchounwou

Received: 11 April 2021

Accepted: 11 June 2021

Published: 13 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: transfer service; COVID-19; Wuhan city lockdown; high-speed rail network; intercity population flow; spatial transmission

1. Introduction

For a respiratory infectious disease, person-to-person contact plays a vital role in the transmission of an epidemic, and a large population flow raises the probability of contact between people [1–6]. An imported case of the disease could trigger a serious local outbreak. Border control policies could be an effective measure to prevent imported cases from causing local outbreaks. Therefore, recent studies have focused on assessing the effectiveness of national border control policies in containing the domestic and global spread of respiratory infectious diseases, such as body temperature screening in airports and comprehensive travel history investigations [7–13]. Some studies have found evidence that supports this approach because it can delay the peak of the pandemic and reduce the number of confirmed cases; however, some have argued that the efficiency of disease control is still limited [14]. The primary reason is that the implementation of border control is too slow and that local transmission has already occurred by the time it is implemented. Another reason is that the accuracy of positive detection during border screenings is too low to efficiently differentiate infected persons because fever is not a consistent symptom

of influenza, or symptoms may develop after arrival [15–17]. For instance, only 13.8% of infected people were detected in Australia during the SARS epidemic [18], and only 6.6% of imported cases were detected in Japan during the H1N1 period [17].

Due to its high transmissibility, COVID-19 is an ongoing pandemic that was reported at the end of 2019 in Wuhan, China [19–21] and rapidly disseminated to all provinces in around one month [22]. The virus was primarily spread through the massive annual cross-province migration during the Spring Festival from 24 to 30 January in 2020 [23]. The Spring Festival travel season, also referred to as Chunyun, is from 10 January to February 18 in 2020 [24]. During this period, the demand for domestic transport (railways, air transportation, and bus systems) reaches its peak [25]; furthermore, railways are one of the most crucial transport modes because of their high accessibility and low cost. The city of Wuhan, which is a critical transfer hub in Chinese railway transport, has two primary inter-city rails (the Jingguang railway and the Shanghai–Wuhan–Chengdu passenger railway) connected to most of the major cities (e.g., Beijing, Shanghai, and Guangzhou) [26,27]. In consideration of both the severe situation of COVID-19 and the hub characteristics during Chunyun, the Chinese government enacted emergency measures in Wuhan city and the surrounding cities and counties of Hubei Province on 23 January effective from 25 January to 8 April in order to restrict all population outflows, control pandemic transmission, and expand social distancing. Hence, the Wuhan city lockdown might have seriously affected population flows in railway transport.

Recent studies have assessed the impacts of the Wuhan city lockdown on pandemic transmission through time and space. By simulating different levels of travel reduction from Hubei or Wuhan, researchers could estimate the spatial progress of the pandemic within neighboring provinces and overseas airports before and after the Wuhan city lockdown [28]. These studies have pointed out that greater reductions in travel efficiently decreased the development of the pandemic in the surrounding provinces or countries connected by roads or flights. Moreover, some studies have also indicated that city lockdown policies could delay the timing of the infection peak [29–34]. In other words, city lockdowns controlled pandemic transmission and bought time for the healthcare system to address the outbreaks [24,35]. Due to Wuhan's status as an important railway hub in mainland China, the Wuhan city lockdown has reduced many infections because of the resultant sizable cross-province population flow changes. These studies have indicated that the provinces bordering Hubei had a higher number of confirmed cases than other provinces, even after the Wuhan city lockdown [24]. However, the possible reasons of a large number of confirmed cases concentrated in the neighboring provinces of Hubei have not been fully discussed. Therefore, this study aims to assess the changes in railway passenger transport on the early spatial transmission of COVID-19 in mainland China. Daily railway schedule data were used in this study to measure intercity population mobility patterns in order to capture the impact of the Wuhan city lockdown. Examining the role of railway transport properties in disease transmission could help quantify the spatial spillover effects of large-scale travel restriction interventions, such as city lockdowns. It could improve our understanding of the spatial diffusion patterns of COVID-19 during early transmission in mainland China.

2. Data

2.1. Railway Schedule Data

To compare the changes in railway transportation before and after the Wuhan city lockdown, we collected the daily train timetable from the official Chinese railway reservation website (<https://www.12306.cn/index/>, accessed on 22 January 2020) by web crawler from 22 to 26 January 2020. Each train's unique key is collected daily from the official Chinese train reservation website; then, the complete train timetable of the specific train could be fetched by the unique key and save to the local train timetable database. There are 12 columns within train timetable data, including the day of arrival (e.g., today or tomorrow), station name, train type, base station indicator (1 represents the first station,

while “NaN” indicates other stations), station name of the final destination, the arrival time of the station, base station name, train code, the day offset of arrival (0 means the train will arrive today, while 1 indicates the train will arrive tomorrow), departure time of this station, station index (i.e., the station index of this train), and duration (from the previous station to this station). There are approximately 8540 timetables for each day. The daily schedule data cover 3129 stations in 325 cities. Four fields for each record are used in this study, including the names of the departure and destination stations, the name of the station stops along the route, and the train code. The Chinese railway includes 10 different train types (G, C, D, Z, T, K, S, L, Y, and N). Train types G, C, and D are cross-province transport with high-speed trains; Z, T, and K are intercity transport with lower-speed trains; and the rest of the train types are within-city transport. In this study, we focused on the cross-province movement of people; therefore, two major types of high-speed trains were incorporated into the following analyses: the D-series and G-series high-speed trains. (Train types G and C are of the same train type; therefore, we combined them into one train type, G.) Within the dataset collected, 32% and 42% of trains are G-Series and D-Series High-Speed Train each day, respectively; furthermore, G-Series and D-Series High-Speed Train cover 960 stations and 256 cities.

We further transformed the daily train tables into city-to-city origin-destination (OD) networks to visualize and analyze the spatial patterns of intercity railway transport. The nodes of the OD network represent each city, and the link weights represent the frequency with which trains travel between those cities. We obtained the differences in train frequency before and after the Wuhan city lockdown (22 January/26 January 2020) to represent the changes in intercity passenger movement due to the large-scale travel restriction policy.

2.2. COVID-19 Cases

The reported date for each confirmed COVID-19 case was collected from summary reports from the National Health Commission of China [36]. The data collection period ranges from 22 January 2020 to 25 March 2020. To depict the temporal epidemic progression in the early transmission period, we adopted the study conducted by Li, Guan, Wu, Wang, Zhou, Tong, Ren, Leung, Lau, Wong, Xing, Xiang, Wu, Li, Chen, Li, Liu, Zhao, Liu, Tu, Chen, Jin, Yang, Wang, Zhou, Wang, Liu, Luo, Liu, Shao, Li, Tao, Yang, Deng, Liu, Ma, Zhang, Shi, Lam, Wu, Gao, Cowling, Yang, Leung, and Feng [19]) to estimate the date of symptom onset for each confirmed case from symptom onset distributions and reported date distributions in Figure 1.

