PNAS
www.pnas.org

Supplementary Information for

Quantifying COVID-19 importation risk in a dynamic network of domestic cities and international countries

Xiaoyi Han, Yilan Xu, Linlin Fan, Yi Huang, Minhong Xu, and Song Gao

Corresponding Author: Yilan Xu.
E-mail: yilanxu@illinois.edu

This PDF file includes:

Supporting Information Text

Data and sampling

Sampling and Domestic confirmed cases. Our analytical data include information for 284 Chinese cities from January 20, 2020, to April 28, 2020. January 20, 2020 is chosen as day 0 in our sample because it is the date that human-to-human transmission of COVID-19 was first confirmed in China. April 28, 2020 is chosen as the last day of our sample because Baidu Mobility data for 2019 are not available after May 7, 2019, the same day on the lunar calendar as April 25, 2020. Mudanjiang and Hulunbeier are excluded from our sample because they are land border cities with imported cases but without international flights hence abroad infection index. Cities with missing Baidu Mobility or weather data are also excluded from the analysis. Acknowledging the average COVID-19 incubation period of 5 days (1), we use Baidu Mobility data between January 13, 2020 and April 25, 2020 to construct the travel network for our sample; data over the same period on the 2019 lunar calendar are used as a proxy for the counterfactual unrestricted travel flows of 2020. We obtain the city-level daily newly confirmed cases from the China Data Lab (CDL) Dataverse accessed via the Harvard Dataverse (https://projects.iq.harvard.edu/chinadatalab/resources-covid-19). Starting from February 29, 2020, the CDL has recorded imported cases at provincial levels, for the four centrally-administered municipalities (CAMs) – Beijing, Shanghai, Chongqing, and Tianjin, as well as for the city of Guangzhou. Hence, we collect city-level daily imported cases from the province-level health commission daily updates and news reports to separate imported cases and domestic cases for all cities. Our main dependent variable is the city-level total newly confirmed cases including imported cases. In a robustness check, we use only domestic newly confirmed cases as the dependent variable. Table S1 presents the summary statistics of the variables used in the empirical analysis.

City-level weather condition. The daily weather information is collected from the Global Surface Summary of Day (GSOD) data accessed via the National Oceanic and Atmospheric Administration (NOAA). We use data from the 352 weather stations in China that reported data every day from December 23, 2019, to April 28, 2020, so that we can construct both contemporaneous and lagged weather conditions averaged over the past 7 days, 14 days, 21 days, and 28 days for each city over our sample period. The variables we used include the average daily values of temperature (°C) and wind speed (m/s), daily maximum sustained wind speed (m/s), and daily precipitation amount (mm) (2).

To construct city-level daily weather conditions, we first select weather stations within a 100km radius centered on each city's centroid with a 50km incremental increase, up to a radius of 250km. Next, we adopt the inverse-distance weighting method to compute the weighted average daily weather conditions for each city using daily data collected from the matched nearby weather stations. The weights are inversely proportional to the distance between a city's centroid and its nearby stations.

Hospital information. We obtain the numbers of hospitals for 20 categories of hospitals at the city level from the CDL. We match the data with the number of doctors, GDP per capita, and population density at the city-level from (2). We remove the following 14 cities (their names with city code) after matching: Xingan 1522, Liangshanyizu 5134, Daxinganlingdiqu 2327, Dalibaizu 5329, Bayinguolengmenggu 6528, Yanbianchaoxianzu 2224, Dehongdaizujingpozu 5331, Wenshanzhuangzumiaozu 5326, Chuxiongyizu 5323, Ganzicangzu 5133, Honghehanizuyizu 5325, Xishuangbannadaizu 5328, Xilinguole 1525, Akesudiqu 6529.

The network of domestic cities and international countries

Domestic travel flow network. Baidu Mobility (https://qianxi.baidu.com/) serves as the source for inter-city population movement. Baidu records real-time location from its mobile mapping software to track daily human mobility. For inter-city travels, Baidu Mobility provides the daily percentages of population moving into a city from each of the top 100 origination cities (the inflow percentage) and the daily percentages of population moving out of a city to each of the top 100 destination cities (the outflow percentage). The daily inflow/outflow percentages from the top 100 origination/destination cities sum up to more than 90%, suggesting that Baidu Mobility represents China's daily population movement well (3, 4). Baidu Mobility portal also provides the daily inflow and outflow movement intensity indices for 389 geographic units, including cities and provinces. These intensity indices are consistent and comparable over time and across cities.

Given that the average incubation period of the COVID-19 is approximately five days (1), a lagged 5-day average of percentage weights, from $t-7$ to $t-3$, is constructed for each city i on each date t taking the following steps.

Step 1: For each date t and city i, we construct the lagged 5-day moving average as follows:

$$\tilde{w}_{ijt} = \frac{\sum_{l=3}^{7} p_{ij,t-l}}{5}, \quad \text{Lagged 5-day average inflow percentage}$$

Step 2: we row normalise $\tilde{w}_{ijt}$.

$$\bar{w}_{ijt} = \frac{\tilde{w}_{ijt}}{\sum_j \tilde{w}_{ijt}}, \quad \text{Normalized 5-day average inflow percentage}$$

We construct the spatial weights matrices $\bar{W}_t = [\bar{w}_{ijt}]$ to describe the time-varying travel flow network across 283 cities excluding Wuhan. We also include the inflow intensity indices MI_{it}'s as control variables in the model. To capture the influence of travel flows from Wuhan to other cities, we further pick out the row and column of Wuhan from $\bar{W}_t$, which consists of flows initiated from Wuhan, as an additional regressor in the model.

Table S1. Summary Statistics - 284 Chinese Cities Including Wuhan

Variables	Mean	Std.Dev.	Min	Max
Time-Varying Variables, Jan 20 - Apr 28				
Daily Total Newly Confirmed Cases	2.836	91.042	-107	13436
5-day Average Daily Inflow Intensity Index 2020	0.781	0.952	0.007	9.230
5-day Average Daily Inflow Intensity Index 2019	1.461	1.779	0.066	20.708
5-day Average Daily Abroad Infection Index 2020	0.234	1.104	0	11.778
5-day Average Daily Abroad Infection Index 2019	0.563	1.720	0	12.869
Daily Temperature, ℃	9.596	8.337	-25.333	31.778
Weekly Average Temperature, 1 week lag	8.995	8.007	-22.524	29.643
Weekly Average Temperature, 2 week lag	8.120	8.297	-22.976	28.516
Weekly Average Temperature, 3 week lag	7.341	8.419	-23.754	28.516
Weekly Average Temperature, 4 week lag	6.661	8.658	-25.349	28.516
Daily Minimum Temperature, ℃	3.687	8.966	-33.389	24.751
Weekly Average Minimum Temperature, 1 week lag	3.208	8.854	-30.595	22.993
Weekly Average Minimum Temperature, 2 week lag	2.413	9.084	-30.595	22.993
Weekly Average Minimum Temperature, 3 week lag	1.791	9.300	-30.595	22.993
Weekly Average Minimum Temperature, 4 week lag	1.137	9.596	-31.198	22.993
Daily Maximum Temperature, ℃	16.038	8.436	-18.500	39.778
Weekly Average Maximum Temperature, 1 week lag	15.350	7.911	-16.738	38.381
Weekly Average Maximum Temperature, 2 week lag	14.376	8.240	-16.921	38.381
Weekly Average Maximum Temperature, 3 week lag	13.447	8.297	-17.214	38.381
Weekly Average Maximum Temperature, 4 week lag	12.725	8.479	-17.698	38.381
Daily Average Wind Speed, m/s	2.453	1.199	0.154	21.710
Weekly Average Wind Speed, 1 week lag	2.430	0.870	0.397	12.567
Weekly Average Wind Speed, 2 week lag	2.378	0.847	0.397	12.567
Weekly Average Wind Speed, 3 week lag	2.342	0.840	0.397	12.486
Weekly Average Wind Speed, 4 week lag	2.321	0.835	0.397	12.486
Daily Maximum Sustained Wind Speed, m/s	4.231	2.004	0.514	28.089
Weekly Average Maximum Sustained Wind Speed, 1 week lag	4.189	1.474	0.955	18.968
Weekly Average Maximum Sustained Wind Speed, 2 week lag	4.095	1.445	0.955	18.968
Weekly Average Maximum Sustained Wind Speed, 3 week lag	4.025	1.437	0.955	18.968
Weekly Average Maximum Sustained Wind Speed, 4 week lag	3.991	1.430	0.955	18.968
Daily Precipitation Amount, mm	2.180	7.101	0	104.140
Weekly Average Precipitation Amount, 1 week lag	2.202	4.227	0	54.320
Weekly Average Precipitation Amount, 2 week lag	2.144	4.300	0	54.320
Weekly Average Precipitation Amount, 3 week lag	2.232	4.506	0	54.320
Weekly Average Precipitation Amount, 4 week lag	2.177	4.657	0	54.320
City characteristics				
GDP per capita, 10,000RMB	5.239	2.904	1.141	16.741
Population Density, per km^2	446.692	372.878	9.049	3444.092
# of Doctors, 10,000	1.135	1.136	0.032	10.938
Total # of Hospitals	2151.996	1005.824	111	5790
# of Tier-3 A Hospitals, Sanjia	43.887	98.025	0	877
# of Specialized Hospitals, Zhuanke	2.461	3.168	0	30
# of Infectious Disease Hospitals, Chuanran	4.116	4.815	0	27
# of Medical Care Service Facilities, Yiliaobaojian	663.250	212.421	22	1155
# of Health Centers, Weishengyuan	248	197.758	5	892
# of Stomatology Hospitals, Kouqiang	220.690	212.340	5	877
# of Obstetrical and Gynecological Hospitals, Fuke	13.246	16.097	0	126
# of Emergency Care Centers, Jijiu	16.465	19.466	0	202
# of Plastic Surgery HospitalS, Zhengxing	26.570	50.189	0	466
# of Disease Prevention HospitalS, Jibingfangyu	32.032	23.340	0	143
# of Ophthalmology Hospitals, Yanke	22.415	22.164	0	138
# of Psychiatric Hospitals, Jingshen	2.662	4.688	0	61
# of General Hospitals, Zonghe	216.173	142.735	12	674
# of Otolaryngology Hospitals, Erbihou	4.915	6.212	0	39
# of Tumor Hospitals, Zhongliu	3.539	5.683	0	49
# of horacic Hospitals, Xiongke	1.940	3.594	0	34
# of Brian Health Hospitals, Naoke	1.845	3.067	0	22
# of Clinics, Zhensuo	607.806	209.326	43	1027
# of Orthopedics Hospitals, Guke	19.982	23.818	0	170
Observations		28400		

Note: Time-varying variables are observed daily for each city. 5-day average indices are averaged over $t-7$ to $t-3$. Weekly averages are averaged over the proceeding 7 days, 14 days, 21 days, and 28 days, respectively. City characteristics are obtained from (2) and CDL.

Counterfactual travel flows. We use the 2019 Baidu inflow indices to construct proxies for unrestricted 2020 travel flows. The 2019 daily inflow percentages are not available, thus we have to assume that the inter-city daily travel percentage patterns do not change by calendar year, which is not unreasonable given the consistent seasonality across years in China (5). We obtain the inflow movement intensity indies from January 24, 2019, to May 7, 2019 (comparable to January 13, 2010, to April 25, 2020, on the lunar calendar) from the CDL. Figure S1 plots the daily inflow intensity indices, aggregate over the 284 cities, for 2019 and 2020.

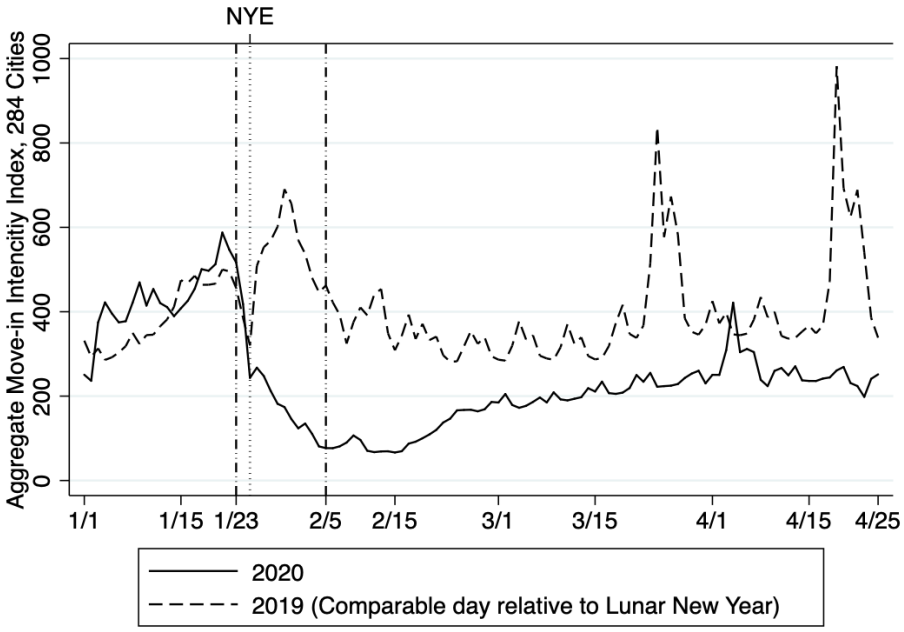


Fig. S1. Aggregate inflow intensity indices. Notes: The figure plots the daily inflow intensity indices aggregated over the 284 cities for 2019 and 2020 using data from the CDL Dataverse. The dotted vertical line indicates New Year's Eve in 2020 (January 24, 2020). The first dash-dotted vertical line marks the date of the Wuhan lockdown (January 23, 2020). The second dash-dotted vertical line corresponds to the date on which the 2020 aggregate inflow intensity index reached its minimum.

In Figure S1, the dotted vertical line indicates New Year's Eve in 2020 (January 24, 2020). The first dash-dotted vertical line marks the date of the Wuhan lockdown (January 23, 2020). The second dash-dotted vertical line corresponds to the date on which the 2020 aggregate inflow intensity index reached its minimum. As shown in Figure S1, the aggregate inflow intensity indices increased during ChunYun (the Spring Festival Travel Rush) and dropped right before New Year's Eve in both years. The inflow intensity index saw a sharp increase in 2019 after the New Year as people returned to work. However, the index continued to decrease in 2020 due to the Wuhan lockdown and travel restrictions imposed in other cities. The 2019 index spiked in late March and late April, which corresponded to the Qingming Festival and the International Labor's Day Holiday in 2019.

Given the incubation period of COVID-19, we smooth the inflow intensity indices by taking the average over $t-7$ to $t-3$ for each city and label the corresponding lagged index for city i as ξ_{it}. In the SDPD model, the vector of inflow intensity indices $\xi_t = (\xi_{1t}, \xi_{2t}, \cdots, \xi_{nt})'$ with $n = 283$.

International flight network and abroad infection index. Our analysis includes flight routes between 48 origins, including Hong Kong, Macau, and Taiwan in China and 45 foreign countries, and 44 destination cities in China with international airports.* We aggregate the daily number of international flights for each origin-destination pair. In Figure S2, Panel A plots the time trends in the total number of international flights aggregated over all the destination cities. The 2020 average daily number of flights started to decrease by two thirds from the beginning of February. By contrast, the 2019 daily flight volume stayed at a relatively stable level before increasing around the mid-March of 2020, which corresponded to the end of March in 2019, driven by the seasonal trends. Panel B plots the time trends of the total number of origins. Consistently, the 2020 number of origins began to decrease from the early February and then remained at a low level before going up in April, whereas the 2019 trends stayed at a stable level during the entire period.

