Travel Medicine and Infectious Disease 42 (2021) 102097

Contents lists available at ScienceDirect

Travel Medicine and Infectious Disease

journal homepage: www.elsevier.com/locate/tmaid

Original article

The role of high-speed rail and air travel in the spread of COVID-19 in China

Pengyu Zhu *, Yuqing Guo

Hong Kong University of Science and Technology, Hong Kong, China

ARTICLE INFO

Keywords:
High-speed railway
HSR
Air travel
Covid-19
Infectious diseases
Transport suspension

ABSTRACT

Background: Public transportation is a major facilitator of the spread of infectious diseases and has been a focus of policy interventions aiming to suppress the current COVID-19 epidemic.
Methods: We use a random-effects panel data model and a Difference-in-Differences in Reverse (DDR) model to examine how air and rail transport links with Wuhan as well as the suspension of these transport links influenced the development of the epidemic in China.
Results: We find high-speed rail (HSR) and air connectivity with Wuhan resulted in 25.4% and 21.2% increases in the average number of daily new confirmed cases, respectively, while their suspension led to 18.6% and 13.3% decreases in that number. We also find that the suspension effect was dynamic, growing stronger over time and peaking 20–23 days after the Wuhan lockdown, then gradually wearing off. It took approximately four weeks for this effect to fully materialize, roughly twice the maximum incubation period, and similar dynamic patterns were seen in both HSR and air models. Overall, HSR had a greater impact on COVID-19 development than air transport.
Conclusions: Our research provides important evidence for implementing transportation-related policies in controlling future infectious diseases.

1. Introduction

At the end of 2019, the first confirmed case of COVID-19 was reported in Wuhan, Hubei. Case numbers soon exploded due to the strong infectivity of this new coronavirus and the extensive travel by road, rail and air in China during the Spring Festival travel rush. Wuhan is located in central China and is a major transportation hub connecting many cities. These factors were highly conducive to the spread of the virus. To curb the spread, Wuhan and other cities in Hubei implemented a complete lockdown on January 23, shutting down all inter-city transport.

Transportation networks accelerate the spread of viruses and disease [1–3]. Existing studies on the role of transportation networks in China's COVID-19 outbreak fall into two broad categories: those that focus on the role of transportation in the domestic spread of COVID-19 before the Wuhan lockdown and those that examine the effect of lockdown policies and travel restrictions. Studies of the first kind have suggested that high-speed rail (HSR) played an important role in the spread of the virus. A preliminary study by Zhao et al., based on data from six cities, found that HSR connectivity is significantly and positively correlated with the spread of COVID-19, while no significant correlation was found for air and road connectivity [4]. The significance of HSR connectivity was corroborated by two recent studies. One study found that the connection of HSR with Wuhan led to a significantly higher total number of cumulative confirmed cases by February 6 [5]. Another study using a gravity model also found that HSR and flight frequencies from Wuhan had a significant impact on the total number of cumulative cases by February 15th in each city [6]. Meanwhile, research on the impact of lockdown measures and travel restrictions has found that such measures were effective in hindering the spread of the virus [7,8]. However, there are currently no studies that look comprehensively at the role of transportation in the development of COVID-19 in China. How did connection to the epicenter of the outbreak via different transport networks speed up early-stage transmission? How did prevention and control measures such as suspending transportation networks help curb the spread of the virus? This paper examines the role of HSR and air travel in virus transmission from both of these angles.

To evaluate the effect of the operation and suspension of transport links with Wuhan on virus transmission, we apply a random effects

* Corresponding author.
E-mail address: pengyuzhu@ust.hk (P. Zhu).

https://doi.org/10.1016/j.tmaid.2021.102097
Received 18 December 2020; Received in revised form 27 May 2021; Accepted 28 May 2021
Available online 31 May 2021

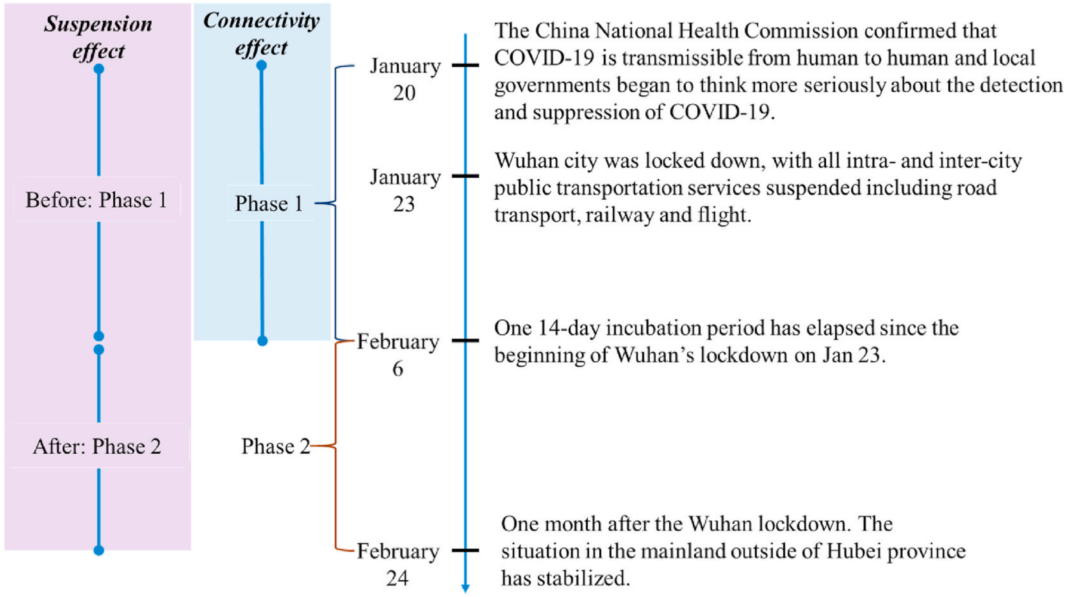


Fig. 1. Timeline indicating important events and phases of the study.

panel data model and a Difference-in-Differences in Reverse (DDR) model [9], respectively. Two major government actions demarcate different sub-phases of our study. First, on January 20, the China National Health Commission confirmed the human-to-human transmissibility of the novel coronavirus, and local governments began to take action to detect and suppress the virus. Second, on January 23, Wuhan implemented a complete lockdown; all inter-city transportation services, including flights, rail services and bus routes, were suspended, and all major freeway checkpoints leading out of the city were shut down. Based on these two government actions, we delineate two main phases (and a pre-phase) and estimate the impact of HSR and air travel on the development of the epidemic in each phase. The pre-phase refers to the period from January 11 to January 19, the day before the government announcement of human-to-human transmissibility. During this phase, efforts to detect COVID-19 outside of Hubei province had not yet begun; therefore, the HSR and air connectivity effects are not measurable. We designate Phase One as the period from January 20 to February 6. In this period, the connectivity effects are measurable (widespread virus testing in different cities began on January 20) and case numbers potentially include cases directly imported from Wuhan. While travel between Wuhan and other cities ceased on January 23, given the 14-day incubation period of COVID-19, case numbers up until February 6 may still include directly imported cases. We designate February 7 as the beginning of Phase Two. At this point, case numbers should no longer reflect cases directly imported from Wuhan and the effect of suspending HSR and flight links with Wuhan can be fully measured. We choose February 24, one month after the lockdown of Wuhan, as the end of our study period. By this point, the situation in the mainland, outside of Hubei province, had stabilized [10]. Fig. 1 presents a timeline of the relevant events and phases of the study.