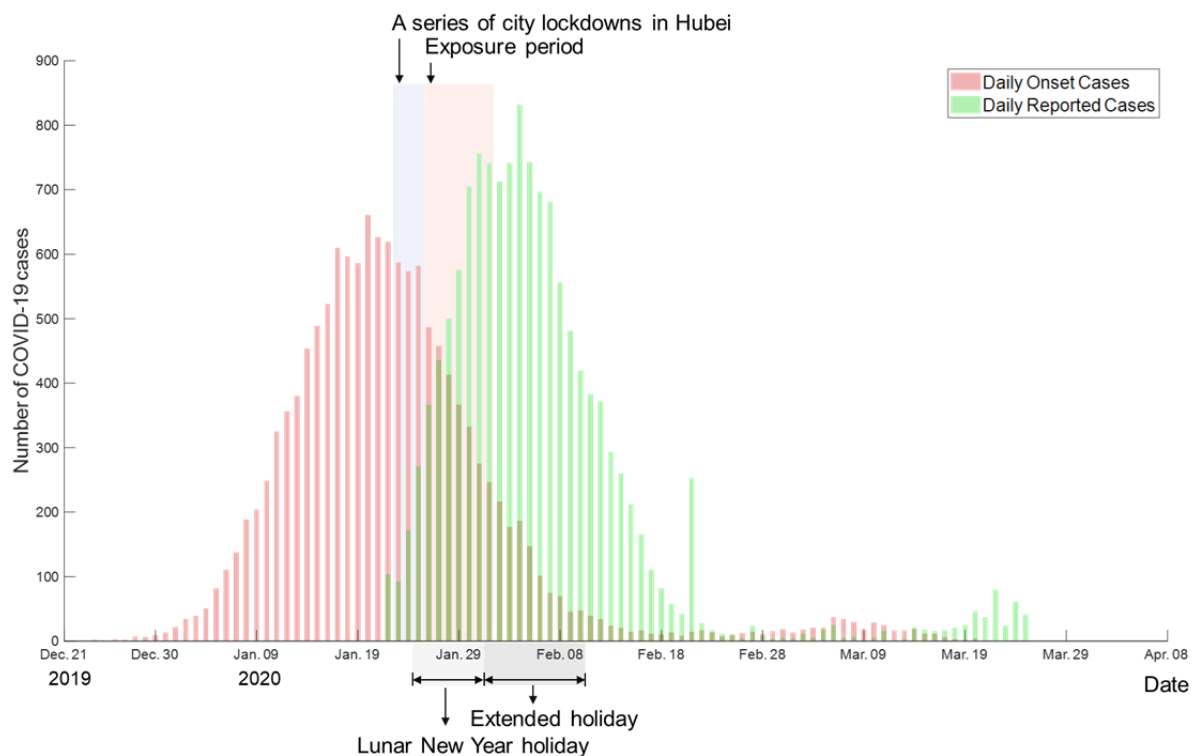


Figure 1. The timeline of pandemic transmission in mainland China (Hubei province was excluded).

3. Methods

In this cross-sectional study, we used a daily train timetable (960 stations) to observe the changes in intercity passenger movement (256 cities) before and after the Wuhan city lockdown. Two network properties, PageRank and betweenness centrality were used to quantify the structural changes in the intercity railway OD network after the Wuhan city lockdown. The spatial distribution of intercity transportation network properties was characterized by bivariate k function and k -nearest neighbor statistic. Eight province-level socioeconomic indicators were considered as confounders. Bayesian multivariate regression was used to examine the association between structural changes in the railway OD network and the incidence of COVID-19 cases.

3.1. Structural Changes in Intercity Railway Transport

Wuhan is an important railway hub during Chunyun; thus, the Wuhan city lockdown substantially affected the cross-province population flow. This could be reflected in the structural changes to the intercity railway OD network. We used network centrality metrics to measure the nodal properties of the OD network, including PageRank centrality, which represents the destination arrival probability (DAP), and betweenness centrality, which reflects the potential for transfer activity (PTA) for each city.

PageRank centrality (PR) measures the probability that a person randomly clicks on a particular link from anywhere; hence, a higher value indicates higher importance [37,38]. The definition of the PR of a particular city C_i is as follows:

$$PR(C_i) = \sum_{C_j \in M(C_i)} \frac{PR(C_j)}{L(C_j)}, \quad (1)$$

where $M(C_i)$ is the set of cities that connect to city C_i , and $L(C_j)$ is the number of outbound connections with city C_j . The recursive equation starts with rank $PR(C_j)$ and then stops the iteration when it converges. PageRank represents the staying probability of a particular city when starting from any city in the railway OD network. Thus, we use this metric as an

indicator to evaluate the DAP of the city. A higher value of the DAP represents a higher probability of staying in that specific city [39].

Betweenness centrality (BC) indicates the ratio of the number of shortest paths through a particular city to the total number of shortest paths between two given cities in the entire city network as follows [40,41]:

$$BC(C_i) = \sum_{k \neq i \neq j \in N} \sigma_{kj}(i) / \sigma_{kj}, \quad (2)$$

where σ_{kj} is the total number of shortest paths between city C_k and city C_j , and $\sigma_{kj}(i)$ is the number of shortest paths through a particular city C_i . Betweenness centrality is adopted to measure the PTA of the city because a higher betweenness centrality value stands for a higher number of transfers passing through the city [42,43].

We then used spatial statistics to compare the spatial patterns in the nodal metrics (PageRank and betweenness centrality) before and after the Wuhan city lockdown. A city with rising nodal metrics represents that city becoming a more influential node in the railway OD network after the Wuhan city lockdown. In contrast, a city with declining nodal metrics represents that city becoming a less influential node. The bivariate k function, a spatially summarized statistic, is used to evaluate the spatial clustering of cities with declining nodal metrics around cities with rising ones [44]. The k -nearest neighbor (kNN) statistic is used to assess the spatial proximity between cities with rising and declining nodal metrics [45,46].

3.2. The Association between Intercity Transportation Network Properties and the Number of Confirmed COVID-19 Cases

The changes in nodal metrics in the railway OD network could alter the contact probabilities within intercity population flows, which may influence the temporal progression of the COVID-19 epidemic. The changes in nodal metrics, PageRank, and betweenness centrality were aggregated into the province level, as were the destination arrival and transfer service statistics, respectively. Bayesian multivariate regression was used to measure the province-level association between the changes in the nodal metrics and the incidence of COVID-19 cases. The province-level socioeconomic indicators are incorporated as control variables in the regression model. These confounders include three dimensions: demographics, economics, and healthcare conditions. The variables in the demographics dimension include the total household population, the resident population percentage (the percentage of people who live in a specific area for six months or more), and the percentage of the household population living in other provinces (the proportion of the difference between the total population and the resident population). The variables in the economics dimension include the average gross domestic product (GDP) per capita. The variables in the healthcare conditions dimension consist of life expectancy, average local health expenditure per person, birth insurance coverage, and average birth insurance expenditure. All data are from the National Bureau of Statistics of China [47–49].

To avoid collinearity among the predictor variables, principal component analysis (PCA) was conducted to reduce the dimensions of the socioeconomic indicators by extracting the principal components. PCA is a statistical method used to obtain principal components from observations through orthogonal transformation. The first principal component (PC) can be regarded as the greatest amount of variance that can explain the most variance in the observations. To prevent the biased estimation of model parameter coefficients due to a small sample size ($n = 28$), a Bayesian multivariate regression is applied to estimate the effect of changes in nodal metrics on the incidence of confirmed COVID-19 cases by controlling for the PCs of the socioeconomic indicators. PCA and the Bayesian linear regression were performed by the R packages BAS 1.5.5 and STATS 3.6.2.

4. Results

4.1. Descriptive Statistics

Figure 2 shows the daily number of G-series and D-series high-speed trains that stopped in Wuhan/Hubei, ranging from 19 January to 2 February 2020. When the city lockdown took effect, the number of each type of train dropped sharply. Within the dataset used, each day, 32% and 42% of trains are G-series and D-series high-speed trains, respectively; furthermore, G-series and D-series high-speed trains cover 960 stations in 256 cities. A significant decrease in the frequency of D-series trains centered on Wuhan city and extended widthwise (e.g., toward Shanghai, Jiangsu, Anhui, Hubei, Chongqing, and Sichuan) and in longitudinal directions (e.g., toward Beijing, Henan, Hubei, Hunan, and Guangdong), which is shown in Figure 2a. The frequency of trains between Guangdong and its surrounding provinces (e.g., Fujian, Yunnan, Guizhou, and Sichuan) also decreased. For G-series trains (Figure 2b), the decreasing pattern extended in a longitudinal direction (Beijing, Shandong, and Fujian) and two widthwise directions (Shanghai, Henan, and Shaanxi; and Zhejiang, Hunan, and Guizhou). Figure 2c shows that the Wuhan city lockdown mainly affected Hubei, Jiangsu, Chongqing, Jiangxi, and Anhui for the D-series high-speed trains. The major provinces affected by the G-series high-speed trains were Hubei, Henan, Shandong, Anhui, and Jiangsu (Figure 2d).