We construct a destination- and time-specific abroad infection index using the method as below:

$$b_{it} = \sum_k \#cases_{kt} * \#flights_{ikt} \quad [1]$$

*According to the policy of the Civil Aviation Administration of China, all international flights into Beijing have been rerouted to one of 12 cities prior to landing in the capital since March 23, 2020. See more at http://www.caac.gov.cn/XWZX/MHYW/202003/P020200322518497533717.pdf and https://www.airchina.com.cn/cn/info_and_services/138279.shtml

where i indexes a destination city in China, k indexes the origin of international/regional flights, and t indexes time. $\#cases_{kt}$ represents the number of daily newly confirmed cases of origin k on day t, sourced from the European Centre for Disease Prevention and Control (6). $\#flights_{ikt}$ represents the number of international flights for each i-k pair on day t. We smooth the index by taking the average over $t-7$ to $t-3$ considering the incubation period of Covid-19. We then take the log transformation of the smoothed index as $\tilde{b}_{it} = log(b_{it}+1)$ in the regression. In the SDPD model, $B_t = (\tilde{b}_{1t}, \tilde{b}_{2t}, ..., \tilde{b}_{nt})'$ with $n = 283$.

To conduct the counterfactual analysis, we also construct the 2019 abroad infection index using the 2020 confirmed cases and the 2019 actual flight volume for the comparable day on 2019 lunar calendar. Figure S3 demonstrates the divergent time trends between the 2020 and the 2019 abroad indices since March, indicating the dramatically decreasing number of international flights in 2020.

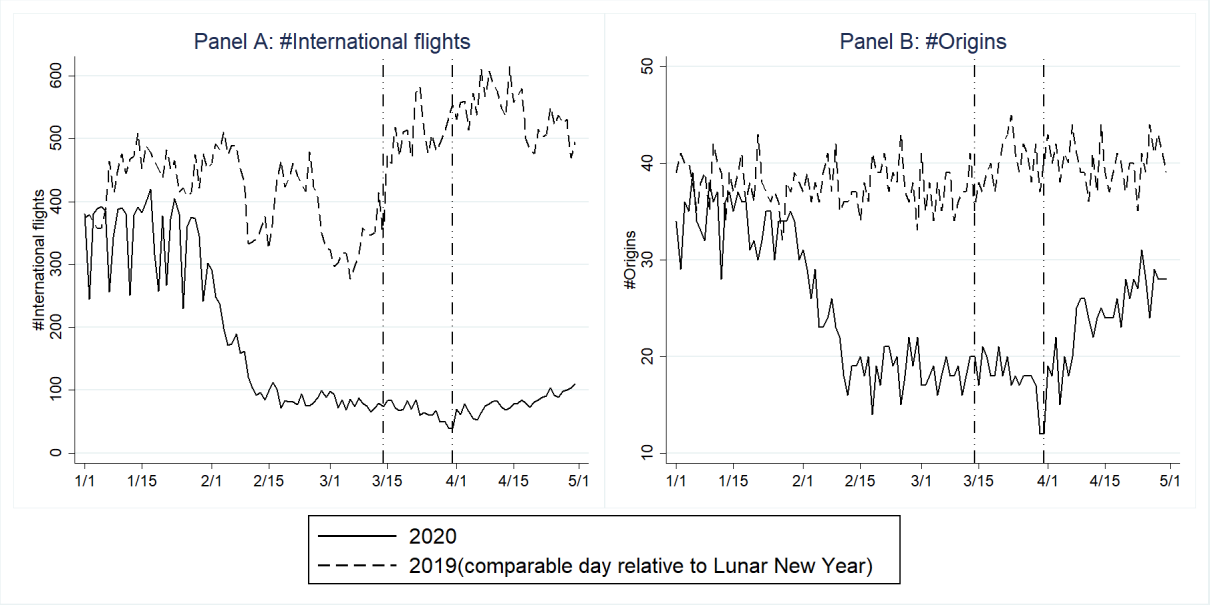


Fig. S2. Time trends in the total number of flights and origin countries by 2020 and 2019. Notes: Panel A and B plot the time trends in the total number of international flights and the time trends in the total number of origins, respectively. The two dashed vertical lines represent the identified change points for the abroad infection index.

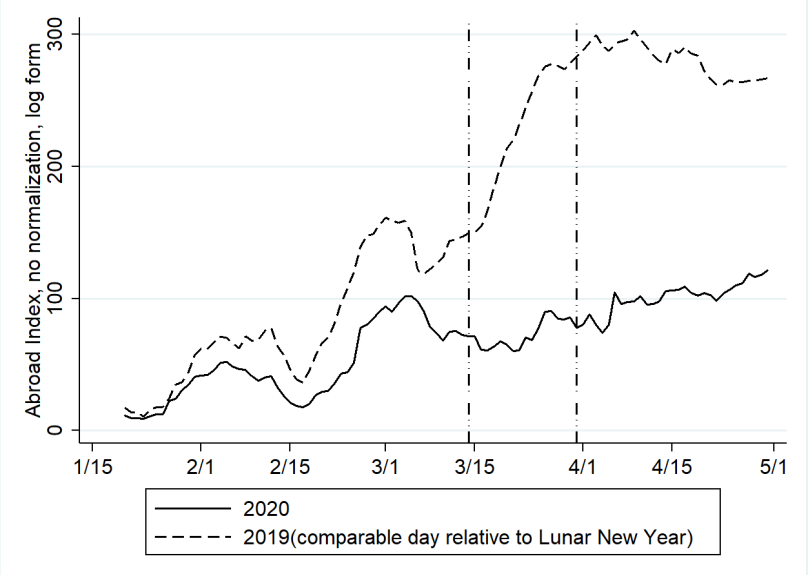


Fig. S3. Time trends in the abroad infection index by 2020 and 2019. Notes: The figure plots the abroad infection index by 2020 and 2019 comparable day relative to Lunar New Year. The two dashed vertical lines represent the identified change points for the abroad infection index.

The Spatial Dynamic Panel Data (SDPD) model

Full model specification. We treat the city of Wuhan as the initial epidemic center, which was hit by the first out-break of COVID-19 in China, and integrate the transmission from Wuhan to other cities, the cross and within city transmission, and the international transmission across borders to cities in a spatial panel framework. The SDPD model considered in the scalar form is

$$y_{it} = \lambda(t)\sum_{j=1}^{n}\bar{w}_{ijt}y_{jt} + \gamma(t)\bar{w}_{i,wht}y_{wh,t} + \rho(t)y_{i,t-1} + \mu(t)\sum_{j=1}^{n}\bar{w}_{ij,t-1}y_{j,t-1} + \tilde{b}_{it}\delta(t) + \xi_{it}\phi(t) + x'_{it}\beta_1 + c'_i\beta_2 + \alpha_t + u_{it}, \quad [2]$$

where y_{it} denotes the number of newly confirmed cases of COVID-19 in city i at day t for $i = 1, 2, \cdots, n$ with $n = 283$, $y_{wh,t}$ represents the number of newly confirmed cases in Wuhan at day t, and, x_{it} and c_i are, respectively, time-varying $k_1 \times 1$ and time-invariant $k_2 \times 1$ vector of controls for city i. Different from (2), the cross city transmission effect $\lambda(t)$, does not represent a casual transmission relation from city i's neighbors to i itself. Rather, it captures the spillover effect of virus transmission in an equilibrium steady state among cities through the travel flow network. With variables stacked as $Y_t = (y_{1t}, y_{2t}, \cdots, y_{nt})'$, the 283×1 vector of newly confirmed cases for 283 Chinese cities outside Wuhan; $\bar{W}_t = [\bar{w}_{ijt}]$, the 283×283 spatial weights matrix representing domestic network, constructed by 5-day lagged average travel flows between cities; $M_{wh,t} = (\bar{w}_{1,wht}, \bar{w}_{2,wht}, \cdots, \bar{w}_{n,wht})'$, the 283×1 vector of 5-day lagged average share of travel inflow to a city from Wuhan; $B_t = (\tilde{b}_{1t}, \tilde{b}_{2t}, \cdots, \tilde{b}_{nt})'$, $\xi_t = (\xi_{1t}, \xi_{2t}, \cdots, \xi_{nt})'$, $X_t = (x_{1t}, x_{2t}, \cdots, x_{nt})'$, includes contemporaneous and lagged city level weather conditions; $C = (c_1, c_2, \cdots, c_n)'$, time-invariant city level variables including the numbers of various types of hospitals; $U_t = (u_{1t}, u_{2t}, \cdots, u_{nt})'$; l_n is a 283×1 vector of ones and α_t is the scalar time effect at time t, we can derive the matrix expression of the model as in the main text:

$$Y_t = \lambda(t)\bar{W}_tY_t + \gamma(t)M_{wh,t}y_{wh,t} + \rho(t)Y_{t-1} + \mu(t)W_{t-1}Y_{t-1} + B_t\delta(t) + \xi_t\phi(t) + X_t\beta_1 + C\beta_2 + l_n\alpha_t + U_t. \quad [3]$$

Let $I(.)$ be the 0-1 indicator function. We model the time-varying transmissibility parameter via some change points except $\mu(t)$,

$$\begin{aligned}
\lambda(t) &= \lambda_1 \times I(t \le \tau_\lambda) + \lambda_2 \times I(t > \tau_\lambda),\\
\rho(t) &= \rho_1 \times I(t \le \tau_\rho) + \rho_2 \times I(t > \tau_\rho),\\
\gamma(t) &= \gamma_1 \times I(t \le \tau_\gamma) + \gamma_2 \times I(t > \tau_\gamma),\\
\delta(t) &= \delta_1 \times I(t \le \tau_{\delta|1}) + \delta_2 \times I(\tau_{\delta|1} < t \le \tau_{\delta|2}) + \delta_3 \times I(t > \tau_{\delta|2}).
\end{aligned}$$

As will be shown in the following subsection on Bayesian estimation, we do not find a significant diffusion effect when estimating a SDPD model with constant coefficients by month. Hence we simply set $\mu(t)$ to be a time-invariant parameter, i.e., $\mu(t) = \mu$. Moreover, we allow two different change points, $\tau_{\delta|1}$ and $\tau_{\delta|2}$ for importation, to capture the initial emergence and later decrease of the international transmission. After introducing these parameter specific change points for domestic and international transmissibility, our model is more flexible than the typical SDPD model with constant coefficients in (7) and is a further generalization of the spatial panel data model with a common structural break date for all parameters in (8).

The transimission within Wuhan, the epidemic center, is described by the following AR(1) specification:

$$y_{wh,t} = \kappa_1 y_{wh,t-1} + x'_{wh,t}\kappa_2 + u_{wh,t},\ t = 1, 2, \cdots, T, \quad [4]$$

where $x_{wh,t}$ includes contemporaneous and lagged weather variables of Wuhan, and $u_{wh,t} \sim \mathcal{N}(0, \sigma^2_{wh})$ is assumed to be independent across t and of u_{it}'s.

Stability conditions. Let $\lambda = (\lambda_1, \lambda_2)'$, $\rho = (\rho_1, \rho_2)'$, $\gamma = (\gamma_1, \gamma_2)'$, $\phi = (\phi_1, \phi_2)'$ and $\delta = (\delta_1, \delta_2, \delta_3)'$ be vectors of possible parameter values that $\lambda(t)$, $\rho(t)$, $\gamma(t)$, $\phi(t)$ and $\delta(t)$ may take. Denote $S_t(\lambda, \tau_\lambda) = I_n - \lambda(t)\bar{W}_t$ and $A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho) = S_t^{-1}(\lambda, \tau_\lambda)[\rho(t)I_n + \mu\bar{W}_{t-1}]$. Conditional on $y_{wh,t}$, the reduced form of the SDPD model is

$$Y_t = A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho)Y_{t-1} + S_t^{-1}(\lambda, \tau_\lambda)[\gamma(t)M_{wh,t}y_{wh,t} + B_t\delta(t) + \xi_t\phi(t) + X_t\beta_1 + C\beta_2 + l_n\alpha_t + U_t]. \quad [5]$$

To ensure the reduced form exists, stability conditions need to be imposed on $S_t(\lambda, \tau_\lambda)$ and $A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho)$. Specifically, we require $S_t(\lambda, \tau_\lambda)$ to be invertible. According to (9), one sufficient condition is $||\lambda(t)\bar{W}_t||_\infty < 1$, with $||.||_\infty$ being the maximum row sum matrix norm. As $||\bar{W}_t||_\infty \le 1$, we may adopt the following condition:

$$||\lambda(t)\bar{W}_t||_\infty \le |\lambda(t)|||\bar{W}_t||_\infty \le \max\{|\lambda_1|, |\lambda_2|\} < 1. \quad [6]$$

We also require $||A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho)||_\infty < 1$. Notice that

$$\begin{aligned}
||S_t^{-1}(\lambda, \tau_\lambda)||_\infty &= ||I_n + \lambda(t)\bar{W}_t + \lambda(t)^2\bar{W}_t^2 + \cdots||_\infty \le ||I_n||_\infty + |\lambda(t)|||\bar{W}_t||_\infty + |\lambda(t)|^2||\bar{W}_t||^2_\infty + \cdots\\
&\le 1 + \max\{|\lambda_1|, |\lambda_2|\} + \max\{|\lambda_1|^2, |\lambda_2|^2\} + \cdots = \frac{1}{1 - \max\{|\lambda_1|, |\lambda_2|\}}.
\end{aligned}$$

Then, the stability condition on $A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho)$ is

$$||A_t(\lambda, \rho, \mu, \tau_\lambda, \tau_\rho)||_\infty \leq ||S_t^{-1}(\lambda, \tau_\lambda)||_\infty \times [|\rho(t)| + |\mu|||\bar{W}_{t-1}||_\infty] \leq \frac{1}{1 - \max\{|\lambda_1|, |\lambda_2|\}}[\max\{|\rho_1|, |\rho_2|\} + |\mu|] < 1, \quad [7]$$

which suffices to

$$\max\{|\lambda_1|, |\lambda_2|\} + \max\{|\rho_1|, |\rho_2|\} + |\mu| < 1. \quad [8]$$