2. Materials and methods

2.1. Variables

This study draws on data from 271 Chinese prefecture-level cities. The sample consists of 13,550 observations of 271 cities, covering the 50 days from January 11 to February 29. As we study the effects of both HSR and air travel, there are two pairs of treatment and control groups. For HSR, a city is defined as treated if there are direct G-series or D-series trains[1] running between the city and Wuhan; cities that require transfers to reach Wuhan or those without HSR stations are defined as 'untreated' and belong to the control group. For air transport, a city is defined as treated if there are direct flights between the city and Wuhan; those that require transfers or do not have airports are defined as the control group. The HSR treatment group consists of 117 cities and 5850 observations, while the air transport treatment group consists of 56 cities and 2800 observations.

Variables and data sources. The dependent variable in our models is the number of daily new confirmed cases in each city. The data were gathered from the Chinese Center for Disease Control and Prevention (China CDC). Explanatory variables include city-specific variables and city-time variant variables. City-specific variables include the HSR connectivity dummy, air connectivity dummy, highway distance from Wuhan and socio-economic indexes. The data come from a wide range of sources, including ticket booking websites, Baidu Map, the 2019 China City Yearbook and local government websites. We use GDP per capita and population density for 2018 as the measures of the socio-economic situation in different cities. City-time variant variables include population mobility, the weather situation and whether local control policies were implemented. Population mobility indexes are drawn from Baidu mobility data and include in-migration (IMI) and out-migration (OMI) indexes. We also control for the weather situation, including cities' Air Quality Index, relative humidity and wind speed. The presence of

[1] G-series trains represent high-speed railway service with a maximum speed of 350 km/h. D-series trains are the second fastest trains in China with speeds of up to 250 km/h and some of them can reach the speed of over 300 km/h. Together with the C-series train, these three kinds of trains all belong to China Railway High-speed (CRH) services. However, there is no C-series train connecting Wuhan and cities out of Hubei province, this paper will only consider G- and D-series trains.

Table 1
Summary statistics for the continuous control variables.

	Obs.	Mean (SD)							
		MoveIn (unit)	MoveOut (unit)	AQI (unit)	RHU1 (%)	WINms (m/s)	distance (kilometer)	city_CapGDP (100^2 Yuan)	city_den (persons/sqkm)
Full sample	13,550	0.777 (1.071)	0.786 (1.619)	74.302 (50.94)	70.465 (15.85)	2.132 (0.787)	1151.601 (673.961)	6.007 (3.49)	781.495 (659.5813)
HSR-Treated	5850	1.079 (1.322)	1.136 (2.201)	73.531 (53.06)	74.014 (14.31)	2.121 (0.811)	848.935 (406.121)	7.258 (3.805)	972.837 (643.079)
HSR-Control	7700	0.548 (0.754)	0.520 (0.876)	74.888 (49.26)	67.768 (16.42)	2.14 (0.768)	1381.548 (742.677)	5.057 (2.888)	636.124 (634.414)
Flight-Treated	2800	1.196 (1.424)	1.348 (2.568)	68.815 (52.22)	72.122 (14.78)	2.23 (0.849)	1165.409 (487.005)	8.606 (3.822)	893.343 (628.661)
Flight-Control	10,750	0.668 (0.928)	0.639 (1.217)	75.731 (50.5)	70.033 (16.09)	2.106 (0.768)	1148.004 (714.648)	5.330 (3.054)	752.362 (664.344)

lockdown measures or transport restrictions is represented by a dummy variable indicating if a city had implemented local control policies[2] at date t. Information about local COVID-19 control policies is drawn from two relevant studies [8,11] and was verified against announcements from government websites. Our sample includes 94 cities that imposed local epidemic suppression policies (see Appendix for the list of cities). Table 1 presents the summary statistics for the explanatory variables.

2.2. Model specifications

2.2.1. Connectivity effect

In this paper, we quantify the effects of HSR and air connectivity on COVID-19 transmission in China by comparing the average daily new confirmed case numbers of the treated and control group cities between January 20 and February 6. We apply a random-effects panel data model as follows:

$$\ln\left(\mathrm{NewCase}_{i,t}\right) = \beta_0 + \beta_1 \mathrm{Treated}_{\mathrm{HSR}_i} + \beta_2 \mathrm{Treated}_{\mathrm{Flight}_i} + \beta_3 \mathrm{Distance}_i + \beta_4 \sum_{k=0}^{19} \mathrm{RatioMoveIn}_{i,t-k} + \beta_5 \mathrm{Weather}_{i,t} + \beta_6 \mathrm{SocioEcon}_i + v_i + \theta_t + \varepsilon_{it} \tag{1}$$

where i denotes the city and t denotes the date. $\ln(\mathrm{NewCase}_{i,t})$ is the logarithm of one plus the number of new confirmed cases in city i at date t. $\mathrm{Treated}_{\mathrm{HSR}_i}$ is a dummy variable equal to one if there is direct HSR service between city i and Wuhan during the sampling period. $\mathrm{Treated}_{\mathrm{Flight}_i}$ is a dummy variable equal to one if there is direct flight service between city i and Wuhan during the sampling period. $\mathrm{Distance}_i$ is the highway distance between Wuhan and city i and serves as a proxy variable for road transportation. $\mathrm{RatioMoveIn}_{i,t-k}$ represents the ratio of daily in-migration into city i over daily out-migration from city i k days before date t. The model also includes nineteen lag variables for the IMI-OMI ratio as some studies observe that population inflow from Wuhan 19 days before the analyzed date had a statistically significant impact on daily new confirmed case numbers [8]. $\mathrm{Weather}_{i,t}$ is a list of control variables for weather situations, including the Air Quality Index, relative humidity and wind speed of city i at date t. $\mathrm{SocioEcon}_i$ includes two major social-economic indexes: GDP per capita and population density. Variance inflation factors show that there is no multi-collinearity among the explanatory variables. v_i captures the city random effect, which follows a Normal distribution $v_i \sim N(0, \sigma_v^2)$. Random effects panel data model assumes that v_i is uncorrelated with the observed explanatory variables [12,13].θ_t captures all date-specific fixed effects. Finally, ε_{it} is the error term. The coefficient β_1 measures the HSR connectivity effect with other variables controlled, while β_2 measures the air connectivity effect. These two coefficients are the focus of the first part of the analysis. To check for robustness, we add a dummy variable indicating the implementation of local control policies six days before date t in city i to capture the suppression effect of these policies. We choose six days as the lag because the average incubation period for COVID-19 is 5–6 days [14]. Thus, we have Eq. (2):