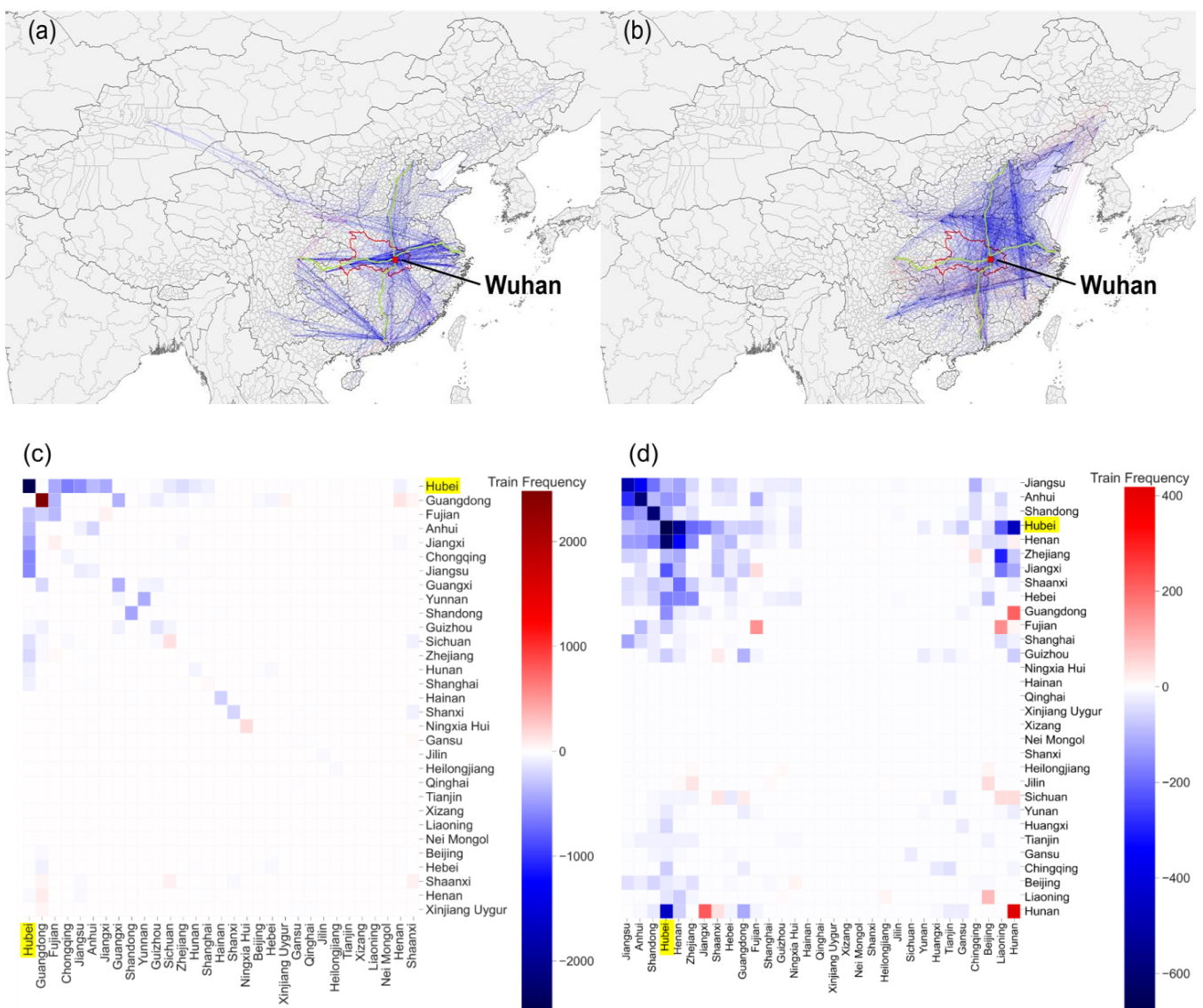


Figure 2. The changes in train frequency in the OD network after the Wuhan city lockdown at the city and province levels,

where (a) D-series and (b) G-series high-speed trains are shown at the city scale. Blue represents decreased train frequency, and red represents increased frequency. The longitudinal railway is the Jingguang railway (green line), and the widthwise railway is the Shanghai–Wuhan–Chengdu railway (green line). The province-level changes for the (c) D-series and (d) G-series high-speed trains are shown to identify the routes most affected by the Wuhan city lockdown.

4.2. The Changes in Nodal Metrics

Figure 3a,b shows the spatial patterns in the changes in DAP. The declining DAP pattern occurs in the major cities lying on the widthwise Shanghai–Wuhan–Chengdu railway (see Figure 2a,b) for the D-series trains. For the G-series trains, a declining DAP pattern occurs in the longitudinal cities along the Jingguang railway (see Figure 2a,b). The declining DAP pattern in the G-series and D-series is consistent with the train frequencies in the railway OD network and shows that the cities with rising DAP surround the cities with declining DAP. Figure 3c,d shows that neither the declining nor rising PTA patterns for the G-series and D-series trains follow a railway route. Most of the cities with declining PTAs are primary cities in mainland China, such as Beijing, Shanghai, and Guangdong. Additionally, the spatial distribution of cities with declining and rising PTA has a pattern similar to that of the distribution of declining and rising DAP.

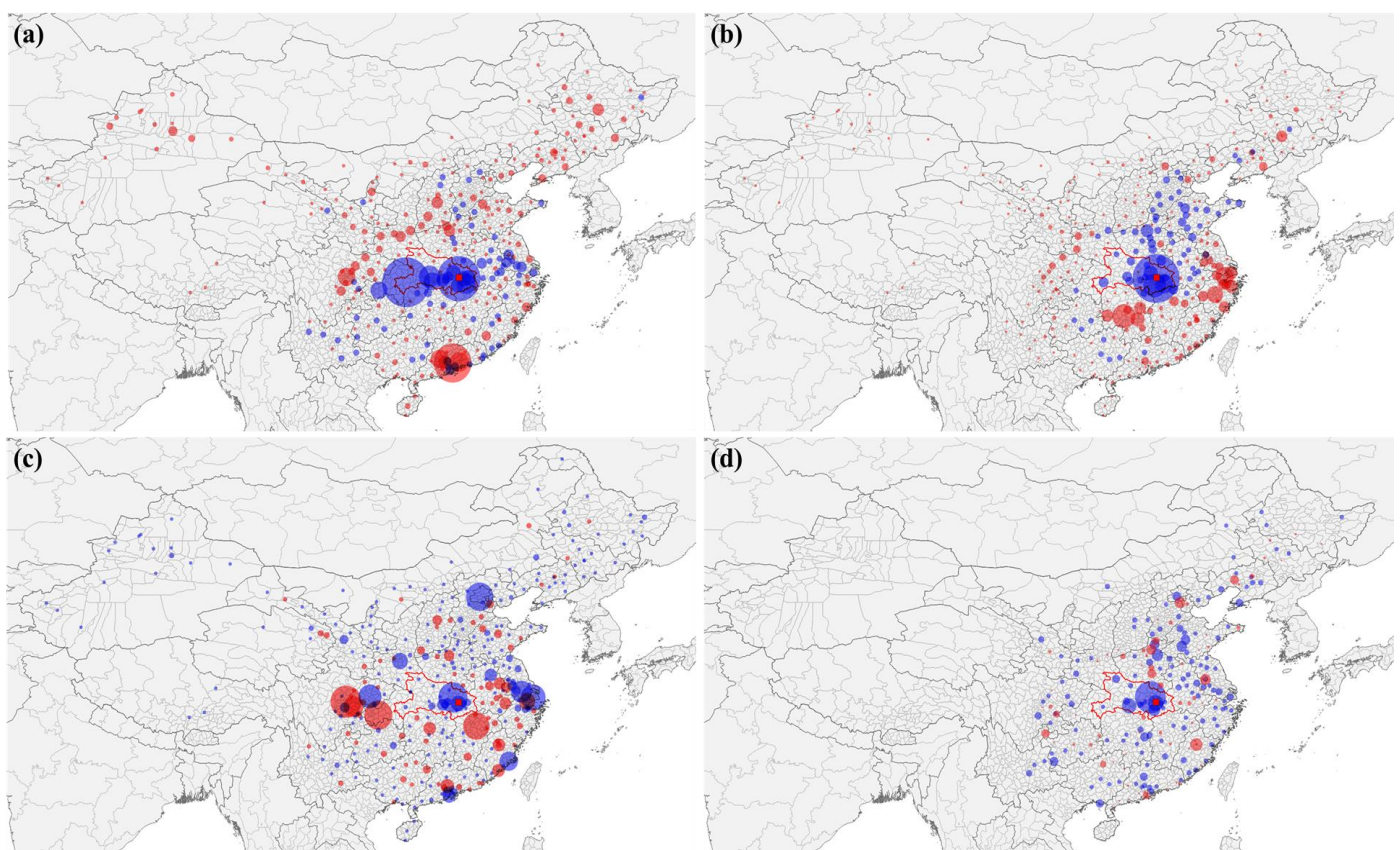


Figure 3. The changes in DAP and PTA after the Wuhan city lockdown. The declining and rising nodal metrics are illustrated by blue and red dots, respectively. The dot size represents the magnitude of the changes in nodal metrics. (a) The DAP changes in D-series trains; (b) the DAP changes in G-series trains; (c) PTA for D-series trains; (d) PTA changes for G-series trains.