These stability conditions would be imposed on λ, ρ and μ in their MCMC sampling steps. Notice that the stability condition on ρ does not necessarily rule out the scenario of rapidly growing infection cases in cities at the early phase of the epidemic in China. This is so, because the marginal effect of one more lagged confirmed case, $y_{i,t-1}$ on the current newly confirmed case y_{it} goes beyond the within-transmission parameter (the AR coefficient) ρ. Specifically, the reduced form of the SDPD model can be rewritten as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{pmatrix} = \begin{pmatrix} S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{11} & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{12} & \cdots & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{1n} \\ S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{21} & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{22} & \cdots & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{n1} & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{n2} & \cdots & S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{nn} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{n,t-1} \end{pmatrix} \quad [9]$$

$$+ S_t^{-1}(\lambda, \tau_\lambda)(\mu \bar{W}_{t-1} Y_{t-1} + \gamma(t) M_{wh,t} y_{wh,t} + B_t \delta(t) + \xi_t \phi(t) + X_t \beta_1 + C\beta_2 + l_n \alpha_t + U_t), \quad [10]$$

where $S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho) = S_t^{-1}(\lambda, \tau_\lambda)\rho(t) I_n$ and $S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{ij}$ denotes the ijth element of the $n \times n$ matrix $S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)$. Consider a simplified case of no diffusion effect $\mu(t) = 0$. Let $G_t = \gamma(t) M_{wh,t} y_{wh,t} + B_t \delta(t) + \xi_t \phi(t) + X_t \beta_1 + C\beta_2 + l_n \alpha_t + U_t$ and $[S_t^{-1}(\lambda, \tau_\lambda)]_i$ be the ith row of $S_t^{-1}(\lambda, \tau_\lambda)$. y_{it}, the total newly confirmed cases in city i at day t, is determined by

$$y_{it} = \sum_{j=1}^{n} S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{ij} y_{j,t-1} + [S_t^{-1}(\lambda, \tau_\lambda)]_i G_t.$$

So the marginal effect of $y_{j,t-1}$ on y_{it} is

$$\frac{\partial y_{it}}{\partial y_{j,t-1}} = S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{ij}, \ j = 1, 2, \cdots, n. \quad [11]$$

This implies the direct influence of $y_{i,t-1}$ on y_{it}, captured by $S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{ii}$, goes beyond the AR coefficient ρ and essentially represents a combined effect of the within city transmission ρ, the cross-city transmission λ, as well as the travel flow network $\bar{w}_{ij}$'s. Conceptually, when $\lambda \neq 0$, in addition to the direct influence of $y_{i,t-1}$ on y_{it} captured by ρ, $y_{i,t-1}$ may go beyond y_{it} to influence y_{jt} due to the cross-city transmission effect λ, and then y_{jt} can go back to affect y_{it} again through λ. This feedback effect, initialized by $y_{i,t-1}$, also holds for longer paths where y_{jt} may first affect y_{kt} then go back to y_{it}(10). Thus, as long as the cross-city transmission effect λ is significantly positive, the direct influence of $y_{i,t-1}$ on y_{it} would be larger than the AR coefficient ρ. Moreover, from (11), with a significant cross-city transmission effect, the lagged newly confirmed cases in other cities, $y_{j,t-1}$ may also contribute to the increase in y_{it} through an indirect effect of $S_{rt}(\lambda, \rho, \tau_\lambda, \tau_\rho)_{ij}$ for $j \neq i$, which may arise because $y_{j,t-1}$ affects y_{jt}, and then y_{jt} affects y_{it}, along with other longer paths initialized from $y_{j,t-1}$ to y_{it}, not to mention the influence from possible diffusion effect $\mu(t)\bar{w}_{ij,t-1} y_{j,t-1}$ if $\mu \neq 0$. Therefore, although the AR coefficient (within-city transmission) ρ is less than one, the direct and indirect effects of all cities' lagged infections, could lead to a much larger influence on y_{it} than ρ and accommodate the scenario of rapidly growing infection cases in many cities.

Overview of the Bayesian estimation. Let $\theta = (\lambda', \rho', \mu, \gamma', \phi', \delta', \beta_1', \beta_2', \{\alpha_t\}, \sigma^2)'$ be the vector of parameters of the SDPD model and $\tau = (\tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)'$ with $\tau_\delta = (\tau_{\delta|1}, \tau_{\delta|2})'$ be the collection of change points. Let $H_t(\theta, \tau) = S_t(\lambda, \tau_\lambda) Y_t - \gamma(t) M_{wh,t} y_{wh,t} - \rho(t) Y_{t-1} - \mu \bar{W}_t Y_{t-1} - B_t \delta(t) - \xi_t \phi(t) - X_t \beta_1 - C\beta_2 - l_n \alpha_t$, $\kappa = (\kappa_1', \kappa_2')'$ and $h(\kappa) = y_{wh,t} - \kappa_1 y_{wh,t-1} - x_{wh,t}' \kappa_2$. Conditional on $y_{wh,t}$, the likelihood function of the SDPD model at time t is

$$f(Y_t | y_{wh,t}, \theta, \tau) \propto (\sigma^2)^{-\frac{n}{2}} \times |S_t(\lambda, \tau_\lambda)| \times \exp(-\frac{H_t(\theta, \tau)' H_t(\theta, \tau)}{2\sigma^2}). \quad [12]$$

The likelihood function of the full model at time t is

$$\begin{aligned} f(Y_t, y_{wh,t} | \theta, \tau, \kappa_1, \kappa_2, \sigma_{wh}^2) &\propto f(Y_t | y_{wh,t}, \theta, \tau) \times f(y_{wh,t} | \kappa_1, \kappa_2, \sigma_{wh}^2) \\ &\propto (\sigma^2)^{-\frac{n}{2}} \times |S_t(\lambda, \tau_\lambda)| \times \exp(-\frac{H_t(\theta, \tau)' H_t(\theta, \tau)}{2\sigma^2}) \times (\sigma_{wh}^2)^{-\frac{1}{2}} \exp(-\frac{h_t(\kappa)' h_t(\kappa)}{2\sigma_{wh}^2}). \end{aligned} \quad [13]$$

For simplicity, the exogenous and pre-determined control variables ξ_t, B_t, X_t, C and $x_{wh,t-1}$, as well as the pre-determined spatial weights matrices of travel flows $\bar{W}_t$'s, $M_{wh,t}$'s and lagged dependent variables Y_{t-1} and $y_{wh,t-1}$ are suppressed from the conditional set of the likelihood function. We focus on the Bayesian estimation of θ and τ in the conditional likelihood function

of [12]. The parameters in the AR(1) specification of Wuhan in [4] can be directly sampled by two Gibbs steps for $\kappa = (\kappa_1', \kappa_2')$ and σ_{wh}^2. Denote $\bar{\beta}_1 = (\gamma_1, \gamma_2, \delta_1, \delta_2, \delta_3, \phi_1, \phi_2, \beta_1')'$. Assume the following priors for θ and τ:

$$\lambda_l \sim \mathcal{U}(-1,\ 1),\ \rho_l \sim \mathcal{U}(-1,\ 1), l = 1, 2;\ \mu \sim \mathcal{U}(-1,\ 1),$$
$$\bar{\beta}_1 \sim \mathcal{N}(\bar{\beta}_{10}, B_1);\ \beta_2 \sim \mathcal{N}(\beta_{20}, B_2),\ \sigma^2 \sim IG(\frac{a}{2}, \frac{b}{2}),\ \alpha_t \sim \mathcal{N}(\alpha_0, \sigma_\alpha^2),\ t = 1, 2, \cdots, T;$$
$$\tau_\lambda \sim U(14,\ T-1),\ \tau_\rho \sim U(14,\ T-1),\ \tau_\gamma \sim U(14,\ T-1),\ \tau_\phi \sim U(14,\ T-1),\ \tau_{\delta|1} \sim U(40,\ 70),\ \tau_{\delta|2} \sim U(71,\ T-1). \quad [14]$$

We mostly follow (10) to specify the priors for parameters in θ, such as the uniform priors for λ and ρ, normal priors for $\bar{\beta}_1$ and β_2, and inverse-gamma prior for σ^2. Furthermore, we assume independent uniform priors for τ_λ, τ_ρ, τ_γ, τ_ϕ, $\tau_{\delta|1}$ and $\tau_{\delta|2}$. Our prior elicitation on those change points is based upon the by-month estimations of a typical SDPD model with constant transmissibility reported in Table S2.

Table S2. By-month estimations of a constant coefficient SDPD model.

	Jan	Feb	Mar	Apr
Cross-city transmission	0.143	0.225	-0.010	-0.045
	(0.031)	(0.022)	(0.014)	(0.016)
Within city transmission	0.584	0.487	0.490	0.195
	(0.015)	(0.009)	(0.009)	(0.010)
Diffusion effect	0.062	0.024	-0.003	-0.029
	(0.038)	(0.024)	(0.019)	(0.019)
Importation	0.833	-0.107	0.088	0.027
	(0.498)	(0.243)	(0.010)	(0.010)
Wuhan transmission	0.103	0.023	0.002	0.000
	(0.006)	(0.001)	(0.001)	(0.002)
Inflow index	-0.027	0.305	0.052	0.064
	(0.217)	(0.326)	(0.020)	(0.021)
Time effect	Yes	Yes	Yes	Yes
Weather control	Yes	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes	Yes

Note: n=283 excluding Wuhan; Jan: Jan 21-Feb 3; Feb: Feb 4-Feb 29; Mar: Mar 1-Mar 31, Apr: Apr 1-Apr 28; We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

It is evident that the parameters for cross and within city transmission, and the effect of inflow index and Wuhan transmission all changed after January, so we set the lower bounds of $\lambda(t)$, $\rho(t)$, $\phi(t)$ and $\gamma(t)$ to be 14. Furthermore, we notice multiple structure changes for importation risk, the parameter of which becomes significant in March but soon disappears in April. So we require the first change point $\tau_{\delta|1}$ to be between February and March and put the second change point $\tau_{\delta|2}$ in April. Additionally, in the following empirical and simulation study, we specify the remaining hyperparameters of the prior distribution as follows: $\bar{\beta}_{10} = 0,\ B_1 = 10I_{k_1},\ \beta_{20} = 0,\ B_2 = 10I_{k_2},\ a = 0.001,\ b = 0.001$ and $\alpha_0 = 0,\ \sigma_\alpha^2 = 1$. These hyperparameters generally lead to relatively flat (uninformative) prior distributions over the parameter spaces. With the priors $\pi(.)$ specified in [14], by Bayes' theorem, the posterior distribution of all parameters θ and change-points τ is

$$p(\theta, \tau | \{Y_t\}, \{y_{wh,t}\}) \propto \pi(\theta) \times \pi(\tau) \times f(\{Y_t\} | \{y_{wh,t}\}, \theta, \tau). \quad [15]$$

We extend the MCMC sampler in (11) by incorporating sampling steps for τ, in addition to the steps for θ. With conjugate priors, $\bar{\beta}_1$, β_2, σ^2 and α_t's can be sampled directly from their posterior distributions, which is either normal or inverse-gamma distributions, via Gibbs sampling steps. But Metropolis-Hasting (M-H) steps are needed to sample the 5×1 parameter vector $\psi = (\lambda_1, \lambda_2, \rho_1, \rho_2, \mu)'$ and the change points τ_λ, τ_ρ, τ_γ, τ_ϕ and τ_δ because the corresponding conditional posterior distributions do not take a known form. Below we list the full MCMC sampling steps:

- Step 1: Conditional on τ_λ, τ_ρ, τ_γ, τ_ϕ and τ_δ, we obtain the posterior draws of ψ, $\bar{\beta}_1$, β_2, σ^2 and α_t's.
  - Step 1.1: Sample ψ from $p(\psi | \{Y_t\}, \{y_{wh,t}\}, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a M-H step;
  - Step 1.2: Sample $\bar{\beta}_1$ from $p(\bar{\beta}_1 | \{Y_t\}, \{y_{wh,t}\}, \psi, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a Gibbs step;
  - Step 1.3: Sample β_2 from $p(\beta_2 | \{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a Gibbs step;
  - Step 1.4: Sample σ^2 from $p(\sigma^2 | \{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a Gibbs step;

- Step 1.5: Sample α_t from $p(\alpha_t|\{Y_t\}, y_{wh,t}, \psi, \bar{\beta}_1, \beta_2, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a Gibbs step for all t's.

- Step 2: Conditional on ψ, $\bar{\beta}_1$, β_2, σ^2 and α_t's, we obtain the posterior draws of τ_λ, τ_ρ, τ_γ, τ_ϕ and τ_δ.
  - Step 2.1: Sample τ_λ from $p(\tau_\lambda|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a M-H step;
  - Step 2.2: Sample τ_ρ from $p(\tau_\rho|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\gamma, \tau_\phi, \tau_\delta)$ using a M-H step;
  - Step 2.3: Sample τ_γ from $p(\tau_\gamma|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\phi, \tau_\delta)$ using a M-H step;
  - Step 2.4: Sample τ_ϕ from $p(\tau_\phi|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\delta)$ using a M-H step;
  - Step 2.5: Sample $\tau_\delta = (\tau_{\delta|1}, \tau_{\delta|2})'$ from $p(\tau_\delta|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi)$ using a M-H step.

The length of the MCMC sampler is 20000 with a 20% burn-in ratio. We treat the posterior mean of MCMC draws as point estimates for parameters in θ and use the Bayesian 95% credible interval (CI) to check the statistical significance of parameter estimates. We demonstrate the convergence of the MCMC sampler by the trace plots of λ_1, λ_2, ρ_1, ρ_2, ϕ_1 and ϕ_2 in Figure S4. For unknown change points in τ, we simply treat their posterior modes as corresponding point estimates. We plot the posterior distributions of some importation transmissibility parameters, i.e., δ_1, δ_2, $\tau_{\delta|1}$ and $\tau_{\delta|2}$ in Figure (S5).

Computational details of the MCMC algorithm. The MCMC sampler outlined above consists of several M-H steps for $\psi = (\lambda_1, \lambda_2, \rho_1, \rho_2, \mu)'$ and all change points in τ, as well as some Gibbs steps for $\bar{\beta}_1$, β_2, σ^2 and $\{\alpha_t\}$'s. Particularly, we apply the Adaptive Metropolis (AM) algorithm in (12) and (13) to sample ψ. Unlike the standard M-H algorithm, which uses the random walk proposal with a covariance matrix equaling to an identity matrix, the AM algorithm uses the historical MCMC draws to construct the covariance matrix of the proposal distribution. At the gth iteration, the historical draws of ψ is $(\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)})$. The AM proposal suggestd by(12) and (13) is:

$$\Omega(\psi|\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)}) = \begin{cases} \mathcal{N}_5(\psi^{(g-1)}, 0.1^2 I_5/5) & g \le g_0 \\ \mathcal{N}_5(\psi^{(g-1)}, Var_\psi) & g > g_0 \end{cases} \quad [16]$$

where $Var_\psi = (1-\omega)^2 \times 2.38^2 \times Cov(\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)})/5 + \omega^2 \times 0.1^2 I_5/5$. Here $Cov(\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)}) = \frac{1}{g}\sum_{i=0}^{g-1} \psi^{(g)}\psi^{(g)'} - \overline{\psi}^{(g-1)}\overline{\psi}^{(g-1)'}$ is the empirical covariance of $(\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)})$ with $\overline{\psi}^{(g-1)} = \frac{1}{g}\sum_{i=0}^{g-1} \psi^{(g)}$, g_0 represents the length of initial sampling period, and 2.38^2 optimizes the mixing properties of the Metropolis search for the Gaussian proposals(14). Notice that when $g > g_0$, the variance covariance matrix Var_ψ consists of two components, where the second component $\omega^2 \times 0.1^2 I_5/5$ can prevent us from generating some singular variance-covariance matrix due to some problematic values of $(\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)})$. In this paper we follow (13) to set $\omega = 0.05$. In our MCMC sampler, we apply the AM step only to the burn-in draws. After burn-in, we fix the variance-covariance matrix at Var_ψ and use a normal proposal to continue the M-H step for ψ.