RE models assume that the observed predictors in the model are

$$\ln\left(\mathrm{NewCase}_{i,t}\right) = \beta_0 + \beta_1 \mathrm{Treated}_{\mathrm{HSR}_i} + \beta_2 \mathrm{Treated}_{\mathrm{Flight}_i} + \beta_3 \mathrm{Distance}_i + \beta_4 \sum_{k=0}^{19} \mathrm{RatioMoveIn}_{i,t-k} + \beta_5 \mathrm{Weather}_{i,t} + \beta_6 \mathrm{SocioEcon}_i + \beta_7 \mathrm{Restriction}_{i,t-6} + v_i + \theta_t + \varepsilon_{it} \tag{2}$$

To examine the variation in the HSR connectivity effect on cities of different distances from Wuhan, we replace the HSR connectivity variable with the interaction terms between this variable and the three distance dummies: D1 for a short distance of up to 500 km, D2 for a medium distance of 501–1000 km, and D3 for a long distance of over 1000 km, all measured in great-circle distance from Wuhan.

$$\ln\left(\mathrm{NewCase}_{i,t}\right) = \beta_0 + \beta_1 (D1 \times \mathrm{Treated}_{\mathrm{HSR}_i}) + \beta_2 (D2 \times \mathrm{Treated}_{\mathrm{HSR}_i}) + \beta_3 (D3 \times \mathrm{Treated}_{\mathrm{HSR}_i}) + \beta_4 \mathrm{Treated}_{\mathrm{Flight}_i} + \beta_5 \mathrm{Distance}_i + \beta_6 \sum_{k=0}^{19} \mathrm{RatioMoveIn}_{i,t-k} + \beta_7 \mathrm{Weather}_{i,t} + \beta_8 \mathrm{SocioEcon}_i + \beta_9 \mathrm{Restriction}_{i,t} + v_i + \theta_t + \varepsilon_{it} \tag{2.1}$$

2.2.2. Suspension effect

To measure the effect of suspending HSR and flight service with Wuhan, we use difference-in-differences in reverse (DDR) estimation and the dataset from January 20 to February 24. We estimate the following specifications for HSR and air travel.

Eq. (3) captures how daily new confirmed case numbers responded to the suspension of HSR service, controlling for the effects of air and

[2] Local control policies might include (1) the shut-down of public transportation and prohibition of private cars; (2) or the set-up of checkpoints and quarantine zone; (3) or the ban on any gathering by residents; (4) or restriction of commercial activities or any combination of these measures.

Table 2
Model results for HSR and flight connectivity effects.

Lncase	Model (1)	Model (2)
Connected via HSR before January 23 (dummy variable)	0.227*** (0.06)	0.227*** (0.06)
Connected via flight routes before January 23 (dummy variable)	0.192** (0.09)	0.192** (0.09)
Ratio of move-in index over move-out index	0.094** (0.04)	0.094** (0.04)
Highway distance to Wuhan	−0.00018*** (0.00005)	−0.00018*** (0.00005)
Air quality index	−0.00033* (0.00018)	−0.00033* (0.00018)
Relative humidity	−0.002** (0.00)	−0.002** (0.00)
Wind speed	0.011 (0.02)	0.011 (0.02)
City GDP per capita	0.050*** (0.01)	0.050*** (0.01)
City population density	0.00015*** (0.00005)	0.00015*** (0.00005)
Presence of local pandemic control and restriction policies six days earlier (dummy variable)		0.038 (0.08)
Time fixed effect	Yes	Yes
City fixed effect	No	No
Intercept	−0.510** (0.23)	−0.510** (0.23)
N	4878	4878
R-sq	0.32	0.32

Notes: Standard errors are clustered at city level and shown in parentheses.
Date dummies and the nineteen lag variables of population inflow ratio are not presented to save space.
*p $<$ 0.05; **p $<$ 0.01; ***p $<$ 0.001.

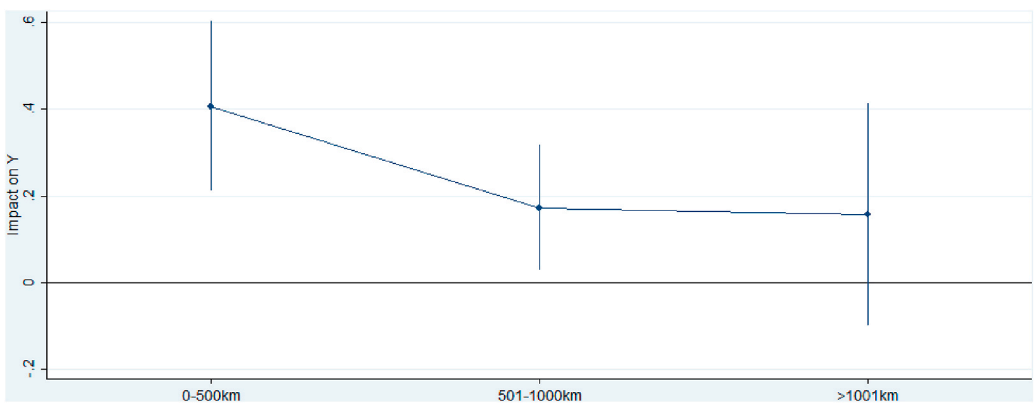


Fig. 2. Impact of HSR connectivity over different distances.

road transportation (e.g., cars and buses):

$$\ln(\mathrm{NewCase}_{i,t}) = \beta_0 + \beta_1 \mathrm{HSR}_i + \beta_2 \mathrm{Post}_t + \beta_3 \mathrm{DDR\_HSR}_{i,t} + \beta_4 \mathrm{Flight}_i + \beta_5 \mathrm{Distance}_i + \sum_{k=0}^{19} \beta_{6k} \mathrm{RatioMoveIn}_{i,t-k} + \beta_7 \mathrm{Weather}_{i,t} + \beta_8 \mathrm{SocioEcon}_i + \beta_9 \mathrm{Restriction}_{i,t} + \varepsilon_{it} \tag{3}$$

where HSR_i is a dummy variable indicating if city i is directly connected with Wuhan by HSR. The intervention variable, $\mathrm{DDR\_HSR}_{i,t}$, is coded as follows: 1) for cities connected with Wuhan before lockdown via HSR (or air), i.e. the treatment group, it takes the value of 0 during the pre-intervention period and 1 during the post-intervention period; 2) for cities not connected with Wuhan via HSR (or air) before lockdown, i.e. the control group, the intervention variable is always coded as 1 during the entire sampling period as these cities share the same status as the treatment group after the intervention. Note that the intervention date for our DDR model is set as February 7, the 15th day after Wuhan's lockdown on Jan 23. Given the 14-day incubation period of COVID-19, all cases directly associated with travel from Wuhan should have been exposed by February 6. The coefficient $\beta_3$measures the average effect of the suspension of HSR service on daily new confirmed case numbers and is the focus of the second part of the analysis. Flight_i is a dummy variable equal to one if city i is connected with Wuhan via direct flight routes, which controls for the effect of air connectivity on the dependent variable during the sampling period. As in the connectivity model, the highway distance, IMI-OMI ratio, daily weather situation, social-economic indexes and implementation of local control policies are also included in the suspension model as control variables. θ_t captures all date-specific fixed effects, and ε_{it} is the error term.