4.3. Spatial Patterns in the Nodal Metrics Before and After the Wuhan City Lockdown

Figure 4 represents the results of bivariate spatial analysis, showing spatial clustering patterns of the nodal metrics. It indicates that the cities with declining nodal metrics were

significantly clustered near the cities with rising metrics for all types of high-speed trains and nodal metrics. Then, the median distance between the k -nearest city with rising railway network properties and that with declining railway network properties is presented in Figure 5. The results show that the distance between the two types of cities was larger when considering changes in PTA than changes in DAP for both types of high-speed trains, and they show that the extent of the impact on transfer services is geographically larger than that on destination arrivals. Taking D-series high-speed trains as an example, the distances from the cities with declining DAP and PTA to the 3rd nearest city with rising DAP and PTA were 146.2 km and 258.5 km, respectively. The gray area is the 95% confidence interval according to the $K_{D,R}^{theo}$ value.

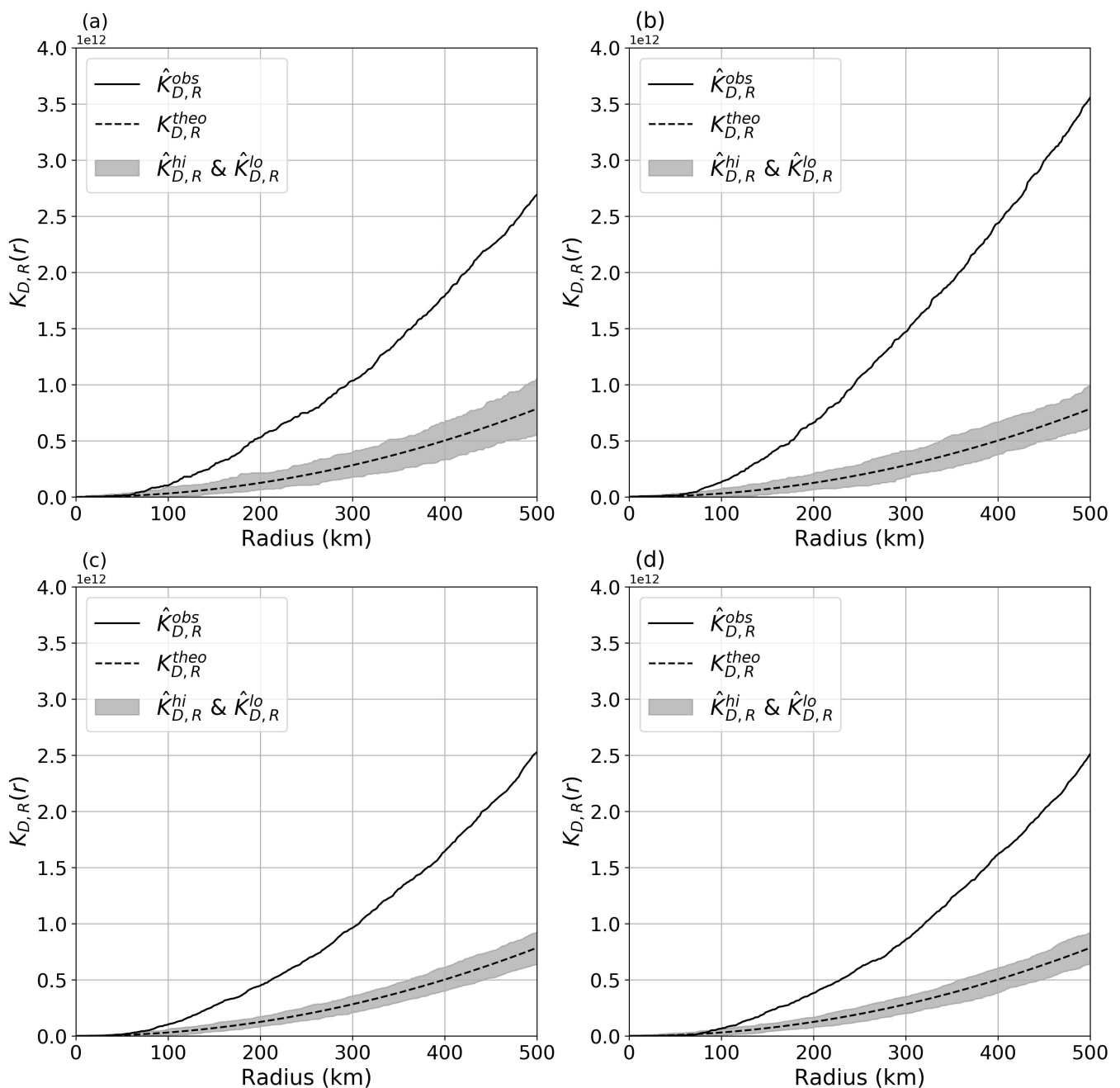


Figure 4. The results of the bivariate k function representing a significant spatial clustering tendency of cities with declining railway network properties around cities with rising properties. The network properties include the potential for transfer activity (PTA) for (a) D-series high-speed trains and (b) G-series high-speed trains and destination arrival probability (DAP) for (c) D-series high-speed trains and (d) G-series high-speed trains.

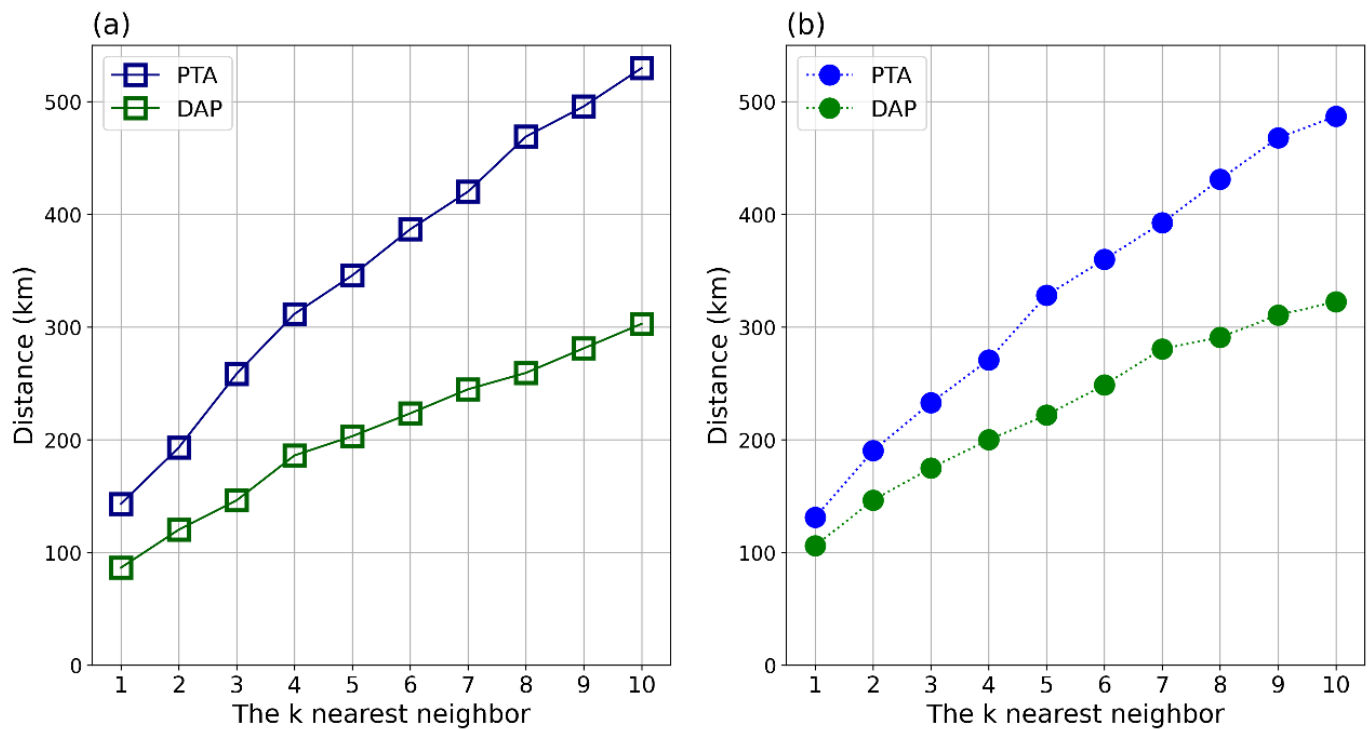


Figure 5. Spatial proximity of cities with declining railway network properties to cities with rising properties, where k varies from one to ten. (a) The square lines are the D-series high-speed trains, and (b) the circle lines are the G-series high-speed trains. Blue is PTA, and green is DAP.