Below we list the set of conditional posterior distributions required in the MCMC sampler. To simplify notation, exogenous and pre-determined controls ξ_t, B_t, X_t and C, pre-determined spatial weights matrices $\bar{W}_t$'s, $M_{wh,t}$'s, lagged dependent variables Y_{t-1}'s and $y_{wh,t-1}$'s, as well as initial values $\bar{W}_0$, Y_0, $y_{wh,0}$ are suppressed from the conditional set. For each step, the full conditional is conditioned on the rest of parameters with the most updated values at the current iteration.

Step 1.1 Sample ψ from $p(\psi|\{Y_t\}, \{y_{wh,t}\}, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$;

By Bayes'theorem,

$$p(\psi|\{Y_t\}, \{y_{wh,t}\}, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \propto \pi(\psi) \times f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta).$$

Let g_b be the burn-in period of the MCMC sampler with $g_0 < g_b$.

(**Sub-step 1.1a: burn-in stage**)

For $g < g_b$, generate a new candidate $\tilde{\psi}$ from the AM proposal $\Omega(\psi|\psi^{(0)}, \psi^{(1)}, \cdots, \psi^{(g-1)})$ in [16], which is a symmetric density. Check whether $\tilde{\psi}$ satisfies the stability conditions in [6] and [8]. If not, redraw $\tilde{\psi}$ until it meets those conditions. With acceptance probability equaling to

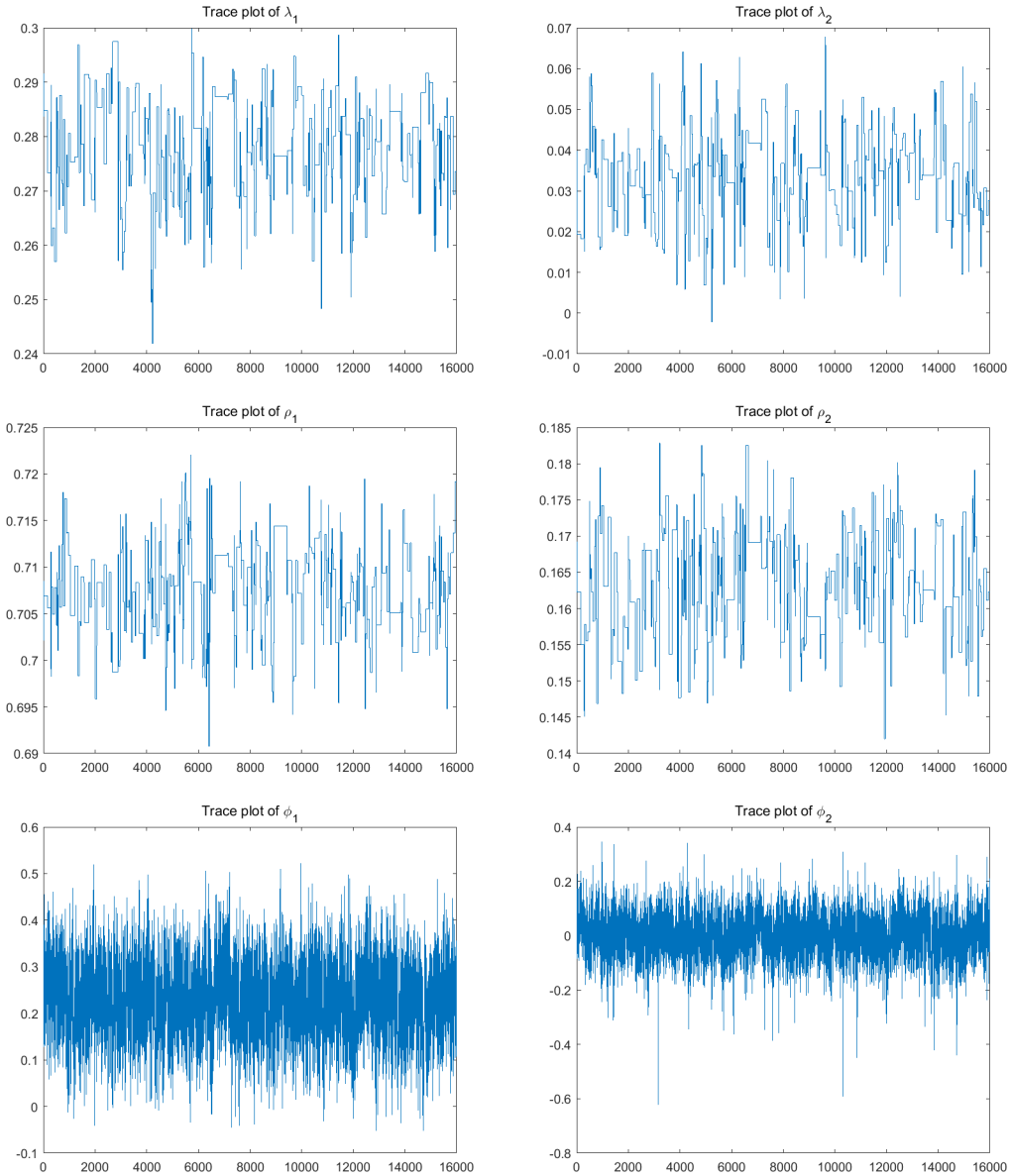
$$Pr(\psi^{(g-1)}, \tilde{\psi}) = \min\{1, \frac{f(\{Y_t\}|\{y_{wh,t}\}, \tilde{\psi}, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)}{f(\{Y_t\}|\{y_{wh,t}\}, \psi^{(g-1)}, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)} \times \frac{\pi(\tilde{\psi})}{\pi(\psi^{(g-1)})}\}$$

set $\psi^{(g)}$ equal to $\tilde{\psi}$, otherwise set it equal to $\psi^{(g-1)}$.

(**Sub-step 1.1b: After burn-in**)

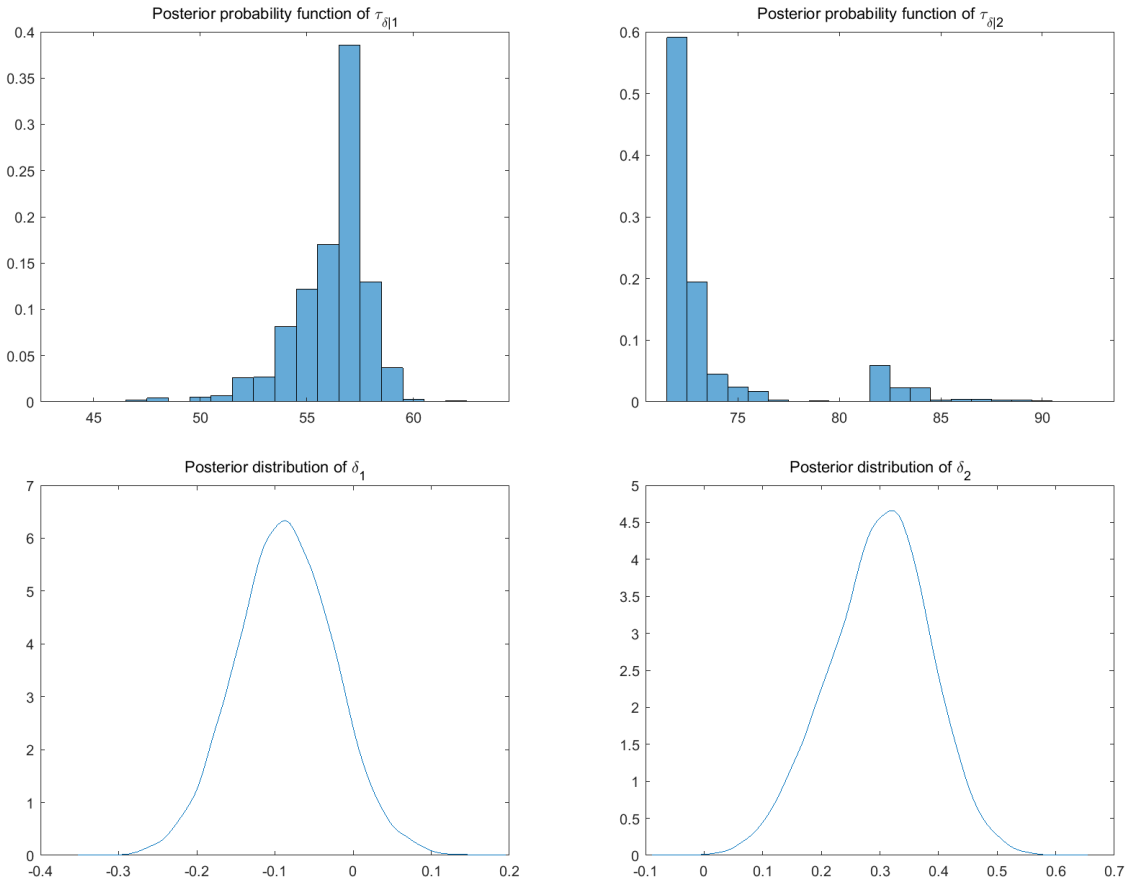
For $g > g_b$, fix the variance-covariance matrix of the proposal at Var_ψ in [16] and generate $\tilde{\psi}$ from $\mathcal{N}(\psi^{(g-1)}, Var_\psi)$. Check whether $\tilde{\psi}$ satisfies the stability conditions in [6] and [8]. If not, redraw $\tilde{\psi}$ until it meets those conditions. With the acceptance probability specified in Sub-step 1.1a, set $\psi^{(g)}$ equal to $\tilde{\psi}$, otherwise set it equal to $\psi^{(g-1)}$.

Step 1.2: Sample $\bar{\beta}_1$ from $p(\bar{\beta}_1|\{Y_t\}, \{y_{wh,t}\}, \psi, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$;



h

Fig. S4. Trace plots of some parameters in the empirical study



h

Fig. S5. Posterior distributions of $\tau_{\delta|1}$, $\tau_{\delta|2}$, δ_1 and δ_2 in the empirical study

Let $\bar{X}_t = (M_{wh,t}y_{wh,t}I(t \le \tau_\gamma), M_{wh,t}y_{wh,t}I(t > \tau_\gamma), B_tI(t \le \tau_{\delta|1}), B_t(1 - I(t \le \tau_{\delta|1}) - I(t > \tau_{\delta|2})), B_tI(t > \tau_{\delta|2}), \xi_tI(t \le \tau_\phi), \xi_tI(t > \tau_\phi), X_t)$ and $H_t(\theta, \tau) = S_t(\lambda, \tau_\lambda)Y_t - \rho(t)Y_{t-1} - \mu\bar{W}_tY_{t-1} - C\beta_2 - l_n\alpha_t - \bar{X}_t\bar{\beta}_1$. By Bayes'theorem,

$$\begin{aligned} p(\bar{\beta}_1|\{Y_t\}, \{y_{wh,t}\}, \psi, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) &\propto \pi(\bar{\beta}_1) \times f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \\ &\propto \exp\left(-\frac{1}{2}(\bar{\beta}_1 - \beta_{10})'B_1^{-1}(\bar{\beta}_1 - \beta_{10})\right) \times \exp(-\frac{1}{2\sigma^2}\sum_{t=1}^{T} H_t(\theta, \tau)'H_t(\theta, \tau)) \\ &\sim \mathcal{N}(T_{\bar{\beta}_1}, \Sigma_{\bar{\beta}_1}), \end{aligned}$$

where

$$\begin{aligned} \Sigma_{\bar{\beta}_1} &= (B_1^{-1} + \sum_{t=1}^{T} \frac{\bar{X}_t'\bar{X}_t}{\sigma^2})^{-1}, \\ T_{\bar{\beta}_1} &= \Sigma_{\bar{\beta}_1}(B_1^{-1}\bar{\beta}_{10} + \sum_{t=1}^{T} \frac{\bar{X}_t'(S_t(\lambda, \tau_\lambda)Y_t - \rho(t)Y_{t-1} - \mu\bar{W}_tY_{t-1} - C\beta_2 - l_n\alpha_t)}{\sigma^2}). \end{aligned}$$

Step 1.3: Sample β_2 from $p(\beta_2|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$;
By Bayes'theorem,

$$\begin{aligned} p(\beta_2|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) &\propto \pi(\beta_2) \times f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \\ &\propto \exp(-\frac{1}{2}(\beta_2 - \beta_{20})'B_2^{-1}(\beta_2 - \beta_{20})) \times \exp(-\frac{1}{2\sigma^2}\sum_{t=1}^{T} H_t(\theta, \tau)'H_t(\theta, \tau)) \\ &\sim \mathcal{N}(T_{\beta_2}, \Sigma_{\beta_2}), \end{aligned}$$

where

$$\begin{aligned} \Sigma_{\beta_2} &= (B_2^{-1} + T\frac{C'C}{\sigma^2})^{-1}, \\ T_{\beta_2} &= \Sigma_{\beta_2}(B_2^{-1}\beta_{20} + \sum_{t=1}^{T} \frac{C'(S_t(\lambda, \tau_\lambda)Y_t - \rho(t)Y_{t-1} - \mu\bar{W}_tY_{t-1} - \bar{X}_t\bar{\beta}_1 - l_n\alpha_t)}{\sigma^2}). \end{aligned}$$

Step 1.4: Sample σ^2 from $p(\sigma^2|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$;
By Bayes's theorem,

$$\begin{aligned} p(\sigma^2|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) &\propto \pi(\sigma^2) \times f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \\ &\propto (\sigma^2)^{-\frac{a}{2}+1}\exp(-\frac{b}{2\sigma^2}) \times (\sigma^2)^{-\frac{nT}{2}}\exp(-\frac{1}{2\sigma^2}\sum_{t=1}^{T} H_t(\theta, \tau)'H_t(\theta, \tau)) \\ &\sim \mathcal{IG}(\frac{a_p}{2}, \frac{b_p}{2}), \end{aligned}$$

where $a_p = a + nT$ and $b_p = b + \sum_{t=1}^{T} H_t(\theta, \tau)'H_t(\theta, \tau)$.
Step 1.5: Sample α_t from $p(\alpha_t|\{Y_t\}, y_{wh,t}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$for $t = 1, 2, \cdots, T$;
By Bayes's theorem,

$$\begin{aligned} p(\alpha_t|\{Y_t\}, y_{wh,t}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) &\propto \pi(\alpha_t) \times f(Y_t|y_{wh,t}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \alpha_t, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \\ &\propto \exp(-\frac{(\alpha_t - \alpha_0)^2}{2\sigma_\alpha^2}) \times \exp(-\frac{1}{2\sigma^2}H_t(\theta, \tau)'H_t(\theta, \tau)) \\ &\sim \mathcal{N}(T_{\alpha_t}, \Sigma_{\alpha_t}), \end{aligned}$$

with $\Sigma_{\alpha_t} = (\frac{1}{\sigma_\alpha^2} + \frac{n}{\sigma^2})^{-1}$ and $T_{\alpha_t} = \Sigma_{\alpha_t}[\frac{\alpha_0}{\sigma_\alpha^2} + \frac{l_n'(S_t(\lambda,\tau_\lambda)Y_t - \rho(t)Y_{t-1} - \mu\bar{W}_tY_{t-1} - \bar{X}_t\bar{\beta}_1 - C\beta_2)}{\sigma^2}]$.
Step 2.1: Sample τ_λ from $p(\tau_\lambda|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)$
By Bayes'theorem,

$$p(\tau_\lambda|\{Y_t\}, \{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta) \propto \pi(\tau_\lambda) \times f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta).$$

At the gth iteration, generate a new candidate $\tilde{\tau}_\lambda$ from a uniform proposal $\mathcal{U}(14, T-1)$, which is independent of $\tau_\lambda^{(g-1)}$. With acceptance probability

$$Pr(\tau_\lambda^{(g-1)}, \tilde{\tau}_\lambda) = \min\{1, \frac{f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tilde{\tau_\lambda}, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)}{f(\{Y_t\}|\{y_{wh,t}\}, \psi, \bar{\beta}_1, \beta_2, \sigma^2, \{\alpha_t\}, \tau_\lambda^{(g-1)}, \tau_\rho, \tau_\gamma, \tau_\phi, \tau_\delta)} \times \frac{\pi(\tilde{\tau}_\lambda)}{\pi(\tau_\lambda^{(g-1)})}\}$$

set $\tau_\lambda^{(g)}$ equal to $\tilde{\tau}_\lambda$, otherwise stay at $\tau_\lambda^{(g-1)}$.