Eq. (4) is adopted from Eq. (3) with the HSR treatment and intervention variables replaced by the corresponding flight variables and the flight control variable replaced by the HSR counterpart. The equation analyzes the effect of flight suspension on the spread of COVID-19, controlling for the effects of HSR and road transport (e.g., cars and buses):

$$\ln(\mathrm{NewCase}_{i,t}) = \beta_0 + \beta_1 \mathrm{Flight}_i + \beta_2 \mathrm{Post}_t + \beta_3 \mathrm{DDR\_Flight}_{i,t} + \beta_4 \mathrm{HSR}_i + \beta_5 \mathrm{Distance}_i + \sum_{k=0}^{19} \beta_{6k} \mathrm{RatioMoveIn}_{i,t-k} + \beta_7 \mathrm{Weather}_{i,t} + \beta_8 \mathrm{SocioEcon}_i + \beta_9 \mathrm{Restriction}_{i,t} + \varepsilon_{it} \tag{4}$$

To check for robustness, we modify Eq. (3) and Eq. (4) into Eq. (5) and Eq. (6) by replacing the post variable and all time-invariant city-specific characteristic variables with date dummies θ_t and city dummies u_i:

$$\ln(\mathrm{NewCase}_{i,t}) = \beta_0 + \beta_1 \mathrm{DDR\_HSR}_{i,t} + \sum_{k=0}^{19} \beta_{2k} \mathrm{RatioMoveIn}_{i,t-k} + \beta_3 \mathrm{Weather}_{i,t} + \beta_4 \mathrm{Restriction}_{i,t} + \theta_t + u_i + \varepsilon_{it} \tag{5}$$

Table 3
Model results for HSR and flight suspension effects.

			with city- and date-fixed effects	
	HSR Model (3)	Flight Model (4)	HSR Model (5)	Flight Model (6)
HSR suspension with Wuhan (dummy variable)	−0.252*** (0.04)		−0.206*** (0.03)	
Flight suspension with Wuhan (dummy variable)		−0.256*** (0.06)		−0.143*** (0.05)
Connected via HSR before January 23 (dummy variable)	0.049 (0.04)	0.175*** (0.05)		
Connected via Flight before January 23 (dummy variable)	0.116 (0.07)	−0.011 (0.06)		
Post: February 7 - February 23 (dummy variable)	−0.136*** (0.02)	−0.180*** (0.03)		
Ratio of move-in index over move-out index	−0.022 (0.04)	−0.020 (0.04)	0.047* (0.02)	0.048* (0.02)
Highway distance to Wuhan	−0.0001*** (0.00003)	−0.0001*** (0.00003)		
Air quality index	−0.00039 (0.00030)	−0.00031 (0.00029)	−0.00002 (0.00016)	0.00007 (0.00016)
Relative humidity	0.003*** (0.00)	0.003*** (0.00)	0.002*** (0.00)	0.002*** (0.00)
Wind speed	0.001 (0.02)	0.003 (0.02)	0.026*** (0.01)	0.030*** (0.01)
City GDP per capita	0.036*** (0.01)	0.036*** (0.01)		
City population density	0.00012*** (0.00004)	0.00012*** (0.00004)		
Presence of local pandemic control and restriction policies six days earlier (dummy variable)	−0.128** (0.05)	−0.137** (0.05)	−0.134*** (0.04)	−0.153*** (0.05)
Time fixed effect	No	No	Yes	Yes
City fixed effect	No	No	Yes	Yes
Intercept	0.162 (0.11)	0.163 (0.11)	0.102 (0.07)	0.050 (0.07)
N	9756	9756	9756	9756
adj. R-sq	0.204	0.201	0.529	0.525

Notes: Standard errors are clustered at city level and shown in parentheses.
Date dummies and the nineteen lag variables of population inflow ratio are not presented to save space.
*p < 0.05; **p < 0.01; ***p < 0.001.

$$\ln\left(\mathrm{NewCase}_{i,t}\right)=\beta_0+\beta_1\left(D1\times \mathrm{DDR\_HSR}_{i,t}\right)+\beta_2\left(D2\times \mathrm{DDR\_HSR}_{i,t}\right)+\beta_3\left(D3\times \mathrm{DDR\_HSR}_{i,t}\right)+\sum_{k=0}^{19}\beta_{4k}\mathrm{RatioMoveIn}_{i,t-k}+\beta_5\mathrm{Weather}_{i,t}+\beta_6\mathrm{Restriction}_{i,t}+\theta_t+u_i+\varepsilon_{it} \tag{5.1}$$

$$\ln\left(\mathrm{NewCase}_{i,t}\right)=\beta_0+\beta_1\mathrm{DDR\_Flight}_{i,t}+\sum_{k=0}^{19}\beta_{2k}\mathrm{RatioMoveIn}_{i,t-k} + \beta_3\mathrm{Weather}_{i,t}+\beta_4\mathrm{Restriction}_{i,t}+\theta_t+u_i+\varepsilon_{it} \tag{6}$$

As we did in the HSR connectivity model, we also adjust Eq. (5) to quantify the variation in the effect of HSR suspension over different distances:

3. Results

3.1. Connectivity effect

Table 2 presents the coefficient estimates and standard errors for the two transportation variables and major control variables of interest. The nineteen lagged variables for the population inflow ratio[3] are reported in the complete regression results in the Appendix. Model (1) in Table 2 applies Eq. (1) to examine the effect of HSR and air transportation on the spread of COVID-19 using data between January 20 and February 6. Both HSR and air connectivity are shown to have a significant positive impact on the daily number of new confirmed cases. The coefficient for HSR shows that cities directly connected with Wuhan reported on average a 25.5% higher number of daily new confirmed cases than those not directly connected with Wuhan.[4] The magnitude of the impact was smaller for air transport: on average, cities directly connected with Wuhan by air reported a 21.2% higher number of daily new confirmed cases than those not directly connected with Wuhan. To check for robustness, we control for the confounding effects of local pandemic control and restriction policies[5] and modify Eq. (1) into Eq. (2) by adding a dummy variable indicating whether a city had implemented any local control measures six days earlier. Model (2) in Table 2 presents the robustness check results. The results are nearly unchanged compared to the baseline model, indicating that HSR connectivity increased the average daily number of new confirmed cases by 25.4%, while air connectivity increased case numbers by 21.2%. Thus, HSR and air transportation both had a significant effect on the COVID-19 epidemic in China, although the effect of HSR travel was slightly stronger than that of air travel.

We further test whether the effect of HSR connectivity on virus transmission varied according to the distance of cities from Wuhan. Fig. 2 illustrates how the HSR connectivity effect differed for the different distance groups: HSR connectivity had a statistically significant

[3] The models include nineteen lagged variables for the ratio of population inflow into each city, as some research suggests that population inflow from Wuhan nineteen days earlier had a significant impact on the spread of COVID-19 [8].

[4] The percentage change of daily new confirmed cases for a discrete change in any dummy variable is calculated by $100*\left(\frac{Y_1-Y_0}{Y_0}\right) = 100*(EXP(\beta)-1)$.