4.4. The Effect of Changes in Nodal Metrics on COVID-19 Transmission

4.4.1. Spatial Distribution of Nodal Metrics

Figure 6a,b shows the spatial changes in destination arrival (DAP) and transfer service (PTA) after the Wuhan city lockdown. These results show that after the Wuhan city lockdown, most provinces had rising DAP, and the provinces with rising PTA were around Hubei Province, including the widthwise provinces from Zhejiang to Qinghai and the longitudinal provinces from Shanxi to Guangxi. This result indicates that the Wuhan city lockdown caused rising destination arrival and high levels of transfer service in neighboring provinces during the early COVID-19 transmission stages.

4.4.2. Clarification of the Role of the Frequency of Transfer Activities in COVID-19 Transmission

Figure 7 shows the proportion of variance explained by eight PCs obtained by PCA. The first three PCs were selected for our regression model because they explain 92% of the variance. According to the variable loadings in Table 1, the first principal component (PC1) was mainly dominated by the resident population percentage. The second principal component (PC2) was driven by the average local health expenditure per person. The total household population and life expectancy represented the third principal component (PC3). Hence, PC1, PC2, and PC3 captured the dimensions of floating populations, healthcare availability, and general demographics, respectively.

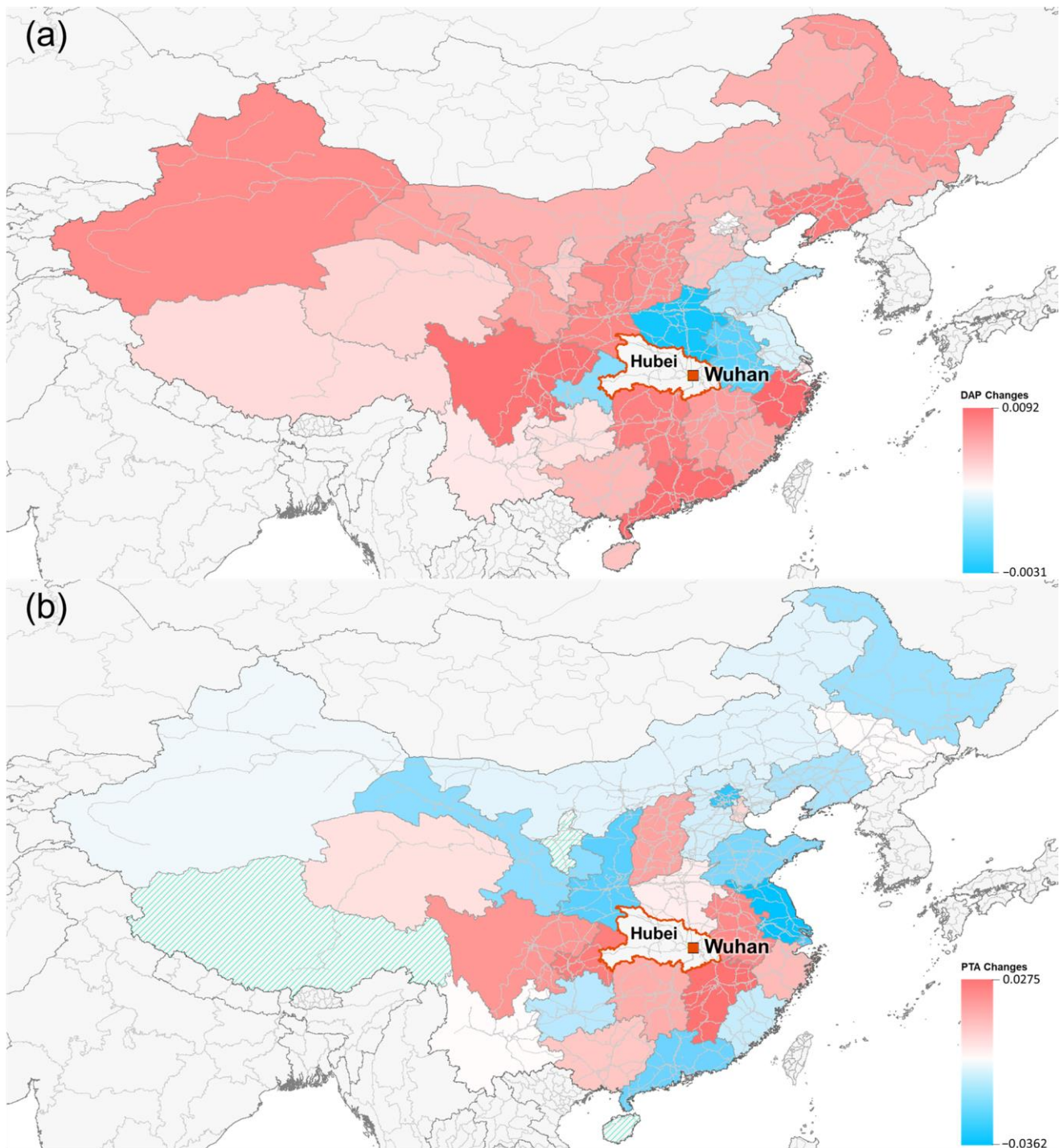


Figure 6. The spatial distribution of the provinces with changes in nodal metrics: (a) DAP and (b) PTA. The red provinces indicate those with rising nodal metrics; the blue provinces indicate those with declining nodal metrics.

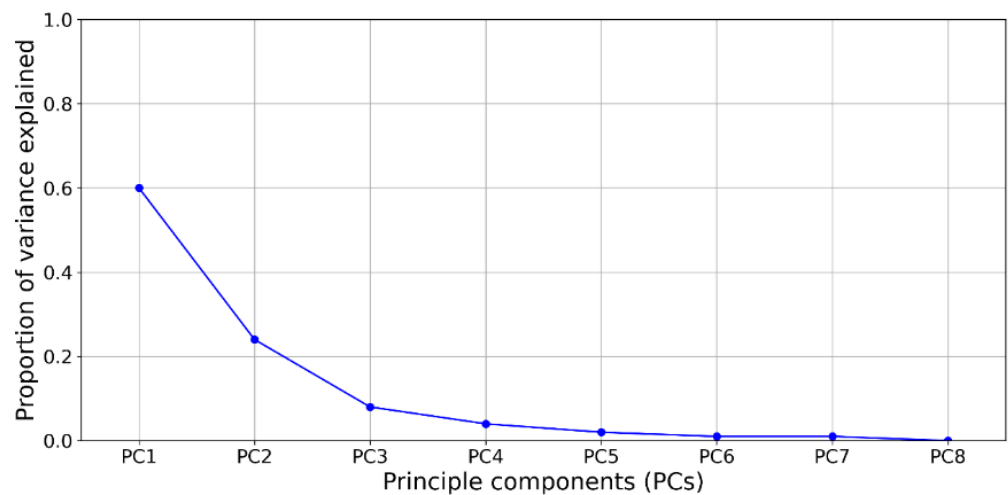


Figure 7. The proportion of variance explained by each principal component.

Table 1. The variable loadings for each principal component.