Step 2.2: Sample τ_ρ from $p(\tau_\rho|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\gamma,\tau_\phi,\tau_\delta)$
By Bayes'theorem,

$$p(\tau_\rho|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\gamma,\tau_\phi,\tau_\delta) \propto \pi(\tau_\rho)\times f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tau_\delta).$$

At the gth iteration, generate a new candidate $\tilde{\tau}_\rho$ from a uniform proposal $\mathcal{U}(14,T-1)$, which is independent of $\tau_\rho^{(g-1)}$. With acceptance probability

$$Pr(\tau_\rho^{(g-1)},\tilde{\tau}_\rho)=\min\{1,\frac{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tilde{\tau}_\rho,\tau_\gamma,\tau_\phi,\tau_\delta)}{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho^{(g-1)},\tau_\gamma,\tau_\phi,\tau_\delta)}\times\frac{\pi(\tilde{\tau}_\rho)}{\pi(\tau_\rho^{(g-1)})}\}$$

set $\tau_\rho^{(g)}$ equal to $\tilde{\tau}_\rho$, otherwise remain at $\tau_\rho^{(g-1)}$.

Step 2.3: Sample τ_γ from $p(\tau_\gamma|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\phi,\tau_\delta)$
By Bayes'theorem,

$$p(\tau_\gamma|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\phi,\tau_\delta) \propto \pi(\tau_\gamma)\times f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tau_\delta).$$

At the gth iteration, generate a new candidate $\tilde{\tau}_\gamma$ from a uniform proposal $\mathcal{U}(14,T-1)$, which is independent of $\tau_\gamma^{(g-1)}$. With acceptance probability

$$Pr(\tau_\gamma^{(g-1)},\tilde{\tau}_\gamma)=\min\{1,\frac{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tilde{\tau}_\gamma,\tau_\phi,\tau_\delta)}{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma^{(g-1)},\tau_\phi,\tau_\delta)}\times\frac{\pi(\tilde{\tau}_\gamma)}{\pi(\tau_\gamma^{(g-1)})}\}$$

set $\tau_\gamma^{(g)}$ equal to $\tilde{\tau}_\gamma$, otherwise remain at $\tau_\gamma^{(g-1)}$.

Step 2.4: Sample τ_ϕ from $p(\tau_\phi|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\delta)$
By Bayes'theorem,

$$p(\tau_\phi|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\phi,\tau_\delta) \propto \pi(\tau_\phi)\times f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tau_\delta).$$

At the gth iteration, generate a new candidate $\tilde{\tau}_\phi$ from a uniform proposal $\mathcal{U}(14,40)$, which is independent of $\tau_\phi^{(g-1)}$. With acceptance probability

$$Pr(\tau_\phi^{(g-1)},\tilde{\tau}_\phi)=\min\{1,\frac{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tilde{\tau}_\phi,\tau_\delta)}{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi^{(g-1)},\tau_\delta)}\times\frac{\pi(\tilde{\tau}_\phi)}{\pi(\tau_\phi^{(g-1)})}\}$$

set $\tau_\phi^{(g)}$ equal to $\tilde{\tau}_\phi$, otherwise remain at $\tau_\phi^{(g-1)}$.

Step 2.5: Sample $\tau_\delta=(\tau_{\delta|1},\tau_{\delta|2})'$ from $p(\tau_\delta|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi)$
By Bayes'theorem,

$$p(\tau_\delta|\{Y_t\},\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi) \propto \pi(\tau_{\delta|1})\times\pi(\tau_{\delta|2})\times f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tau_\delta).$$

At the gth iteration, generate new candidates $\tilde{\tau}_{\delta|1}$ and $\tilde{\tau}_{\delta|2}$ from, respectively, uniform proposals $\mathcal{U}(40,70)$ and $\mathcal{U}(71,T-1)$, and set $\tilde{\tau}_\delta=(\tilde{\tau}_{\delta|1},\tilde{\tau}_{\delta|2})'$. With acceptance probability

$$Pr(\tau_\delta^{(g-1)},\tilde{\tau}_\delta)=\min\{1,\frac{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tilde{\tau}_\delta)}{f(\{Y_t\}|\{y_{wh,t}\},\psi,\bar{\beta}_1,\beta_2,\sigma^2,\{\alpha_t\},\tau_\lambda,\tau_\rho,\tau_\gamma,\tau_\phi,\tau_\delta^{(g-1)})}\times\frac{\pi(\tilde{\tau}_{\delta|1})}{\pi(\tau_{\delta|1}^{(g-1)})}\times\frac{\pi(\tilde{\tau}_{\delta|2})}{\pi(\tau_{\delta|2}^{(g-1)})}\}$$

set $\tau_\delta^{(g)}$ equal to $\tilde{\tau}_\delta$, otherwise remain at $\tau_\delta^{(g-1)}$.

Monte Carlo simulation. We further conduct some simulation studies to show that our MCMC sampler can indeed recover the true values of time-varying parameters and identify different change points. We consider the following SDPD model:

$$Y_t=\lambda(t)\bar{W}_tY_t+\rho(t)Y_{t-1}+\mu\bar{W}_tY_{t-1}+B_t\delta(t)+X_{nt}\beta+F+l_n\alpha_t+U_t,\ t=1,2,\cdots,T.$$

where we only allow $\lambda(t)$, $\rho(t)$ and $\delta(t)$ to have their own change points:

$$\begin{aligned}\lambda(t)&=\lambda_1\times I(t\le\tau_\lambda)+\lambda_2\times I(t>\tau_\lambda),\\ \rho(t)&=\rho_1\times I(t\le\tau_\rho)+\rho_2\times I(t>\tau_\rho),\\ \delta(t)&=\delta_1\times I(t\le\tau_\delta)+\delta_2\times I(t>\tau_\delta).\end{aligned}$$

We set the number of cross-sectional units (cities) as $n = 300$ and the time length as $T = 40$ throughout the simulation study. The true values of parameters and their change-points are set as $\lambda_1 = 0.4$, $\lambda_2 = 0.06$, $\rho_1 = 0.2$, $\rho_2 = 0.03$, $\delta_1 = 0$, $\delta_2 = 1$, $\beta = 1$, $\tau_\lambda = 10$, $\tau_\rho = 23$ and $\tau_\delta = 30$. $B_{nt} = (b_{1t}, b_{2t}, \cdots, b_{nt})'$ and $X_t = (x_{1t}, x_{2t}, \cdots, x_{nt})'$ are both $n \times 1$ column vector of exogenous regressors, with $B_t \sim \mathcal{N}(0, 2I_n)$ and $X_t \sim \mathcal{N}(0, 2I_n)$ for $t = 1, 2, \cdots, T$. U_t is the $n \times 1$ vector of disturbances with $U_t \sim \mathcal{N}(0, \sigma^2 I_n)$ and $\sigma^2 = 1$. $F = (f_1, f_2, \cdots, f_n)'$ are the $n \times 1$ vector of individual effects and are generated according to the Mundlak's approach in (15): $f_i = 2\bar{b}_i + 2\bar{x}_i + \epsilon_i$, $\epsilon_i \sim \mathcal{N}(0, 2)$, with $\bar{b}_i$ and $\bar{x}_i$ representing, respectively, the empirical means of b_{it}'s and x_{it}'s over time. α_t is the scalar time effect with $\alpha_t \sim \mathcal{N}(0, 1)$ for $t = 1, 2, \cdots, T$. Additionally, we assume the initial values of Y_0 are generated by

$$Y_0 = B_0 + X_0 + F + \nu_0, \ \nu_0 \sim \mathcal{N}(0, 2I_n),$$

with B_0 and X_0 distributed as $\mathcal{N}(0, 2I_n)$. We construct the spatial weights matrices $\bar{W}_t$'s through the following steps:

- For each cross-sectional unit (city) i on the Euclidean plane $\mathcal{R}^2$. Generate its coordinates xc_i and yc_i from $\chi(3)$, the chi-square distribution with three degrees of freedom, to represent its geographical location on $\mathcal{R}^2$.

- Compute the geographical distance $d(i, j)$ between i and j based on their coordinates and define a $n \times n$ non row-normalized $0 - 1$ indicator matrix $W^u = [w^u_{ij}]$, based upon $d(i, j)$'s. Specifically, for each i, find the nearest m neighbors based on $d(i, j)$, $j \neq i$ and denote the corresponding $w^u_{ij} = 1$, otherwise $w^u_{ij} = 0$. We consider a sparse W^u with $m = 5$ and a denser W^u with $m = 50$.

- Simulate the "income" of city i at the initial period via $z_{i0} \sim \mathcal{N}(0, 3)$. Then generate time-varying incomes of different periods for i by the following stationary AR(1) process:

$$z_{it} = 0.6 \times z_{i,t-1} + \nu_{it}, \ \nu_{it} \sim \mathcal{N}(0, 1), \ t = 1, 2, \cdots, T.$$

- Simulate q_{ijt}, the travel flow from city j to city i at time t by the following gravity equation

$$q_{ijt} = 2 + 2 \times z_{it} + 2 \times z_{jt} + \eta_{1i} + \eta_{2j} + v_{ijt}, \ v_{ijt} \sim \mathcal{N}(0, 2), j \neq i$$

  where $\eta_{1i} \sim \mathcal{N}(2, 2)$ is the fixed effect for destination city i and $\eta_{2i} \sim \mathcal{N}(1, 2)$ is the fixed effect for origin city j. Then specify a $n \times n$ time-varying spatial weights matrix $W^e_t = [w^e_{ijt}]$, with $w^e_{ijt} = \exp(q_{ijt}) + \exp(q_{jit})$

- Set $W_t = [w_{ijt}] = W^u \ o \ W^e_t$ as Hadamard product, i.e., $w_{ijt} = w^u_{ij} \times w^e_{ijt}$, and row-normalize W_t to obtain $\bar{W}_t$.

We implement a simplified version of the Bayesian MCMC sampler outlined in previous subsections. We specify the priors of parameters as in [14]. In particular, we assume independent uniform priors for the unknown change points as $\tau_\lambda \sim \mathcal{U}(1, 39)$, $\tau_\rho \sim \mathcal{U}(1, 39)$ and $\tau_\delta \sim \mathcal{U}(1, 39)$. We run 100 Markov chains (repetitions) with a 20% burn-in ratio, and treat the posterior mean of parameters in each repetition as point estimates. The mean and standard deviation (S.D) across 100 experiments are reported in the following Table S3. We also treat the posterior modes of τ_λ, τ_ρ and τ_δ as their point estimates in each repetition, and report the empirical frequencies that those posterior modes match the true change points in the 1000 experiments. We find that the Bayesian estimates are close to the true values of parameters. And our MCMC sampler can successfully identify true change points of $\lambda(t)$, $\rho(t)$ and $\delta(t)$ in each of the 1000 repetitions in Table S3.

Table S3. Monte Carlo simulation over 1000 repetitions.

		Sparse $\bar{W}_t$		Dense $\bar{W}_t$	
		BE	S.D	BE	S.D
$n = 300$	λ_1	0.401	0.008	0.401	0.012
$T = 40$	λ_2	0.060	0.005	0.059	0.007
	ρ_1	0.194	0.005	0.194	0.005
	ρ_2	0.024	0.005	0.024	0.006
	μ	-0.099	0.005	-0.099	0.006
	δ_1	0.001	0.007	0.000	0.007
	δ_2	0.999	0.011	1.000	0.011
	β	0.999	0.006	0.999	0.007
		FCI with Sparse $\bar{W}_t$		FCI with dense $\bar{W}_t$	
	τ_λ	100%		100%	
	τ_ρ	100%		100%	
	τ_δ	100%		100%	

Note: BE: Bayesian estimates; S.D: Standard Deviation; FCI: Frequency of correctly identifying change-points; $\lambda_1 = 0.4$, $\lambda_2 = 0.06$, $\rho_1 = 0.2$, $\rho_2 = 0.03$, $\delta_1 = 0$, $\delta_2 = 1$, $\beta_2 = 1$, $\tau_\lambda = 10$, $\tau_\rho = 23$ and $\tau_\delta = 30$.

Computation of counterfactual and additional results. Our counterfactual computation relies on the reduced form of the SDPD model in [2] for cities except Wuhan:

$$Y_t = S_t^{-1}(\lambda(t))[(\rho(t)I_n + \mu W_{t-1})Y_{t-1} + \gamma(t)M_{wh,t}y_{wh,t} + B_t\delta(t) + \xi_t\phi(t) + X_t\beta_1 + C\beta_2 + l_n\alpha_t + U_t],$$

with $S_t(\lambda(t)) = I_n - \lambda(t)\bar{W}_t$. Consider the scenario of isolated effect on importation transmissibility where $\delta(t)$ remains at its high level since its emergence, i.e., $\delta(t) = \delta_2$ after the second change point $\tau_{\delta|2}$. Let $\delta_2^{(g)}$ be the MCMC draws of δ_2 at the gth iteration of the Markov chain. We first estimate the model via the MCMC algorithm outlined above and obtain the estimated residual $\hat{U}_t^{(g)}$ at the gth iteration. Assuming the initial newly confirmed cases Y_0 and domestic travel flow network $\bar{W}_0$ are exogenously given, the counterfactual newly confirmed cases $\overline{Y}_t^{(g)}$'s at the gth iteration can be computed recursively via:

$$\begin{aligned}
\overline{Y}_1^{(g)} &= S_1^{-1}(\lambda^{(g)}(1))[(\rho^{(g)}(1)I_n + \mu^{(g)}\bar{W}_0)Y_0 + \gamma^{(g)}(1)M_{wh,1}y_{wh,1} + \xi_1\phi^{(g)}(1) + B_1\delta^{(g)}(1) + X_1\beta_1^{(g)} + C\beta_2^{(g)} + l_n\alpha_1^{(g)} + \hat{U}_1^{(g)}],\\
\overline{Y}_2^{(g)} &= S_2^{-1}(\lambda^{(g)}(2))[(\rho^{(g)}(2)I_n + \mu^{(g)}\bar{W}_1)\overline{Y}_1^{(g)} + \gamma^{(g)}(2)M_{wh,2}y_{wh,2} + \xi_2\phi^{(g)}(2) + B_2\delta^{(g)}(2) + X_2\beta_1^{(g)} + C\beta_2^{(g)} + l_n\alpha_2^{(g)} + \hat{U}_2^{(g)}],\\
\overline{Y}_3^{(g)} &= S_3^{-1}(\lambda^{(g)}(3))[(\rho^{(g)}(3)I_n + \mu^{(g)}\bar{W}_2)\overline{Y}_2^{(g)} + \gamma^{(g)}(3)M_{wh,3}y_{wh,3} + \xi_3\phi^{(g)}(3) + B_3\delta^{(g)}(3) + X_3\beta_1^{(g)} + C\beta_2^{(g)} + l_n\alpha_3^{(g)} + \hat{U}_3^{(g)}],\\
&\vdots
\end{aligned}$$

where $\delta(t)^{(g)} = \delta_2^{(g)}$ as long as $t > \tau_{\delta|2}^{(g)}$ at the gth iteration. With posterior draws $\overline{Y}_t^{(g)}$'s, we may calculate the posterior draws of the national counterfactual newly confirmed case by adding up the counterfactual cases of each city except Wuhan on each date. The corresponding posterior mean, as well as the upper bound (UB) and lower bound (LB) of the Bayesian 95% credible interval for the national counterfactual value can be easily derived. The posterior draws of the national cumulative cases excluding Wuhan, as well as counterfactual of other scenarios, can also be derived in a similar manner.