[5] Such policies might include (1) the shut-down of public transportation and prohibition of private cars; (2) the establishment of checkpoints and quarantine zones; (3) bans on gatherings; (4) restriction of commercial activities.

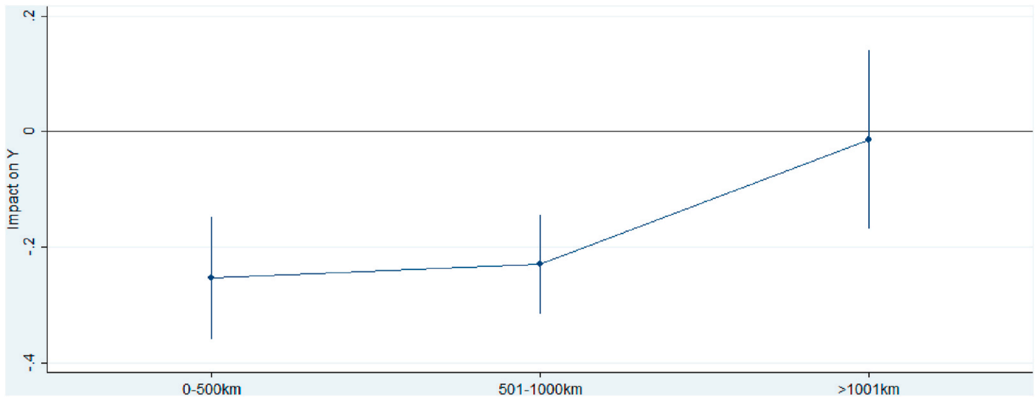


Fig. 3. Impact of HSR suspension over different distances.

impact on the development of COVID-19 in both short- (0–500 km) and medium- (501–1000 km) distance cities, but its impact on the long-distance (over 1001 km) group was not statistically significant. HSR connectivity increased average daily new confirmed cases by 50.2% for short-distance cities and 18.8% for medium-distance cities. The contribution of HSR connectivity to virus transmission thus seems to decline substantially with distance.

Among the control variables, highway distance from Wuhan also had a strong negative impact on the spread of COVID-19: an increase of 100 km in highway distance was associated with a 1.8% reduction in the average daily cases. With longer highway distances and hence longer driving times, people are increasingly likely to prefer HSR or air to bus or car transport. Therefore, highway distance is a good proxy for passenger inflows from Wuhan via inter-city highways and can be used to account for the potential confounding effects of other transportation modes (i.e. inter-city buses, private cars and taxis), ensuring our estimates of HSR and air connectivity effects are unbiased. Also of interest is the "local pandemic control restriction policies" variable. Its coefficient estimate is not statistically significant, suggesting the effects of these policies on virus transmission had not materialized during the time frame analyzed.

3.2. Suspension effect

We employ the DDR approach to model how daily new confirmed case numbers responded to the suspension of HSR and air transport with Wuhan between February 7 and February 24. The intervention date is set as February 7, the 15th day after the lockdown of Wuhan on Jan 23. Models (3) and (4) in Table 3 present the coefficient estimates for major variables of interest. Full regression results are provided in Appendix. We find that shutting down transport links indeed curbed the spread of COVID-19 out of Wuhan: suspending HSR and flight links caused a significant reduction in daily new confirmed cases. On average, suspending HSR service with Wuhan caused a 22.2% decrease in daily new confirmed cases, ceteris paribus. Suspending flight service caused a 22.6% reduction.

We conduct a robustness check by modifying Eq. (3) and Eq. (4) into Eq. (5) and Eq. (6), adding stricter controls by replacing the treatment variable, post variable and all time-invariant city-specific characteristic variables with date dummies θ_t and city dummies u_i to capture all time-specific and city-specific fixed effects. Models (5) and (6) show higher levels of statistical significance for several predictor variables and a smaller magnitude of impact for the key transportation variables relative to models (3) and (4). According to the modified model, suspension of HSR service with Wuhan reduced average daily cases by 18.6%, while suspending flight service decreased average daily cases by 13.3%.

We also examine how the HSR suspension effect varied with distance from Wuhan. Fig. 3 plots the coefficients and 95% confidence intervals for these interaction terms. The variation of the HSR suspension effect over distance demonstrates a pattern similar to the connectivity effect: suspension led to a significant drop in average daily cases in both short- and medium-distance cities, but had little impact in long-distance cities. Suspension of HSR links with Wuhan caused a 22.4% reduction in average daily cases in short-distance cities and a 20.5% reduction in medium-distance cities.

Among the control variables, the coefficient for the "local pandemic control and restriction policies" variable captures the impact of local restrictive measures. Models (3) through (6) show that control policies had a strong negative impact on the number of daily new confirmed cases, in contrast to the results of models (1) and (2). The in-migration index (IMI) to out-migration index (OMI) ratio also shows a significant positive correlation with daily new confirmed cases in models (5) and (6), after time-fixed effects are controlled for: a one-unit increase in this ratio is associated with a 5% increase in daily cases. This is however smaller than the 9% increase observed in the two connectivity effect models. Meanwhile, the coefficients for the "Highway distance to Wuhan" variable are consistent with those from Eq. (1) and Eq. (2) in terms of their signs and statistical significance, but are of much smaller magnitude too. These differences may be attributable to the difference in sampling periods of the connectivity effect and suspension effect models. The connectivity effect models draw on data from January 20 to February 6, the period just before and after Wuhan went into lockdown, when many cities were only just beginning to implement local control policies. The suppression effects of local control policies may have had a limited impact in this period, while population inflow and highway connections with Wuhan would likely have exerted a larger influence on the spread of COVID-19. The suspension effect models analyze the period from February 7 to February 24, when Wuhan's lockdown measures and local control policies had been in effect for some time. Thus, the suppression effects of local control measures would have had more time to take hold, and the effects of population inflow and highway connections with Wuhan would have been weakened.

3.3. Dynamic suspension effect

We further analyze whether suspension effects were dynamic and evolved over time. We added three lags and seven leads[6] of the intervention variable to test the lead and lag effects of the suspension of HSR or air travel. To avoid the problem of multi-collinearity and include as many variables as possible, the models take a lag length of three days instead of one day (results for lag lengths of one day and two days are

[6] During the sampling period between Jan 20th and Feb 24th, all lagged variables with a lag of 18 days or more after Feb 7th will become zero and be omitted in regression. Therefore, in order to include the extra two lagged variables (18 days and 21 days after Feb 7th), the regression models examining the dynamic effect have extended the sampling period to Feb 29th.