Confounding	PC1	PC2	PC3
Household population	0.09	−0.53	−0.77
Resident population percentage	0.42	−0.16	0.27
Percentage of the household population in other provinces	−0.42	0.16	−0.27
GDP per capita	−0.41	−0.24	0.11
Life expectancy	−0.26	−0.51	0.44
Average local health expenditure per person	−0.24	0.57	−0.13
Birth insurance coverage	−0.41	−0.16	−0.08
Average birth insurance expenditure	−0.42	−0.05	0.13

The Bayesian regression models evaluate the impact of destination arrival and transfer services on the incidence of confirmed COVID-19 cases, as shown in Tables 2 and 3, respectively. The R^2 of destination arrival and transfer services models are 0.49 and 0.60, respectively; thus, the transfer services model could explain more proportion of variance than the destination arrival one. The results indicate that destination arrival service does not have a significant impact on the incidence of confirmed COVID-19 cases after controlling for the PCs of the socioeconomic indicators; however, transfer service does have a significantly positive effect on early COVID-19 transmission after the Wuhan city lockdown. In other words, transfer activities could promote the spread of COVID-19. Our results also reveal that healthcare availability has a negative effect on the incidence of confirmed cases, while the total population of the province (PC3) has a positive effect (negative regression coefficient and negative PC variable loadings), which means a large population could be regarded as a risk factor for disease transmission. Corresponding to the results of spatial analyses, the neighboring provinces of Hubei with a high number of confirmed cases have a high transfer service score, including Anhui, Jiangxi, Hunan, and Chongqing. The provinces with a low average local health expenditure per person, a short life expectancy, a high household population result in a high number of confirmed cases, such as Henan, Guangdong, Hunan, Zhejiang, Shandong, Jiangsu, and Sichuan. Nevertheless, the top five provinces (Henan, Anhui, Guangdong, Jiangxi, and Hunan) with severe COVID-19 epidemics are captured via this model.

Table 2. The regression coefficients of destination arrival service and principal components of the socioeconomic indicators in the Bayesian multivariate model.

Variables	Mean	SD	2.5%	97.5%	P ($\beta \neq 0$)
Intercept	96.07	12.13	71.69	121.48	1.00
Destination arrival service	−835.76	2532.82	−9101.57	1831.27	0.26
Resident population (PC1)	0.76	3.00	−3.84	9.78	0.24
Healthcare availability (PC2)	−34.72	8.85	−53.23	−16.92	1.00
General demographics (PC3)	−34.28	18.77	−64.21	0.00	0.88

Mean and SD indicate the average and standard deviation of the posterior distribution of each variable. P ($\beta \neq 0$) is the marginal probability that a coefficient is nonzero.

Table 3. The regression coefficients of transfer service and principal components of the socioeconomic indicators in the Bayesian multivariate model.

Variables	Mean	SD	2.5%	97.5%	P ($\beta \neq 0$)
Intercept	96.07	11.10	74.29	119.52	1.00
Transfer service	2025.33	1319.49	0.00	4160.11	0.82
Resident population (PC1)	−1.15	3.93	−13.27	5.00	0.31
Healthcare availability (PC2)	−36.71	8.54	−53.82	−20.02	1.00
General demongraphics (PC3)	−40.25	16.80	−65.75	0.00	0.94

Mean and SD indicate the average and standard deviation of the posterior distribution of each variable. P ($\beta \neq 0$) is the marginal probability that a coefficient is nonzero.

5. Discussions

High-speed trains can be regarded as a major tool used for cross-province transportation during Chunyun in mainland China. Therefore, high-speed railway schedules could capture cross-province movement patterns. With limited data sources, including COVID-19 reported cases from open data platform and daily high-speed railway schedule by the web crawler, this study compared the differences in city-level network properties (destination arrival and transfer service) before and after the Wuhan city lockdown in the early stages of the spatial transmission of COVID-19 in mainland China. Our results show that the regions with increasing transfer activities had significant numbers of confirmed infected cases, and these regions were located in provinces neighboring Hubei in the widthwise and longitudinal directions. These results indicate that transfer activities enhance the probability of interpersonal transmission and could be a crucial risk factor for increasing epidemic severity after the Wuhan city lockdown. Our study provides another possible pathway to explain why the provinces surrounded by Hubei had a higher number of confirmed COVID-19 cases than other provinces.

In pandemic transmission, considering the network connectivity of a person is crucial because it could reflect the frequency or probability of that person contacts others. Hence, several network indicators have been widely used to measure the network connectivity characteristics of each person and have revealed the associations between network connectivity measures, such as degree centrality, PageRank, and betweenness centrality, and pandemic transmission [50–54]. Degree centrality is used to measure the number of persons connected to the specific person, and so a high value indicates that this person most likely infected others because he or she can reach more people than others can. Previous studies have reported that a high level of degree centrality is positively related to disease incidence [51,55]. However, degree centrality only includes the first-degree neighbors connected to the specific person, and it cannot account for other neighbors who are not directly connected to that person. PageRank takes account not only of all the people in the network but also of the direction and weight of the connection between one person and another [53]. Betweenness centrality represents the mediation property of each person in the network. Previous studies have reported that people with high mediation have greater infection potential due to increased contact with various people [56,57]. Different

from previous studies, our results show no relationship between the changes in PageRank centrality and early COVID-19 transmission, but the changes in betweenness centrality are significantly related to the incidence of confirmed COVID-19 cases after the Wuhan city lockdown. Two factors could explain this. First, the destinations for most railway passengers were tier 1 or 2 cities (as measured by the Chinese government's official city ranking), such as Beijing, Guangdong, or Henan [28]. In addition, travelers usually targeted their hometowns as destinations during Chunyun. Hence, transfer service changes could better explain early COVID-19 transmission than destination arrival service changes. Second, we measured changes in nodal metrics; if a region was not affected by the Wuhan city lockdown, it did not show significant changes. In other words, Wuhan and other cities in Hubei provided railway passengers with a transfer hub, not a destination; therefore, the Wuhan city lockdown did not affect the destination of railway passengers but changed their itinerary routes to their destination. As a result, understanding the effect of transfer services plays an essential role in understanding early COVID-19 transmission.

The definitions of PageRank and betweenness centrality, two network indicators, were used in this study to measure the properties of destination arrival [53,58] and transfer activities [39,42,43]. Our results show that the impact of transfer activities on early COVID-19 transmission was more significant than that of destination arrival. Some regions usually have few transfers; however, those regions provided more transfer services after the Wuhan city lockdown due to the closure of the critical transfer hub of Wuhan during Chunyun. The more people gathered in these regions, the higher the contact probability with various people becomes. A similar result for airport networks has been reported in Gardner and Sarkar [59]). They addressed the fact that transfer passengers are a vital element to monitor in order to avoid disease transmission, especially for airport surveillance. Uninfected passengers could be infected because they contacted infected people during transfers at the airport [6]. Our results further illustrate that the provinces neighboring Hubei had a rising PTA after the Wuhan city lockdown, and remarkably, individuals in these provinces had a greater probability of contact with infected persons than those in provinces with a declining PTA. This indicates that provinces neighboring Hubei provided partial transfer functionality after the Wuhan city lockdown and explains why those provinces had a larger number of confirmed cases.

The demographic, economic, and healthcare dimensions have been reported to play critical roles in pandemic transmission [48,49]. Our results show that transfer activities and the total population had positive impacts on COVID-19 transmission, whereas the resident population percentage and healthcare availability had negative impacts. A large population with a low resident population suggests that the spread of the infection might be driven by people who work in other regions. This implies that most of the workers and students returned to their hometowns during Chunyun, and the number of people and their contact probability in these regions quickly increased. Those provinces with a low GDP per capita, a low level of average birth insurance expenditure, and life expectancy have a higher number of infections, indicating that a lack of healthcare resources increased the incidence of confirmed cases [60]. Moreover, a high total household population might indicate high contact probabilities, while low average local government expenditures might indicate that more time is required to respond to the pandemic [61]. In addition to transfer activity, we further reveal that other risk factors in early COVID-19 transmission, such as a high total household population, low resident population, low GDP, low birth insurance expenditure, low average local government expenditure on healthcare, and short life expectancy, might increase pandemic transmission.