We plot the counterfactual daily newly COVID-19 confirmed cases in Figures S6 to S8, corresponding to the scenarios described in the main text. Figure S6 plots newly confirmed cases by the end of April in six scenarios when importation and domestic transmission can be isolated. Similar to the counterfactual on cumulative cases, importation transmissibility and international travel flow have small effects on daily newly confirmed cases (panels a1-a3), when domestic transmissibility became significantly weaker before importation took place. For instance, the counterfactual newly confirmed case with high importation transmissibility and unrestricted international travel flow would come to about 116 by the end of April, compared with a factual number of 5 (panel a3). On the other hand, when we run similar scenarios for domestic transmission mechanisms holding importation mechanisms as factual, we find a stronger influence from domestic transmission. For instance, the newly confirmed case per day by the end of April would have been 15800 under both high domestic transmissibility and unrestricted domestic flow (panel b3), compared with the factual number of 5 (solid line).

Figure S7 depicts the counterfactual on newly confirmed case per day where importation with domestic transmission are integrated, and domestic transmissibility was first contained as in the reality and then resurged after importation emerged. We use unrestricted international flight and domestic travel flow throughout the sample period and simulate four scenarios by changing the domestic transmissibility parameters to 25%, 50%, 75%, 100% of their pre-containment levels (panels a-d respectively) after the emergence of importation. Similar to the counterfactual on cumulative cases, the counterfactual newly confirmed cases exceed the factual number even before the importation emerged, then would evolve differently depending on the degree of resurgence of domestic transmissibility. Particularly, when the domestic transmissibility resurged to 100% of its pre-containment level, the newly confirmed cases would have been almost 3730 times of the factual cases at the end of the sample period at the height of 18650 per day (panel d). So the importation risk could be more severe if domestic transmissibility fully resurged to its pre-containment level.

Figure S8 plots the counterfactual on newly confirmed case when the domestic transmission was only partially suppressed before importation arose. We simulate the newly confirmed cases in scenarios where the domestic transmissibility parameters changed to 25%, 50%, 75% and 100% of their initial levels after the respective change points (panels a-d, respectively). We still find that counterfactual newly confirmed cases exceeded the factual cases even before the importation arose, yet the gap is wider and takes longer to close than in Figure S7. Specifically, in the worst scenario where domestic transmissibility parameters remained at 100% of their initial levels (panel d), the counterfactual newly confirmed case at the end of April would come to 22690, compared with a factual number of 5. This suggests that importation would have had much larger influence on newly confirmed cases when domestic transmission was not under control.

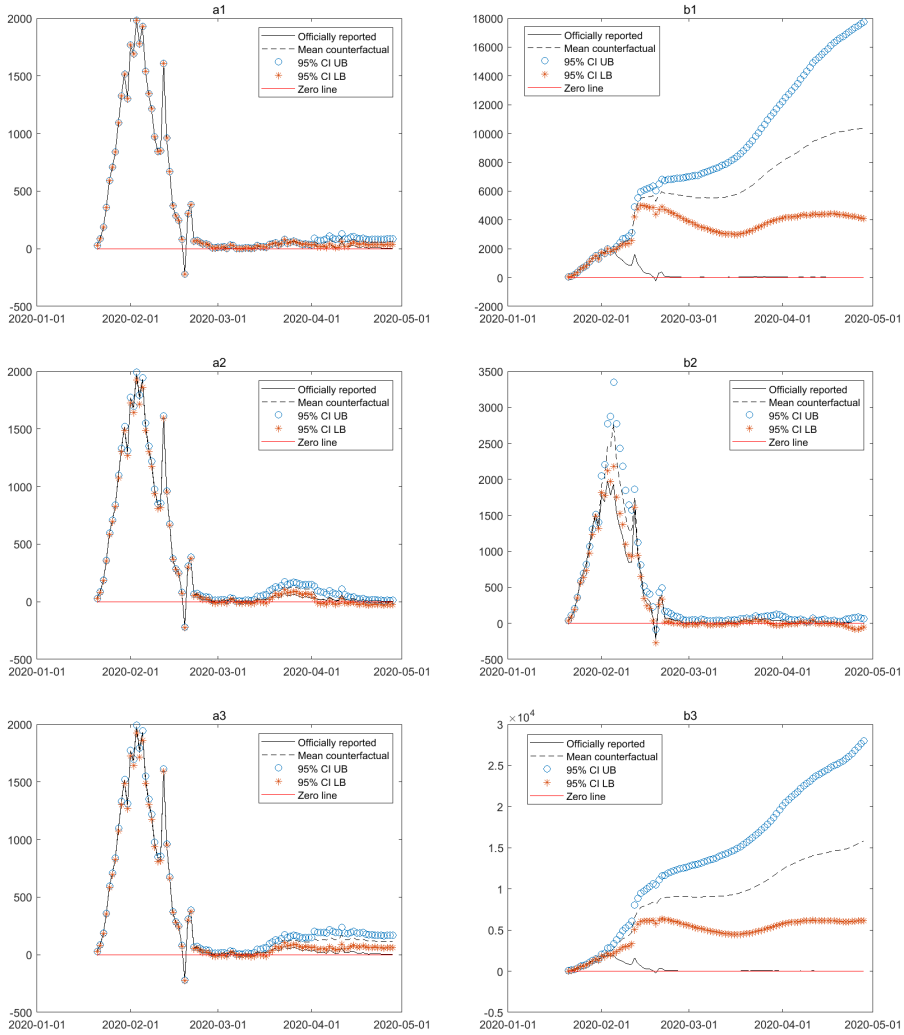


Fig. S6. Simulated newly confirmed cases outside of Wuhan in the scenario when importation and domestic transmission and travel flows can be isolated. Notes: In panel a(b)1, the importation (domestic) transmissibility remained high since its emergence. In panel a(b)2, the international (domestic) travel flow was unrestricted at the 2019 level. In panel a(b)3, both the importation (domestic) transmissibility remained high and the international (domestic) travel flow was unrestricted.

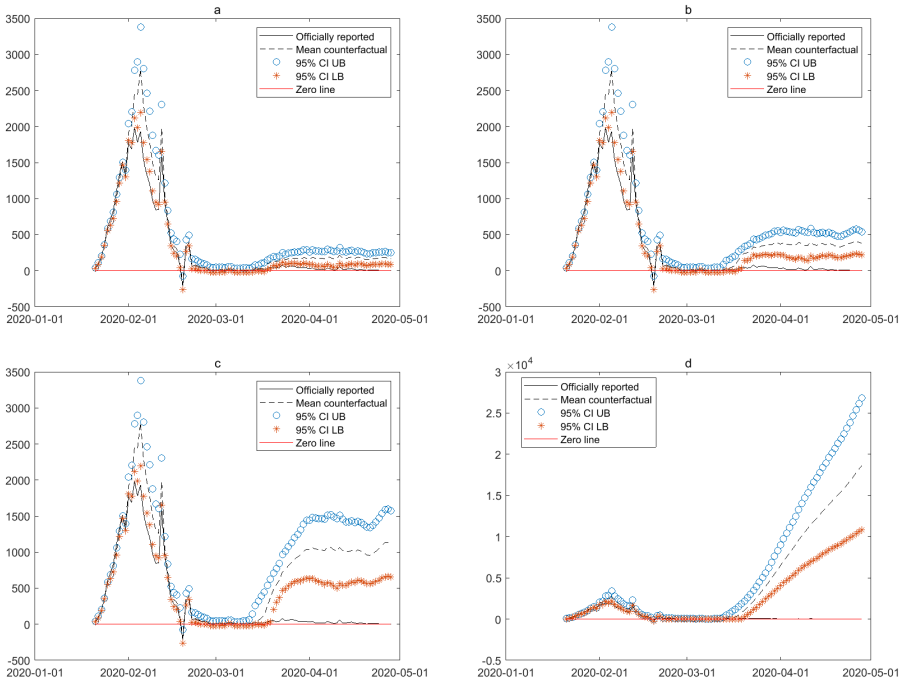


Fig. S7. Simulated newly confirmed cases outside of Wuhan in the scenario when domestic transmission resurged after importation. Notes: In panels a-d, the domestic transmissibility parameters changed to 25%, 50%, 75%, 100% of its pre-containment levels after importation emerged, respectively, while the importation transmissibility remained high since its emergence and international and domestic travel flows were both unrestricted.

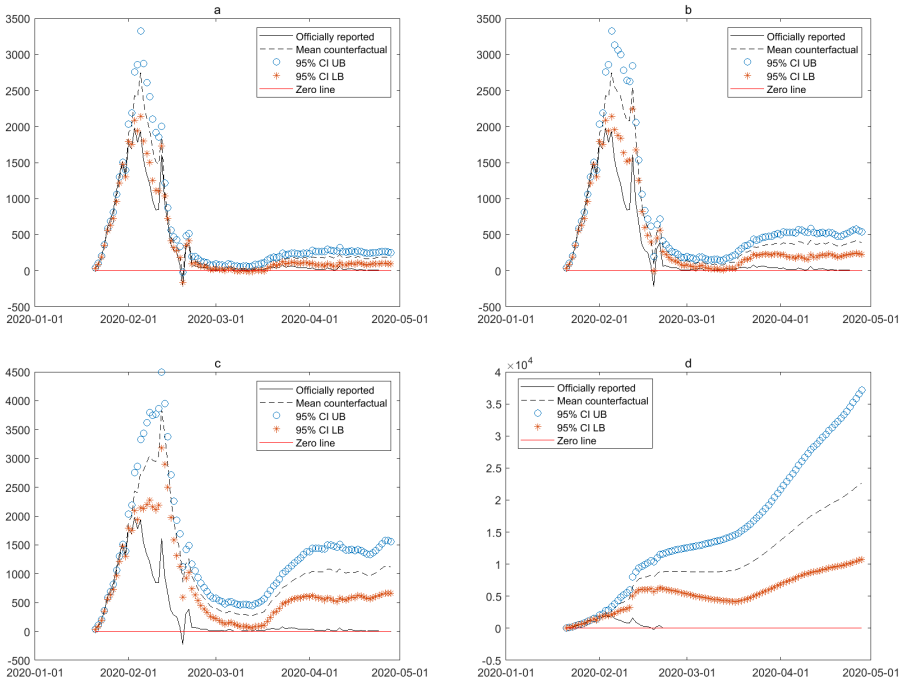


Fig. S8. Simulated newly confirmed cases outside of Wuhan in the scenario when domestic transmissibility was not suppressed before importation. Notes: In panels a-d, the domestic transmissibility parameters changed to 25%, 50%, 75% and 100% of their initial levels after the respective change points, respectively, while the importation transmissibility remained high since its emergence and international and domestic travel flows were both unrestricted.

Chinese COVID-19 Policies Chronicle

January

- Jan 20, the human-to-human transmission was first confirmed in China.
- Jan 23, Wuhan was placed under lockdown with traffic bans for all residents
- Jan 25, 27 out of 31 provinces launched the highest level response to public health emergencies and the Central Government Leadership Group for Epidemic Response was established.
- Jan 26, China State Council extended the Spring Festival holiday to February 2.
- Jan 27, China Ministry of Education postponed the start of the 2020 spring semester until further notice.
- Jan 28, all cities in Hubei province were under complete lockdown with public transit shutdowns, strict limits on the inflows into the city, outflows out of the city, and within-city population movements (16).
- Jan 29, all provinces launched the highest level response to public health emergencies.

February

- Feb 2-Feb 4, seven cities in other provinces (Wenzhou, Hangzhou, Ningbo, Fuzhou, Zhengzhou, Zhumadian, Haerbin) adopted the partial shutdown strategy (16).
- Feb 6, 81 or over 1/5 cities under lockdown as of Feb 6.
- Since Feb 12, there were 115 or about 1/3 cities with different levels of lockdown policies. In addition, closed community management was in effect in 234 cities or over 2/3 of China and family outdoor restrictions were in place in 108 cities or 1/3 of China (2).
- Feb 18, the last day of the Spring Festival transportation.
- Feb 20, all provinces except Hubei resumed within-province long-distance shuttle services and 26 out of 31 provinces resumed shuttle services across provinces.
- Feb 26, the new daily total cases overseas surpassed that of China for the first time. Thirteen provinces downgraded their levels of responses to covid-19 while Hubei cities remained under lockdown.

March

- Mar 1, Chinese Customs required all international entrants to submit COVID-19 related health information.
- Mar 5, out of the 17 daily new cases in China, 16 were imported from other countries.
- Mar 10, last makeshift health facility in Wuhan was closed and President Xi Jinping visited Wuhan.
- Mar 11, the WHO declared covid-19 as a global pandemic.
- Mar 13, President Trump declared national emergency in the U.S. Cities in Hubei other than Wuhan became mid to low risk areas.
- Mar 14, Spain announced lockdown.
- Mar 17, Europe barricaded borders to slow covid-19.
- Mar 18, all new cases in China were imported for the first time. U.S. and Canada agreed to close borders for 'nonessential traffic'.
- Mar 19, thirty-four percent of imported cases in China so far occurred in Beijing. Australia and New Zealand banned non-residents from entry. On the same day, Italy surpassed China as the country with the highest coronavirus death toll.
- Mar 22, Guangzhou and Shanghai started testing all customs entrants.
- Mar 23, Shenzhen and Beijing started testing all customs entrants. All international flights to Beijing had to stop first at other 12 Chinese cities.
- Mar 25, all non-Wuhan Hubei cities unlocked and were allowed to travel to other provinces. The covid-19 risk decreased to middle level and partial public transits resumed in Wuhan.
- Mar 27 Guangzhou required collective quarantine for all customs entrants. The central government required collective quarantine for all international entrants through inland customs.

- Mar 28, ban on foreigners to enter China took effect. Shanghai required collective quarantine for all customs entrants.
- Mar 29, "five one" policy was implemented which allowed only one flight per airline per week between China and a foreign country. The capacity per flight could not exceed 75%.
- Mar 31, the Chinese central government announced that domestic transmission was stemmed successfully. The Ministry of Education required all students to wear masks in classrooms once schools reopened and delayed the national college entrance examination by one month.