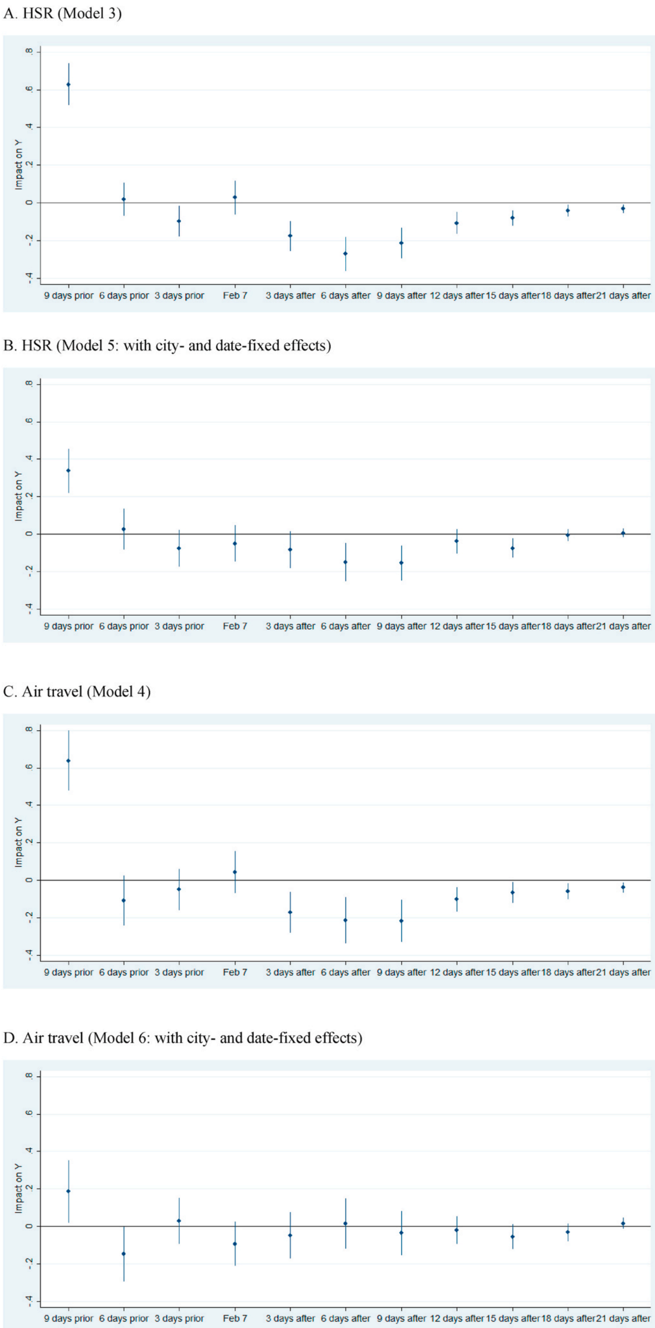


Fig. 4. Impact of transport suspension over time.

attached in Appendix for reference). Fig. 4 plots the coefficients and 95% confidence intervals for the lead and lag variables. Panel A and Panel B in Fig. 4 illustrate the dynamic impact of HSR suspension based on Eq. (3) and Eq. (5), respectively. In Panel A, all of the lag variables after the intervention have significantly negative coefficients, and their magnitude is largest between three and nine days after the intervention. In Panel B, most of the lag variables demonstrate patterns similar to Panel A, but with smaller magnitudes and statistically insignificant coefficients for the 12-, 18- and 21-day lag variables. Panel C and Panel D present the dynamic effect of air transport suspension based on Eq. (4) and Eq. (6), respectively. All coefficients for the lag variables of air transport suspension are statically significant and negative in Panel C, and the time trend is quite similar to that of HSR suspension in Panel A. However, weaker results are obtained with Eq. (5), as shown in Panel D.

Panels A and C suggest that HSR and air travel suspension had similar dynamic effects. Panel A shows that it took time for the suppression effects of suspending HSR service to manifest, and that the reduction in daily new confirmed cases was not observed until the third day after the selected intervention date of February 7 (one maximum incubation period, or 14 days, after the lockdown). This impact grew stronger over time, peaked between six and nine days after the intervention, and then gradually wore off. Panel C shows that the flight suspension effect had a similar dynamic pattern: it started to increase on the third day after the intervention and peaked between six and nine days after the intervention. Although the suspension effects for both HSR and air travel were statistically significant for some time after the peak (as shown in Panels A and C), they approached zero (numerically) on the twelfth or the fifteenth day after the intervention and remained afterwards.

3.4. *Parallel trend assumption*

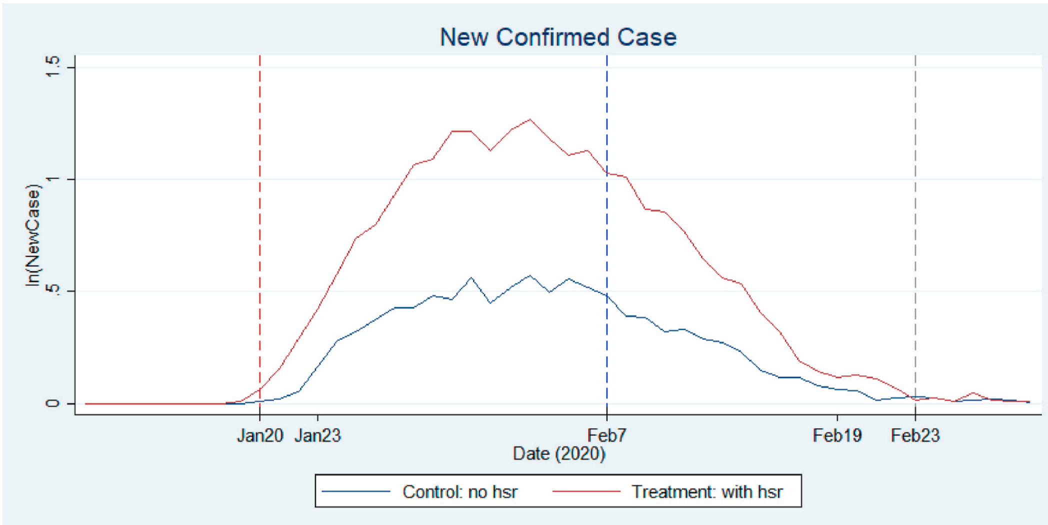
DDR estimation requires that the treatment group should develop in parallel with the control group after the effect of the intervention wears off [9]. While the coefficients for the lag variables in Fig. 4 have already indicated that the suspension effect diminishes, we also verify this parallel trend assumption visually. Fig. 5 depicts average daily cases for cities with HSR (Fig. 5A) or flight (Fig. 5B) connection to Wuhan versus those without. The red lines represent the treatment group and the blue lines represent the control group. The figures show that for both the HSR and flight models, the treatment and control groups have similar date-fixed effects pre-intervention (before January 19) and one month after the lockdown (after February 23), indicating the parallel trend assumption is satisfied. The figures also suggest that by February 23, the suspension effects had neutralized the connectivity effects and the treatment and control groups were evolving in parallel.

4. Discussion

Our study provides a comprehensive view of the relationships between COVID-19 transmission and transportation networks, focusing on both the connectivity and suspension effects of HSR and air transportation. Our findings in general resonate with similar studies on earlier outbreaks, such as H1N1 [2,15], Ebola [16], and SARS [17]. Transportation networks, including rail, air transport and buses, facilitates the virus transmission to new regions and leads to outbreaks [2,17,18]. In the case of COVID-19, our results on the connectivity/connection effects are generally consistent with previous studies concerning the impacts of HSR connection with the epicenter of the COVID-19 outbreak [4–6]. Furthermore, our analysis of flight connections adds to the discussion concerning the role of air transportation in COVID-19 spread, given that there is contradictory evidence about the relationship between aviation and COVID-19 [4,6]. More importantly, our research contributes to the literature on future novel infectious disease (NID) control by estimating how quickly suspension of transportation networks from the epicenter can help reduce COVID-19 transmission at a national scale, and how the suspension effects differ between HSR and air.