This study has several limitations. First, city-level socioeconomic indicators were not incorporated in our regression models. Thus, the characteristics of local transmission within a city could be overlooked in our study. The findings of this study reflect the impact of high-speed railway transport on cross-province COVID-19 transmission. Second, due to a lack of symptom onset data for each confirmed case, we estimated the symptom onset of cases from the published literature. This could lead to a biased temporal trend in estimates regarding

early transmission in mainland China. Third, trains usually carry many passengers during Chunyun; therefore, the number of trains scheduled between two cities was used to represent population flows. It may not reflect the actual volume of passengers. Fourth, not only is high-speed rail transport important but air and road transport are also important cross-province transportation tools during Chunyun. Further investigation is warranted to incorporate more transport modes. Fifth, city-level transfer services were conducted through city-to-city network structures. These network indicators may not capture the actual behaviors of individual railway passengers. Sixth, although this study demonstrates the province-level association between transfer services and the COVID-19 epidemic in the early stages of transmission, it cannot infer individual-level infection risks from transfer behaviors. Last but not least, the spatial heterogeneity of population flow plays an important role in the geographical process of epidemic transmission. Therefore, it could be warranted to use geographically weighted regression to explore spatial heterogeneity of the COVID-19 epidemic in future studies.

6. Conclusions

The impact of the Wuhan city lockdown on railway transportation was measured by destination arrival and transfer activities using city-to-city network metrics, including the PageRank and betweenness centrality scores. Our results show that the provinces with rising transfer activities after the Wuhan city lockdown had more confirmed COVID-19 cases, but changes in destination arrival did not have significant effects. This implies that the destinations of railway passengers might not be affected by the Wuhan city lockdown, but their itinerary routes could be changed due to the replacement of an important transfer hub (Wuhan city) in the Chinese railway transportation network. We conclude that transfer services in the high-speed rail network could be another possible explanation for why the provinces surrounded by Hubei had a higher number of confirmed COVID-19 cases than other provinces.

Author Contributions: Conceptualization, T.-H.W.; methodology, C.-H.C. and T.-H.W.; software, C.-H.C.; validation, C.-H.C. and T.-H.W.; formal analysis, C.-H.C.; investigation, C.-H.C.; resources, C.-H.C. and T.-H.W.; data curation, C.-H.C.; writing—original draft preparation, C.-H.C.; writing—review and editing, C.-H.C. and T.-H.W.; visualization, C.-H.C.; supervision, T.-H.W.; project administration, T.-H.W.; funding acquisition, C.-H.C. and T.-H.W. Both authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by grants from the Ministry of Science and Technology (MOST) of Taiwan (MOST 108-2420-H-002-025-DR and MOST 108-2638-H-002-002-MY2) and National Health Research Institutes (NHRI-108A1-MRCO-0619191) in Taiwan. The authors also acknowledge the financial support provided by the Infectious Diseases Research and Education Center, Ministry of Health and Welfare (MOHW), and National Taiwan University (NTU). The funder had no role in the study design, data collection, analysis, or preparation of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

COVID-19: Coronavirus disease 2019; OD: Origin-destination; DAP: destination arrival probability; PTA: potential for transfer activity; PR: PageRank centrality; BC: Betweenness centrality; kNN: k-nearest neighbor; GDP: gross domestic product; PCA: principal component analysis.

References

1. Jones, K.E.; Patel, N.G.; Levy, M.A.; Storeygard, A.; Balk, D.; Gittleman, J.L.; Daszak, P. Global trends in emerging infectious diseases. *Nature* **2008**, *451*, 990–993. [[CrossRef](#)] [[PubMed](#)]
2. Morens, D.M.; Folkers, G.K.; Fauci, A.S. The challenge of emerging and re-emerging infectious diseases. *Nature* **2004**, *430*, 242–249. [[CrossRef](#)] [[PubMed](#)]
3. Brockmann, D.; Helbing, D. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science* **2013**, *342*, 1337–1342. [[CrossRef](#)] [[PubMed](#)]
4. Viboud, C.; Bjørnstad, O.N.; Smith, D.L.; Simonsen, L.; Miller, M.A.; Grenfell, B.T. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. *Science* **2006**, *312*, 447–451. [[CrossRef](#)]
5. Grenfell, B.T.; Bjørnstad, O.N.; Kappey, J. Travelling waves and spatial hierarchies in measles epidemics. *Nature* **2001**, *414*, 716–723. [[CrossRef](#)]
6. Cai, J.; Xu, B.; Chan, K.K.Y.; Zhang, X.; Zhang, B.; Chen, Z.; Xu, B. Roles of Different Transport Modes in the Spatial Spread of the 2009 Influenza A(H1N1) Pandemic in Mainland China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 222. [[CrossRef](#)] [[PubMed](#)]
7. St John, R.K.; King, A.; de Jong, D.; Bodie-Collins, M.; Squires, S.G.; Tam, T.W.S. Border screening for SARS. *Emerg. Infect. Dis.* **2005**, *11*, 6–10. [[CrossRef](#)] [[PubMed](#)]
8. Hsieh, Y.-H.; King, C.-C.; Chen, C.W.S.; Ho, M.-S.; Hsu, S.-B.; Wu, Y.-C. Impact of quarantine on the 2003 SARS outbreak: A retrospective modeling study. *J. Theor. Biol.* **2007**, *244*, 729–736. [[CrossRef](#)] [[PubMed](#)]
9. Baker, M.G.; Kelly, H.; Wilson, N. Pandemic H1N1 influenza lessons from the southern hemisphere. *Eurosurveillance* **2009**, *14*, 19370. [[CrossRef](#)]
10. Fielding, J.E.; Higgins, N.; Gregory, J.E.; Grant, K.A.; Catton, M.G.; Bergeri, I.; Lester, R.A.; Kelly, H.A. Pandemic H1N1 influenza surveillance in Victoria, Australia, April–September, 2009. *Eurosurveillance* **2009**, *14*, 19368. [[CrossRef](#)]
11. Cohen, N.J. Travel and border health measures to prevent the international spread of Ebola. *MMWR Suppl.* **2016**, *65*, 57–67. [[CrossRef](#)] [[PubMed](#)]
12. Plipat, T.; Buathong, R.; Wacharapluesadee, S.; Siriarayapon, P.; Pittayawonganon, C.; Sangsajja, C.; Kaewpom, T.; Petcharat, S.; Ponpinit, T.; Jumpasri, J.; et al. Imported case of Middle East respiratory syndrome coronavirus (MERS-CoV) infection from Oman to Thailand, June 2015. *Eurosurveillance* **2017**, *22*, 30598. [[CrossRef](#)] [[PubMed](#)]
13. Balcan, D.; Hu, H.; Goncalves, B.; Bajardi, P.; Poletto, C.; Ramasco, J.J.; Paolotti, D.; Perra, N.; Tizzoni, M.; Van den Broeck, W.; et al. Seasonal transmission potential and activity peaks of the new influenza A(H1N1): A Monte Carlo likelihood analysis based on human mobility. *BMC Med.* **2009**, *7*, 45. [[CrossRef](#)]
14. Kelly, H.A.; Priest, P.C.; Mercer, G.N.; Dowse, G.K. We should not be complacent about our population-based public health response to the first influenza pandemic of the 21st century. *BMC Public Health* **2011**, *11*, 78. [[CrossRef](#)] [[PubMed](#)]
15. Selvey, L.A.; Antão, C.; Hall, R. Evaluation of border entry screening for infectious diseases in humans. *Emerg. Infect. Dis.* **2015**, *21*, 197–201. [[CrossRef](#)]
16. Mukherjee, P.; Lim, P.L.; Chow, A.; Barkham, T.; Seow, E.; Win, M.K.; Chua, A.; Leo, Y.S.; Cheng Chen, M.I. Epidemiology of travel-associated pandemic (H1N1) 2009 infection in 116 patients, Singapore. *Emerg. Infect. Dis.* **2010**, *16*, 21–26. [[CrossRef](#)]
17. Sakaguchi, H.; Tsunoda, M.; Wada, K.; Ohta, H.; Kawashima, M.; Yoshino, Y.; Aizawa, Y. Assessment of Border Control Measures and Community Containment Measures Used in Japan during the Early Stages of Pandemic (H1N1) 2009. *PLoS ONE* **2012**, *7*, e31289. [[CrossRef](#)]
18. Samaan, G.; Spencer, J.; Roberts, L.; Patel, M. Border screening for SARS in Australia: What has been learnt? *Med. J. Aust.* **2004**, *180*, 220–223. [[CrossRef](#)]
19. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.M.; Lau, E.H.Y.; Wong, J.Y.; et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [[CrossRef](#)]
20. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727–733. [[CrossRef](#)]
21. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565–574. [[CrossRef](#)]
22. Guan, W.-J.; Ni, Z.-Y.; Hu, Y.; Liang, W.-H.; Ou, C.-Q.; He, J.-X.; Liu, L.; Shan, H.; Lei, C.-L.; Hui, D.S.C.; et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)]
23. Chen, S.; Yang, J.; Yang, W.; Wang, C.; Bärnighausen, T. COVID-19 control in China during mass population movements at New Year. *Lancet* **2020**, *395*, 764–766. [[CrossRef](#)]
24. Tian, H.; Liu, Y.; Li, Y.; Wu, C.-H.; Chen, B.; Kraemer, M.U.G.; Li, B.; Cai, J.; Xu, B.; Yang, Q.; et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **2020**, *368*, eabb6105. [[CrossRef](#)] [[PubMed](#)]
25. Kraemer, M.U.G.; Yang, C.-H.; Gutierrez, B.; Wu, C.-H.; Klein, B.; Pigott, D.M.; du Plessis, L.; Faria, N.R.; Li, R.; Hanage, W.P.; et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **2020**, *368*, 493. [[CrossRef](#)]
26. Chen, X. A Sustainability Analysis on the Wuhan-Guangzhou High-Speed Railway in China. *Int. J. Sustain. Transp.* **2015**, *9*, 348–363. [[CrossRef](#)]
27. Dai, G. The impact of policy networks on the urbanisation around High-Speed Railway stations in China: The case of Wuhan. *Environ. Plan. C Gov. Policy* **2015**, *33*, 533–551. [[CrossRef](#)]