April

- Apr 1, China began to test and collectively quarantine all international entrants.
- Apr 5, the WHO announced that the pandemic in China had entered mitigation phase from containment phase. On the same day, half of the imported cases were from Heilongjiang and the Suifenhe customs had closed due to overwhelmed testing and quarantine capacities.
- Apr 7, China required international entrants from 26 countries to report covid-19 related health and travel histories for 14 days before boarding the airplanes to China.
- Apr 8, flights to and out of Wuhan resumed and Wuhan unlocked.
- Apr 13, a random testing of 143,056 Wuhan residents who returned to work showed that only 113 individuals were positive for covid-19, or 0.08% of the tested population.
- Apr 15, the last assisting medical team from Beijing Xiehe Hospital left Wuhan.
- Apr 19, Beijing started mandatory testing eight categories of people who were at higher risk of contracting covid-19 including (1) covid-19 patients and close contacts (2) patients with fever (3) inpatients who had respiratory or epidemic diseases (4) international entrants (5) residents returned from Wuhan and were about to leave quarantine (6) government officials that returned from places outside of Beijing (7) individuals who came to Beijing and stayed in hotels (8) students and teachers who returned to Beijing from places outside of Beijing.
- Apr 21, many provinces determined to postpone the high school entrance examination dates to July.
- Apr 22, due to 18 confirmed covid-19 cases in Harbin, Harbin started closed community management again.
- Apr 23, Heilongjiang issued within-city health barcodes to every resident to limit the mobility of confirmed, asymptotic, suspected cases and close contacts outside their resident cities. the Central Government Leadership Group for Epidemic Response emphasized that precautions should be taken during the Labor Day holidays and improvement of border controls were needed.
- Apr 25, Yangtze river cruise ships in Wuhan resumed service.
- Apr 27, the Central Government Leadership Group for Epidemic Response left Hubei and returned to Beijing. Italy announced return to work plans in early May.

Robustness checks

In this section we conduct various robustness checks to complement the main results from the SDPD specification, including using average domestic travel flow centered at 7-day and 14-day lags, using a spatial count panel model since total newly confirmed cases for all cities become non-negative, investigating spatial heterogeneity of cities in Heilongjiang province and cities with quarantine facilities, studying the effect of the abroad infection index on domestic and imported cases, incorporating city inflow intensity indices in the travel flow network, relaxing row normalization of travel flow network, and relaxing the stability condition on the within-city transmission coefficient.

Using average travel flow centered at 7-day and 14-day lags . For the spatial weights matrix, the abroad infection index, and the inflow index used in the paper, we smooth the travel flow/importation risk by the five-day average and then lagged the smoothed value by one incubation period. We pick a five-day lag as the center of the average, given the evidence of an average of 5-day incubation period (1). To alleviate the concern about the uncertainty of incubation period and reporting delay, we also conduct some robustness checks using a 7-day lag centered average and a 14-day lag centered average, respectively. For instance, we construct the five-day moving average spatial weights at a 7-day or 14-day lag as

$$\tilde{w}_{ijt} = \frac{\sum_{l=5}^{9} p_{ij,t-l}}{5} \quad \text{or} \quad \frac{\sum_{l=12}^{16} p_{ij,t-l}}{5}.$$

We then specify the spatial weights matrices $\bar{W}_t = [\bar{w}_{ijt}]$ in the SDPD model via row-normalized moving average weights $\bar{w}_{ijt} = \frac{\tilde{w}_{ijt}}{\sum_j \tilde{w}_{ijt}}$ to derive the travel flow network among 283 cities (excluding Wuhan), as well as the flow from Wuhan to other cities. The 7-day lag centered and 14-day lag centered average of the abroad index and the inflow index can be constructed in a similar manner. The estimation results are provided in Tables S4. We find that the parameter and change-point estimates are similar to the main results using the 5 days lagged centered spatial weights. Specifically, the domestic transmission mechanisms, such as cross and within city transmissions, has decreased largely by mid-February, while the influence from the abroad infection index begins to take place from mid-March and quickly turns insignificant by mid-April.

Table S4. SDPD Estimations with alternative lags

		7 days lagged centered	14 days lagged centered
Cross-city transmission	λ_1	0.262	0.199
	S.D	(0.008)	(0.015)
	95% CI	[0.244, 0.278]	[0.167, 0.227]
	λ_2	0.037	0.055
	S.D	(0.012)	(0.013)
	95% CI	[0.012, 0.061]	[0.026, 0.078]
Wuhan influence	γ_1	0.024	0.034
	S.D	(0.003)	(0.000)
	95% CI	[0.023, 0.025]	[0.033, 0.035]
	γ_2	-0.011	-0.009
	S.D	(0.002)	(0.002)
	95% CI	[−0.016, −0.007]	[−0.014, −0.005]
Within-city transmission	ρ_1	0.726	0.711
	S.D	(0.005)	(0.005)
	95% CI	[0.715, 0.735]	[0.701, 0.719]
	ρ_2	0.176	0.163
	S.D	(0.007)	(0.008)
	95% CI	[0.164, 0.192]	[0.148, 0.178]
Abroad index (log)	δ_1	-0.075	-0.083
	S.D	(0.066)	(0.088)
	95% CI	[−0.206, 0.053]	[−0.266, 0.078]
	δ_2	0.264	0.219
	S.D	(0.096)	(0.087)
	95% CI	[0.094, 0.453]	[0.063, 0.407]
	δ_3	-0.066	-0.079
	S.D	(0.058)	(0.065)
	95% CI	[−0.188, 0.041]	[−0.208, 0.044]
Inflow index	ϕ_1	0.261	0.431
	S.D	(0.076)	(0.063)
	95% CI	[0.115, 0.411]	[0.307, 0.553]
	ϕ_2	0.073	0.169
	S.D	(0.084)	(0.077)
	95% CI	[−0.097, 0.226]	[0.017, 0.340]
Change-points	τ_λ	16 (Feb 5)	16 (Feb 5)
	τ_ρ	23 (Feb 12)	22 (Feb 11)
	τ_γ	28 (Feb 17)	28 (Feb 17)
	$\tau_{\delta\|1}$	57 (Mar 17)	57 (Mar 17)
	$\tau_{\delta\|2}$	72 (Apr 1)	82 (Apr 11)
	τ_ϕ	26 (Feb 15)	30 (Feb 19)

Note: $T = 99$, $n = 283$ excluding Wuhan; The dependent variable is the total newly confirmed cases for cities. We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.For change-point estimation, we treat the posterior modes of the change-points as their Bayesian point-estimate.

Using a spatial count panel model since total newly confirmed cases for all cities become non-negative. Throughout the paper we consider a linear SDPD model and treat the dependent variable vector of city level total newly confirmed cases, Y_t, to be continuous random variables. There are two reasons for this treatment: the total newly confirmed cases could be negative in some period due to systematic adjustments in reporting standards or correction for reporting errors, which are addressed by

the day fixed effect, and the coefficients of a linear model are easy to interpret. Nonetheless, to validate the linear model, we also run an estimation assuming a conditional Poisson distribution on y_{it}'s, using a sub-sample between April 8th to April 28th, the period during which the total newly confirmed cases for all cities become non-negative. Particularly, we specify a spatial count panel data model with a latent Gaussian spatial-temporal process. We assume that, conditional on some city latent variable ζ_{it}'s, the total newly confirmed cases y_{it}'s are mutually independent and following Poisson distributions across all cities and time, with non-negative mean $m_{it} = \exp(\zeta_{it})$. So the latent variable ζ_{it} can be viewed as the log of conditional mean of y_{it}. The conditional probability mass function (PMF) of y_{it} is,

$$f(y_{it}|\zeta_{it}) = \frac{\exp(y_{it}\zeta_{it} - \exp(\zeta_{it}))}{y_{it}!}, \ i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, T, \tag{17}$$

We further assume the $n \times 1$ vector of latent variable $\zeta_t = (\zeta_{1t}, \zeta_{2t}, \cdots, \zeta_{nt})'$ follows a linear spatial-temporal Gaussian process,

$$\zeta_t = \rho\zeta_{t-1} + \mu\bar{W}_{t-1}\zeta_{t-1} + \gamma M_{wh,t}y_{wh,t} + B_t\delta + \xi_t\phi + X_t\beta_1 + C\beta_2 + l_n\alpha_t + \epsilon_t, \tag{18}$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2 I_n)$. The coefficients ρ and μ would capture, respectively, the within-city transmission and the diffusion effects, at the log mean level of newly confirmed cases. More ideally, we may specify a full SDPD process on ζ_t and assume a Poisson distribution on y_{it}'s conditional on ζ_t. However, to the best of our knowledge, the Bayesian literature on the SDPD model with Poisson distribution is lacking. The inclusion of a contemporaneous spatial lag (contemporaneous cross-city transmission) would require a latent SDPD process on ζ_t's and call for a new and nontrivial MCMC sampler, whose algorithm stability and convergence need to be carefully examined. The successful implementation of such an algorithm would deserve another publication, which is beyond the scope of this paper. Therefore, we run a spatial model that excludes the contemporaneous cross-city transmission based on Poisson distribution, taking advantage of the finding that contemporaneous cross-city transmission is negligible during this period, and use the diffusion term of the latent variables $\mu\bar{W}_{t-1}\zeta_{t-1}$ to control for the possible cross-city transmission that still exist after early April.

For comparison, we also run the following linear spatial panel data model,

$$Y_t = \rho Y_{t-1} + \mu\bar{W}_{t-1}Y_{t-1} + \gamma M_{wh,t}y_{wh,t} + B_t\delta + \xi_t\phi + X_t\beta_1 + C\beta_2 + l_n\alpha_t + \epsilon_t, \tag{19}$$

where the within and cross city transmission effect on the city newly confirmed cases are captured, respectively, by ρ and μ. We want to see whether Bayesian estimates from the spatial count model in (17) and (18) would be qualitatively similar to the estimates from (19).

We further develop a computationally tractable MCMC algorithm to estimate the spatial count panel model in (17) and (18). Let $X_{\zeta|t} = (\zeta_{t-1}, \bar{W}_{t-1}\zeta_{t-1}, M_{wh,t}y_{wh,t}, B_t, \xi_t, X_t)$, $\beta_c = (\rho, \mu, \gamma, \delta, \phi, \beta_1')'$ and $\theta_\zeta = (\beta_c, \beta_2, \{\alpha_t\}, \sigma_\epsilon^2)$. Assume a normal prior for β_c, i.e., $\pi(\beta_c) \sim \mathcal{N}(\beta_{0|c}, B_c)$. The priors of β_2, α_t's and σ_ϵ^2 can be specified in the same way as (14). We follow the idea of data augmentation in (17) to treat the latent variable ζ_{it}'s as parameters to sample, in addition to θ_ζ. By Bayes' theorem,

$$P(\theta_\zeta, \{\zeta_t\}|\{Y_t\}) \propto \pi(\theta_\zeta) \times f(\{Y_t\}|\{\zeta_t\}) \times \prod_{t=1}^{T} f(\zeta_t|\zeta_{t-1}, \theta_\zeta),$$

where $f(\{Y_t\}|\{\zeta_t\}) = \prod_{i=1}^{n}\prod_{t=1}^{T} f(y_{it}|\zeta_{it})$ and $f(\zeta_t|\zeta_{t-1}, \theta_\zeta) = (2\pi\sigma_\epsilon^2)^{-\frac{n}{2}} \exp(-\frac{H_{\zeta|t}(\theta_\zeta)'H_{\zeta|t}(\theta_\zeta)}{2\sigma_\epsilon^2})$, with $H_{\zeta|t}(\theta_\zeta) = \zeta_t - \rho\zeta_{t-1} - \mu\bar{W}_{t-1}\zeta_{t-1} - \gamma M_{wh,t}y_{wh,t} - B_t\delta - \xi_t\phi - X_t\beta_1 - C\beta_2 - l_n\alpha_t$. We treat the initial latent variable ζ_{i0} as exogenously given and set $\zeta_{i0} = ln(y_{i0})$ for $y_{i0} > 0$, otherwise zero. The MCMC sampler for the spatial count panel model can be realized via the following two blocks:

- Step 1: Conditional on the latent variable ζ_t's, we obtain the posterior draws of β_c, β_2, σ_ζ^2 and α_t's;
  - Step 1.1: Sample β_c from $P(\beta_c|\{Y_t\}, \{\zeta_t\}, \beta_2, \sigma_\zeta^2, \{\alpha_t\})$ using a Gibbs step;
  - Step 1.2: Sample β_2 from $P(\beta_2|\{Y_t\}, \{\zeta_t\}, \beta_c, \sigma_\zeta^2, \{\alpha_t\})$ using a Gibbs step;
  - Step 1.3: Sample σ_ζ^2 from $P(\sigma_\zeta^2|\{Y_t\}, \{\zeta_t\}, \beta_c, \beta_2, \{\alpha_t\})$ using a Gibbs step;
  - Step 1.4: Sample α_t from $P(\alpha_t|\{Y_t\}, \{\zeta_t\}, \beta_c, \beta_2, \sigma_\zeta^2)$ using a Gibbs step for all t's.
- Step 2: Conditional on β_c, β_2, σ_ζ^2 and α_t's, we obtain the posterior draws of the latent variables ζ_t for all t's.
  - Step 2.1: Sample $\zeta_1 = (\zeta_{11}, \zeta_{21}, \cdots, \zeta_{n1})'$ from $P(\zeta_1|Y_1, \zeta_2, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_1)$ using a M-H step;
  - Step 2.2: Sample $\zeta_t = (\zeta_{1t}, \zeta_{2t}, \cdots, \zeta_{nt})'$ from $P(\zeta_t|Y_t, \zeta_{t-1}, \zeta_{t+1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_t)$ using a M-H step for $t = 2, \cdots, T-1$;
  - Step 2.3: Sample ζ_T from $\zeta_T = (\zeta_{1T}, \zeta_{2T}, \cdots, \zeta_{nT})'$ from $P(\zeta_T|Y_T, \zeta_{T-1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_t)$ using a M-H step.

Conditional on ζ_t's, the Gibbs sampling steps for β_c, β_2, σ_ζ^2, and α_t's are similar to the MCMC sampler of the linear SDPD model, thus omitted here. Below, we list the detailed MCMC sampling steps for latent variable ζ_t's. For each step, the full conditional is conditioned on the rest of parameters and latent variables with the most updated values at the current iteration.