First, from our connectivity and suspension models, we can infer that HSR and air transport links both significantly influenced the development of COVID-19 in China: both HSR and air connectivity with Wuhan led to a significantly higher number of daily new confirmed cases, while suspension of HSR and flight services both led to a significant decrease in daily cases. This has implications not only for controlling the current COVID-19 outbreak, but also for controlling the spread of any emerging infectious diseases in the future. Our study provides evidence that policymakers should work closely with the transportation sector whenever there are outbreaks of severe infectious diseases, as suspension of HSR and air transport links is an effective means of suppressing disease transmission. While such measures are likely to be unwarranted except in extreme cases, policymakers can also work closely with railways, airlines and airports to monitor traffic inflows from any epicentres of

A. with HSR connectivity with Wuhan versus those without



B. with flight connectivity with Wuhan versus those without

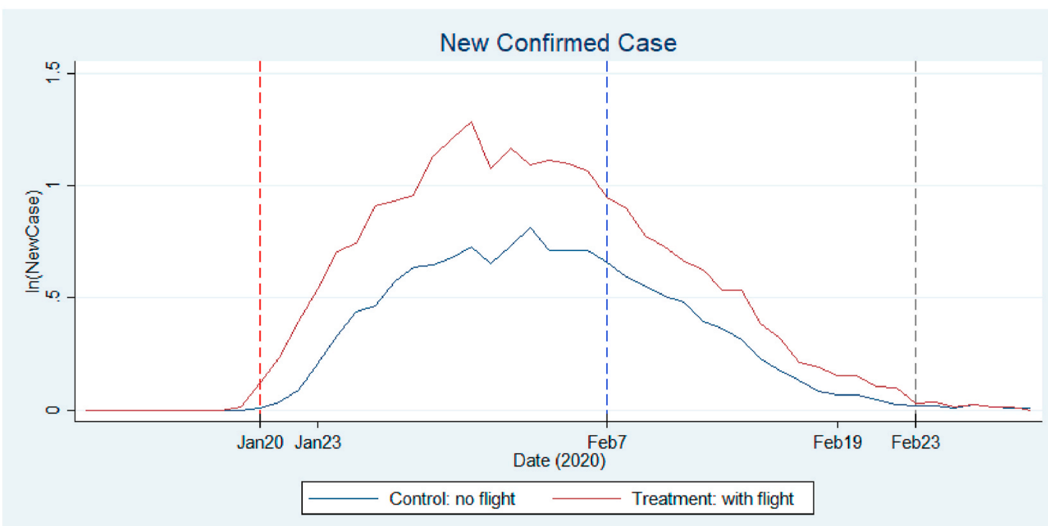


Fig. 5. Evolution of average new confirmed cases for cities.

disease. Setting up checkpoints in railway stations and airports to detect infected passengers and upgrading ventilation systems in trains, planes, train stations and airports are potential measures that could weaken the connectivity effects.

Second, in all models, the suspension effects were of smaller magnitude than the connectivity effects, suggesting that even complete suspension of transport links was insufficient to fully counteract the negative impact of earlier connectivity to the center of the outbreak. This may be due to the fact that connectivity effects had already generated a large number of cases locally, leading to higher rates of local transmission that could not be fully counteracted just by suspending transport links, suggesting that early intervention is imperative.

Third, HSR was associated with stronger amplifying effects on the COVID-19 outbreak than air. Connectivity effects were of larger magnitude for HSR than for air transport, and suspension effects were also more pronounced for HSR. The stronger connectivity effect was likely due to the fact HSR has a larger passenger capacity than air transport and plays a more important role in the domestic transportation network [20–22]. Thus policymakers should focus in particular on working with railways to control the spread of disease. While it is important for both HSR and air transport service providers to take preventative actions in the face of an epidemic, governments should prioritize HSR when allocating resources for epidemic control, and if the suspension of transport services is necessary, prioritize the suspension of HSR over the suspension of flight services to minimize non-essential travel. Certainly, this may cause equity issues as air travel is typically more expensive than HSR.

Fourth, our results indicate that the HSR connectivity effect diminished with greater distance from the epicentre of the outbreak. The HSR connectivity effect was strongest for cities a short distance from Wuhan (0–500 km) and was significantly weaker for medium-distance cities (501–1000 km); there was no significant effect for cities in the long-distance (over 1000 km) group. Suspension effects followed a similar pattern. This may be partially explained by different frequencies of train service to cities of different distances from Wuhan. We used train schedule information from the official China Railway website to calculate the average frequency of train service from and to Wuhan for each distance group between August 8 and August 13, 2020, when train service had fully resumed. The frequency data are consistent with the regression results: Group 1 (0–500 km) had the highest frequency of train service to and from Wuhan with an average train frequency of 19.2, followed by Group 2 (501–1000 km) with a frequency of 8.61, while

Group 3 (>1000 km) had the lowest average frequency of 3.5. The stronger effect of HSR connectivity on COVID-19 case numbers in cities closer to Wuhan can therefore be partially explained by the frequency of train service with Wuhan.

Lastly, we want to acknowledge two limitations. First, when estimating the connectivity effects of transportation networks, the inclusion of detailed frequencies of HSR trains and flights could have provided more precise estimates on their marginal effects. However, as such daily scheduling data is not available to us, we instead use dummy variables representing HSR and flight connectivity to capture the influence of transport connections. If available, future research is encouraged to use detailed scheduling data for more precise results. For models on the suspension effects, there were no more trains or flights scheduled after the Wuhan lockdown, hence it is most appropriate to use dummy variables to indicate suspension status.

Second, as discussed above, we use highway distance to Wuhan as a proxy for passenger inflows from Wuhan via road networks, with the purpose of ruling out the influence of transportation modes other than HSR or flights, such as inter-city buses, private cars and taxis. We understand it might not be the best measure of road transportation (buses, private cars or taxi), as it does not directly reflect the actual number of passengers via bus services or cars. However, passenger capacity of inter-city buses services is highly correlated with the highway distance between cities as it influences travel demand. Travel needs via private cars are less predictable, but this general trend is arguably still valid. Our connectivity models indicate that longer highway distance lead to fewer average daily confirmed cases. Hence, we believe it provides a good approximation of the magnitude of the road transportation impact. Note that our study mainly focuses on the roles of HSR and air travel. In future research that aims to provide more accurate estimates about the impacts of road transportation on infectious disease transmission and control, detailed data on the daily number of passengers travelled via road networks should be used. This would have important and more direct implications for many developing countries, where collective public transport using buses is more common, air travel is expensive, and HSR or rail transport is limited.