28. Chinazzi, M.; Davis, J.T.; Ajelli, M.; Gioannini, C.; Litvinova, M.; Merler, S.; Pastore y Piontti, A.; Mu, K.; Rossi, L.; Sun, K.; et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **2020**, *368*, 395–400. [CrossRef]
29. Wu, J.T.; Leung, K.; Leung, G.M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *Lancet* **2020**, *395*, 689–697. [CrossRef]
30. Wells, C.R.; Sah, P.; Moghadas, S.M.; Pandey, A.; Shoukat, A.; Wang, Y.; Wang, Z.; Meyers, L.A.; Singer, B.H.; Galvani, A.P. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 7504–7509. [CrossRef]
31. Min, W.F.; Huizhi, G.; Jessica, Y.W.; Jingyi, X.; Eunice, Y.C.S.; Sukhyun, R.; Benjamin, J.C. Nonpharmaceutical Measures for Pandemic Influenza in Nonhealthcare Settings—Social Distancing Measures. *Emerg. Infect. Dis. J.* **2020**, *26*, 976.
32. Colizza, V.; Barrat, A.; Barthelemy, M.; Valleron, A.-J.; Vespignani, A. Modeling the Worldwide Spread of Pandemic Influenza: Baseline Case and Containment Interventions. *PLoS Med.* **2007**, *4*, e13. [CrossRef]
33. Cooper, B.S.; Pitman, R.J.; Edmunds, W.J.; Gay, N.J. Delaying the International Spread of Pandemic Influenza. *PLoS Med.* **2006**, *3*, e212. [CrossRef]
34. Ferguson, N.M.; Cummings, D.A.T.; Fraser, C.; Cajka, J.C.; Cooley, P.C.; Burke, D.S. Strategies for mitigating an influenza pandemic. *Nature* **2006**, *442*, 448–452. [CrossRef] [PubMed]
35. Wu, Z.; McGoogan, J.M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72/314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [CrossRef] [PubMed]
36. CCDC China COVID-19 Weekly Report. Available online: <http://weekly.chinacdc.cn/> (accessed on 6 October 2020).
37. Brin, S.; Page, L. The anatomy of a large-scale hypertextual Web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [CrossRef]
38. Page, L.; Brin, S.; Motwani, R.; Winograd, T. *The PageRank Citation Ranking: Bringing Order to the Web*; Stanford InfoLab, Stanford University: Stanford, CA, USA, 1999.
39. Xu, J.; Li, A.; Li, D.; Liu, Y.; Du, Y.; Pei, T.; Ma, T.; Zhou, C. Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Appl. Geogr.* **2017**, *87*, 85–96. [CrossRef]
40. Freeman, L.C. A Set of Measures of Centrality Based on Betweenness. *Sociometry* **1977**, *40*, 35–41. [CrossRef]
41. Anthonisse, J.M. The rush in a directed graph. In *Stichting Mathematisch Centrum. Mathematische Besliskunde*; BN 9/71; SMC: Amsterdam, The Netherlands, 1971.
42. Wang, J.; Mo, H.; Wang, F.; Jin, F. Exploring the network structure and nodal centrality of China’s air transport network: A complex network approach. *J. Transp. Geogr.* **2011**, *19*, 712–721. [CrossRef]
43. Cats, O. Topological evolution of a metropolitan rail transport network: The case of Stockholm. *J. Transp. Geogr.* **2017**, *62*, 172–183. [CrossRef]
44. Dixon, P.M. Ripley’s K Function. In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons: Chichester, UK, 2014.
45. Yang, Y. An Evaluation of Statistical Approaches to Text Categorization. *Inf. Retr.* **1999**, *1*, 69–90. [CrossRef]
46. Yang, Y.; Liu, X. A re-examination of text categorization methods. In Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Berkeley, CA, USA, 15 August 1999; pp. 42–49.
47. NBS. *The Sixth National Population Census of China in 2010*; National Bureau of Statistics of China: Beijing, China, 2010.
48. GLEAM GLEAM Project. Available online: <https://www.gleamproject.org/> (accessed on 22 August 2020).
49. EpiRisk EpiRisk. Available online: <https://epirisk.net/> (accessed on 22 August 2020).
50. Saraswathi, S.; Mukhopadhyay, A.; Shah, H.; Ranganath, T.S. Social Network Analysis of COVID-19 Transmission in Karnataka, India. *Epidemiol. Infect.* **2020**, *148*, E230. [CrossRef] [PubMed]
51. Christakis, N.A.; Fowler, J.H. Social Network Sensors for Early Detection of Contagious Outbreaks. *PLoS ONE* **2010**, *5*, e12948. [CrossRef]
52. Yang, C.H.; Jung, H. Topological dynamics of the 2015 South Korea MERS-CoV spread-on-contact networks. *Sci. Rep.* **2020**, *10*, 4327. [CrossRef]
53. Mata, A.S.D. Complex Networks: A Mini-review. *Braz. J. Phys.* **2020**, *50*, 658–672. [CrossRef]
54. Azad, S.; Devi, S. Tracking the spread of COVID-19 in India via social networks in the early phase of the pandemic. *J. Travel Med.* **2020**, *27*, taaa130. [CrossRef]
55. Kitsak, M.; Gallos, L.K.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H.E.; Makse, H.A. Identification of influential spreaders in complex networks. *Nat. Phys.* **2010**, *6*, 888–893. [CrossRef]
56. Marcelino, J.; Kaiser, M. Critical paths in a metapopulation model of H1N1: Efficiently delaying influenza spreading through flight cancellation. *PLoS Curr.* **2012**, *4*, e4f8c9a2e1fca8. [CrossRef]
57. Wang, Y.; Wu, W.; Cheng, Z.; Tan, X.; Yang, Z.; Zeng, X.; Mei, B.; Ni, Z.; Wang, X. Super-factors associated with transmission of occupational COVID-19 infection among healthcare staff in Wuhan, China. *J. Hosp. Infect.* **2020**, *106*, 25–34. [CrossRef] [PubMed]
58. Jiang, B.; Zhao, S.; Yin, J. Self-organized natural roads for predicting traffic flow: A sensitivity study. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P07008. [CrossRef]

-
59. Gardner, L.; Sarkar, S. A Global Airport-Based Risk Model for the Spread of Dengue Infection via the Air Transport Network. *PLoS ONE* **2013**, *8*, e72129.
 60. Hosseini, P.; Sokolow, S.H.; Vandegrift, K.J.; Kilpatrick, A.M.; Daszak, P. Predictive Power of Air Travel and Socio-Economic Data for Early Pandemic Spread. *PLoS ONE* **2010**, *5*, e12763. [[CrossRef](#)] [[PubMed](#)]
 61. Cox, N.J.; Subbarao, K. Global Epidemiology of Influenza: Past and Present. *Annu. Rev. Med.* **2000**, *51*, 407–421. [[CrossRef](#)] [[PubMed](#)]