In conclusion, our study sheds light on the effectiveness and time efficiency of transportation suspension in disease control and has policy implications for transportation-related risk management in future pandemics. Early interventions and decisive action to impose travel restrictions when necessary are imperative to successfully impede the epidemic progression. Our results show that transportation suspension is indeed effective in curbing the virus spread, but it takes time for the effects to take place. Conditional on other local prevention and control measures, the time needed for suspension effects to fully materialize was almost twice the maximum incubation period. Admittedly, travel restrictions can be costly, exacerbating the economic shock caused by the epidemic and disproportionately hurting the poor and low-skilled. As a matter of fact, China had gone through a major economic contraction in the first quarter of 2020, with an acute and sharp reduction in GDP, fixed asset investment and employment rates [19]. However, swift suspension of transportation networks to the epicenter can minimize a nation's economic disruption, help other cities win time to implement preventative and control measures, and thereby aid a speedy economic recovery. This is one of the main reasons that China has only suffered for a relatively short time frame from the pandemic. China's GDP rebounded in the second quarter of 2020 and its growth rate reached 6.5% in the fourth quarter, 0.5% higher than the same period in 2019. This provides valuable lessons for other developing countries to overcome the current crisis..

Declarations

Travel Medicine and Infectious Disease requires that all authors sign a declaration of conflicting interests. If you have nothing to declare in any of these categories then this should be stated

Funding source

This work was supported by Innovation and Technology Commission, the Government of the Hong Kong Special Administrative Region [Grant Number: GSP/035/20]; the Natural Science Foundation of Guangdong Province [Grant Number: 2021A1515011250]; and the National Natural Science Foundation of China [Grant Number: 71573232].

CRediT authorship contribution statement

Pengyu Zhu: Conceptualization, Methodology, Validation, Investigation, Data curation, Writing – original draft, Writing – review & editing, Supervision. **Yuqing Guo:** Formal analysis, Software, Writing – original draft, Writing – review & editing, Visualization.

Declaration of competing interest

The authors declare no competing interests. Funding sources had no role in the design of this study and did not have any role during its execution, analyses, interpretation of the data, or decision to submit results. Any opinions, findings, conclusions or recommendations expressed in this material/event (or by members of the project team) do not reflect the views of the Government of the Hong Kong Special Administrative Region, the Innovation and Technology Commission or the General Support Programme Vetting Committee of the Innovation and Technology Fund.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tmaid.2021.102097.

References

[1] Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc Natl Acad Sci U S A 2006;103:20–2015. https://doi.org/10.1073/pnas.0510525103.

[2] Browne A, Ahmad SStO, Beck CR, Nguyen-Van-Tam JS. The roles of transportation and transportation hubs in the propagation of influenza and coronaviruses: a systematic review. J Trav Med 2016;23:1–7. https://doi.org/10.1093/jtm/tav002.

[3] Xu F, Connell McCluskey C, Cressman R. Spatial spread of an epidemic through public transportation systems with a hub. Math Biosci 2013;246:164–75. https://doi.org/10.1016/j.mbs.2013.08.014.

[4] Zhao S, Zhuang Z, Ran J, Lin J, Yang G, Yang L, et al. The association between domestic train transportation and novel coronavirus (2019-nCoV) outbreak in China from 2019 to 2020: a data-driven correlational report. Trav Med Infect Dis 2020;33:101568. https://doi.org/10.1016/j.tmaid.2020.101568.

[5] Wan L, Wan Q. The high-speed railway network and the cross-town transmission of epidemics: evidence from Chinese COVID-19. Lab. Rev. Lab. Econ. Ind. Relat. 2020: 37–48. https://doi.org/10.19313/j.cnki.cn10-1223/f.2020.02.005.

[6] Zhang Y, Zhang A, Wang J. Exploring the roles of high-speed train, air and coach services in the spread of COVID-19 in China. Transport Pol 2020;94:34–42. https://doi.org/10.1016/j.tranpol.2020.05.012.

[7] Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 2020:1–12. https://doi.org/10.1126/science.aba9757. 9757.

[8] Fang H, Wang L, Yang Y. Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China. https://doi.org/10.1101/2020.03.24.20042424; 2020.

[9] Kim K, Lee M jae. Difference in differences in reverse. Empir Econ 2019;57:705–25. https://doi.org/10.1007/s00181-018-1465-0.

[10] China Daily. White paper - fighting Covid-19: China in action. https://covid-19.chinadaily.com.cn/a/202006/08/WS5edd8bd6a3108348172515ec.html; 2020. accessed July 23, 2020.

[11] He G, Pan Y, Tanaka T. The short-term impacts of COVID-19 lockdown on urban air pollution in China. J. Nat. Sci. Sustain. Technol. 2020. https://doi.org/10.1038/s41893-020-0581-y.

[12] Vaisey S, Miles A. What you can—and can't—do with three-wave panel data. Socio Methods Res 2017;46:44–67.

[13] Wooldridge J. Econometric analysis of cross SEction and panel data. MIT Press; 2002.

[14] World Health Organization. Coronavirus disease 2019 (COVID-19) situation report – 73. https://www.who.int/docs/default-source/coronaviruse/situation-reports/

20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_6; 2020. accessed August 12, 2020.

[15] Cai J, Xu B, Chan KKY, Zhang X, Zhang B, Chen Z, et al. Roles of different transport modes in the spatial spread of the 2009 influenza A(H1N1) pandemic in mainland China. Int J Environ Res Publ Health 2019;16:222. https://doi.org/10.3390/ijerph16020222.

[16] Peak CM, Wesolowski A, zu Erbach-Schoenberg E, Tatem AJ, Wetter E, Lu X, et al. Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. Int J Epidemiol 2018;47: 1562–70. https://doi.org/10.1093/ije/dyy095.

[17] Luo K, Lei Z, Hai Z, Xiao S, Rui J, Yang H, et al. Transmission of SARS-CoV-2 in public transportation vehicles: a case study in Hunan Province, China. Open Forum Infect Dis 2020;7. https://doi.org/10.1093/ofid/ofaa430.

[18] Gurjav U, Burneebaatar B, Narmandakh E, Tumenbayar O, Ochirbat B, Hill-Cawthorne GA, et al. Spatiotemporal evidence for cross-border spread of MDR-TB along the Trans-Siberian Railway line. Int J Tubercul Lung Dis 2015;19:1376–82. https://doi.org/10.5588/ijtld.15.0361.

[19] Hale T, Liu X, Yang Y. China's economy shrinks for first time in four decades. Financ Times 2020. https://www.ft.com/content/8f941520-67ad-471a-815a-d6ba649d22ed. accessed March 17, 2021.

[20] Cao Jason, Zhu Pengyu. High-speed rail. Transport Lett Int J Transport Res 2017;9 (4). https://doi.org/10.1080/19427867.2017.1318472.

[21] Zhu Pengyu, Yu Tao, Chen Zhao. High-speed rail and urban decentralization in China. Transport Res Rec 2014;2475(1). https://doi.org/10.3141/2475-03.

[22] Yu Tao, Chen Zhao, Zhu Pengyu. Characteristics and mechanism of high speed rail-driven suburbanization in China: A case study of Beijing-Shanghai high-speed rail. Sci Geogr Sin 2012;32(9).