Step 2.1 Sample $\zeta_1 = (\zeta_{11}, \zeta_{21}, \cdots, \zeta_{n1})'$ from $P(\zeta_1|Y_1, \zeta_2, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_1)$;
By Bayes'theorem,

$$P(\zeta_1|Y_1, \zeta_2, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_1) \propto \pi(\zeta_1|\zeta_0) \times \pi(\zeta_2|\zeta_1) \times \prod_{i=1}^{n} f(y_{i1}|\zeta_{i1}). \quad [20]$$

At the gth iteration, we apply a M-H step to sample ζ_1 . Propose $\tilde{\zeta}_1 \sim \mathcal{N}(\zeta_1^{(g-1)}, e_1 I_n)$, where e_1 is chosen by the user. Let $m_1 = X_{\zeta|1}\beta_c + C\beta_2 + l_n\alpha_1$ and

$$\tilde{m}_2 = \rho\tilde{\zeta}_1 + \mu\bar{W}_1\tilde{\zeta}_1 + \gamma M_{wh,2}y_{wh,2} + B_2\delta + \xi_2\phi + X_2\beta_1 + C\beta_2 + l_n\alpha_2,$$
$$m_2^{(g-1)} = \rho\zeta_1^{(g-1)} + \mu\bar{W}_1\zeta_1^{(g-1)} + \gamma M_{wh,2}y_{wh,2} + B_2\delta + \xi_2\phi + X_2\beta_1 + C\beta_2 + l_n\alpha_2.$$

With acceptance probability

$$Pr(\zeta_1^{(g-1)}, \tilde{\zeta}_1) = \min\{\prod_{i=1}^{n} \frac{f(y_{i1}|\tilde{\zeta}_{i1})}{f(y_{i1}|\zeta_{i1}^{(g-1)})} \times \frac{\mathcal{N}(\tilde{\zeta}_1; m_1, \sigma_\epsilon^2 I_n)}{\mathcal{N}(\zeta_1^{(g-1)}; m_1, \sigma_\epsilon^2 I_n)} \times \frac{\mathcal{N}(\zeta_2; \tilde{m}_2, \sigma_\epsilon^2 I_n)}{\mathcal{N}(\zeta_2; m_2^{(g-1)}, \sigma_\epsilon^2 I_n)}, 1\}$$

update $\zeta_1^{(g)} = \tilde{\zeta}_1$, otherwise set $\zeta_1^{(g)} = \zeta_1^{(g-1)}$.
Step 2.2 Sample $\zeta_t = (\zeta_{1t}, \zeta_{2t}, \cdots, \zeta_{nt})'$ from $P(\zeta_t|Y_t, \zeta_{t-1}, \zeta_{t+1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_t)$ for $t = 2, 3, \cdots, T-1$;
By Bayes'theorem,

$$P(\zeta_t|Y_t, \zeta_{t-1}, \zeta_{t+1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_t) \propto \pi(\zeta_t|\zeta_{t-1}) \times \pi(\zeta_{t+1}|\zeta_t) \times \prod_{i=1}^{n} f(y_{it}|\zeta_{it}). \quad [21]$$

At the gth iteration, we apply a M-H step to sample ζ_t . Propose $\tilde{\zeta}_t \sim \mathcal{N}(\zeta_t^{(g-1)}, e_t I_n)$, where e_t is chosen by the user. Let $m_t = X_{\zeta|t}\beta_c + C\beta_2 + l_n\alpha_t$ and

$$\tilde{m}_{t+1} = \rho\tilde{\zeta}_t + \mu\bar{W}_t\tilde{\zeta}_t + \gamma M_{wh,t+1}y_{wh,t+1} + B_{t+1}\delta + \xi_{t+1}\phi + X_{t+1}\beta_1 + C\beta_2 + l_n\alpha_{t+1},$$
$$m_{t+1}^{(g-1)} = \rho\zeta_t^{(g-1)} + \mu\bar{W}_t\zeta_t^{(g-1)} + \gamma M_{wh,t+1}y_{wh,t+1} + B_{t+1}\delta + \xi_{t+1}\phi + X_{t+1}\beta_1 + C\beta_2 + l_n\alpha_{t+1}.$$

With acceptance probability

$$Pr(\zeta_t^{(g-1)}, \tilde{\zeta}_t) = \min\{\prod_{i=1}^{n} \frac{f(y_{it}|\tilde{\zeta}_{it})}{f(y_{it}|\zeta_{it}^{(g-1)})} \times \frac{\mathcal{N}(\tilde{\zeta}_t; m_t, \sigma_\epsilon^2 I_n)}{\mathcal{N}(\zeta_t^{(g-1)}; m_t, \sigma_\epsilon^2 I_n)} \times \frac{\mathcal{N}(\zeta_{t+1}; \tilde{m}_{t+1}, \sigma_\epsilon^2 I_n)}{\mathcal{N}(\zeta_{t+1}; m_{t+1}^{(g-1)}, \sigma_\epsilon^2 I_n)}, 1\}$$

update $\zeta_t^{(g)} = \tilde{\zeta}_t$, otherwise set $\zeta_t^{(g)} = \zeta_t^{(g-1)}$.

Step 2.3 Sample $\zeta_T = (\zeta_{1T}, \zeta_{2T}, \cdots, \zeta_{nT})'$ from $P(\zeta_T|Y_T, \zeta_{T-1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_t)$;
By Bayes'theorem,

$$P(\zeta_T|Y_T, \zeta_{T-1}, \beta_c, \beta_2, \sigma_\zeta^2, \alpha_T) \propto \pi(\zeta_T|\zeta_{T-1}) \times \prod_{i=1}^{n} f(y_{iT}|\zeta_{iT}). \quad [22]$$

At the gth iteration, we apply a M-H step to sample ζ_T . Propose $\tilde{\zeta}_T \sim \mathcal{N}(\zeta_T^{(g-1)}, e_T I_n)$, where e_T is chosen by the user. Note that $m_T = X_{\zeta|T}\beta_c + C\beta_2 + l_n\alpha_T$. With acceptance probability

$$Pr(\zeta_T^{(g-1)}, \tilde{\zeta}_T) = \min\{\prod_{i=1}^{n} \frac{f(y_{iT}|\tilde{\zeta}_{iT})}{f(y_{iT}|\zeta_{iT}^{(g-1)})} \times \frac{\mathcal{N}(\tilde{\zeta}_T; m_T, \sigma_\epsilon^2 I_n)}{\mathcal{N}(\zeta_T^{(g-1)}; m_T, \sigma_\epsilon^2 I_n)}, 1\}$$

update $\zeta_T^{(g)} = \tilde{\zeta}_T$, otherwise set $\zeta_T^{(g)} = \zeta_T^{(g-1)}$.
The estimation results of the spatial count panel model are summarized in Table S5. Notably, the parameters are not directly comparable in magnitudes between linear and count model because the dependent variable is daily confirmed cases in the linear model and the log mean of that in the count model. Nevertheless, the coefficient on each variable can be directly compared within a model or column. We find qualitatively similar results in the linear and the count model; within city transmission is statistically significant and is larger than the abroad infection index coefficient and inflow index coefficient. In both the linear and count model, Wuhan influence is close to zero and lagged cross-city transmission (diffusion) is negative, suggesting cities are cautious about travel flows from other cities. These results show that the linear model and count model produce similar estimates for the period when the dependent variable is non-negative.

Table S5. Spatial count panel model for April

	Linear model		Count model	
	(I)	(II)	(III)	(IV)
Within city transmission	0.180	0.178	0.684	0.714
	(0.013)	(0.013)	(0.009)	(0.009)
Diffusion	N.A	-0.065	N.A	-0.135
		(0.025)		(0.019)
Wuhan Influence	N.A	0.000	N.A	-0.000
		(0.002)		(0.000)
Abroad Index (log)	0.034	0.035	0.004	0.003
	(0.013)	(0.014)	(0.001)	(0.001)
Inflow index	0.084	0.088	0.008	0.009
	(0.027)	(0.027)	(0.002)	(0.002)
Time effect	Yes	Yes	Yes	Yes
Weather control	Yes	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes	Yes

Note: The dependent variable is total newly confirmed cases. Sample period: April 8th to April 28th, with April 8th being the initial period; Columns (I) and (II): linear panel model without and with diffusion term; Columns (III) and (IV): spatial count panel model without and with diffusion term; We run, respectively, a Markov chain of 20000 and 35000 for the linear and count panel model, with 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

Heterogeneity of spread since importation. Since importation posts a threat to the infection control in China, some cities face extra pressure of transmission controls. One such example could be cities with international airports and therefore quarantine facilities starting from March 17, the first identified change-point after which importation takes effect. Another case is the province of Heilongjiang, where clusters of cases has been reported in April. To gain some insight on the spatial heterogeneity of infection, we allow cities of interest (i.e., international airport cities, or Heilongjiang cities) to have different transmission mechanisms using a sub-sample analysis for the period that importation index has an impact (i.e., March 17- April 1). For instance, we specify the following SDPD model to capture the spatial heterogeneity of cities in Heilongjiang,

$$\begin{aligned} y_{it} &= \lambda_h \sum_{j=1}^{n} \bar{w}_{ijt} \times \nu_j \times y_{jt} + \lambda_{nh} \sum_{j=1}^{n} \bar{w}_{ijt} \times (1-\nu_j) \times y_{jt} + \rho_h \times \nu_i y_{i,t-1} + \rho_{nh} \times (1-\nu_i) y_{i,t-1} \\ &+ \mu_h \sum_{j=1}^{n} \bar{w}_{ij,t-1} \times \nu_j \times y_{j,t-1} + \mu_{nh} \sum_{j=1}^{n} \bar{w}_{ij,t-1} \times (1-\nu_j) \times y_{j,t-1} + \gamma \bar{w}_{i,wht} y_{wht} \\ &+ b_{it}\delta + \xi_{it}\phi + x_{it}'\beta_1 + c_i'\beta_2 + \alpha_t + u_{it}, \end{aligned} \quad [23]$$

where ν_i is a 0-1 dummy variable indicating whether city i belongs to Heilongjiang or not, λ_h and λ_{nh} capture, respectively, the cross-city transmission effect originated from cities in Heilongjiang and cities that are not in Heilongjiang. Similarly, μ_h and μ_{nh} capture, respectively, the diffusion effect originated from cities in and out of Heilongjiang. ρ_h and ρ_{nh} represent the within city transmission effect for Heilongjiang and non-Heilongjiang cities. Let $D_h = diag(\nu_1, \nu_2, \cdots, \nu_n)$ be a diagonal selection matrix for Heilongjiang cities and $D_{nh} = I_n - D_h$. The matrix expression of (23) is

$$\begin{aligned} Y_t &= \lambda_h \bar{W}_t D_h Y_t + \lambda_{nh} \bar{W}_t D_{nh} Y_t + \rho_h D_h Y_{t-1} + \rho_{nh} D_{nh} Y_{t-1} + \mu_h \bar{W}_{t-1} D_h Y_{t-1} + \mu_{nh} \bar{W}_{t-1} D_{nh} Y_{t-1} \\ &+ \gamma M_{wh,t} y_{wh,t} + B_t\delta + \xi_t\phi + X_t\beta_1 + C\beta_2 + l_n\alpha_t + U_t. \end{aligned} \quad [24]$$

The SDPD model that captures the spatial heterogeneity of cities with quarantine facilities can be specified in a similar mannar,

$$\begin{aligned} Y_t &= \lambda_q \bar{W}_t D_q Y_t + \lambda_{nq} \bar{W}_t D_{nq} Y_t + \rho_q D_q Y_{t-1} + \rho_{nq} D_{nq} Y_{t-1} + \mu_q \bar{W}_{t-1} D_q Y_{t-1} + \mu_{nq} \bar{W}_{t-1} D_{nq} Y_{t-1} \\ &+ \gamma M_{wh,t} y_{wh,t} + B_t\delta + \xi_t\phi + X_t\beta_1 + C\beta_2 + l_n\alpha_t + U_t, \end{aligned} \quad [25]$$

where $D_q = diag(\nu_1, \nu_2, \cdots, \nu_n)$ be a diagonal selection matrix for cities with quarantine facilities, with ν_i being a 0-1 dummy variable for indicating whether i has quarantine facilities or not. Denote $D_{nq} = I_n - D_q$. λ_q and λ_{nq} capture, respectively, the cross-city transmission effect originated from cities with and without quarantine facilities. Similarly, μ_q and μ_{nq} capture, respectively, the diffusion effect originated from cities with and without quarantine facilities. ρ_q and ρ_{nq} represent, respectively, the within city transmission effect for cities with and without quarantine facilities.

Table S6 reports the estimation results for the spatial heterogeneity of cities in Heilongjiang. We find that the within city transmissions for cities in Heilongjiang are stronger than those for cities out of Heilongjiang, especially when we restrict the sample period to be April. This spatial heterogeneity confirms that cities in Heilongjiang do act as "superspreaders" and have a much larger transmission dynamics than cities out of Heilongjiang. On the other hand, we do not find significant differences in cross-city transmission between cities in and out of Heilongjiang, which is consistent with our main results that the domestic spread across cities is largely suppressed before the importation risk emerges.

Table S7 collects the estimation results for the spatial heterogeneity of cities with quarantine facilities since importation takes place in Mid March. We find that cities that hosted quarantine facilities do exhibit a larger within-city transmission than cities without quarantine facilities, suggesting that those cities with international airports do face a greater threat from importation risk than other cities. Notice that the sample period in Table S7 covers the period when importation took effect in our baseline results (March 17th-April 1st), and April 1st was the starting date of the centralized quarantine policy. Thus, the stronger within city transmission in the quarantine cities likely was driven by the transmission from imported cases before the centralized quarantine policy was implemented. Additionally, the diffusion effect (lagged cross-city transmission) originated from cities without quarantine facilities is significantly larger than that from cities hosting quarantine facilities.

Table S6. Spatial heterogeneity of cities in Heilongjiang

	Since April		Since importation	
	(I)	(II)	(I)	(II)
Cross city transmission originated from Heilongjiang cities	N.A	-0.067	N.A	-0.050
		(0.065)		(0.059)
Cross city transmission originated from other cities	N.A	-0.038	N.A	-0.017
		(0.067)		(0.013)
Within city transmission (Heilongjiang)	0.657	0.644	0.700	0.686
	(0.043)	(0.042)	(0.045)	(0.046)
Within city transmission (other provinces)	0.165	0.164	0.345	0.347
	(0.011)	(0.011)	(0.009)	(0.009)
Diffusion effect originated from Heilongjiang cities	N.A	-0.059	N.A	-0.058
		(0.069)		(0.066)
Diffusion effect originated from other cities	N.A	-0.020	N.A	0.004
		(0.018)		(0.015)
Wuhan influence	N.A	0.000	N.A	0.000
		(0.002)		(0.002)
Abroad Index (log)	0.028	0.029	0.014	0.013
	(0.011)	(0.010)	(0.009)	(0.009)
Inflow index	0.069	0.076	-0.003	-0.000
	(0.021)	(0.022)	(0.019)	(0.019)
Time effect	Yes	Yes	Yes	Yes
Weather control	Yes	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes	Yes

Note: $n = 283$ excluding Wuhan; The sample period since April and since importation are, respectively, from April 1st to April 28th and from March 17th to April 28th; The dependent variable is the total newly confirmed cases of each city. (I) is a linear panel model without cross-city transmission and diffusion effect and (II) is the SDPD model. We run a Markov chain of 20000 and 35000 for the linear and spatial panel model, with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

Table S7. Spatial heterogeneity of cities that host quarantine facilities

	Linear panel	SDPD
Cross city transmission from cities with quarantine facilities	N.A	-0.029
		(0.015)
Cross city transmission from cities without quarantine facilities	N.A	0.040
		(0.038)
Within city transmission (cities with quarantine facilities)	0.371	0.370
	(0.009)	(0.009)
Within city transmission (cities without quarantine facilities)	0.062	0.076
	(0.046)	(0.045)
Diffusion (cities with quarantine facilities)	N.A	0.002
		(0.015)
Diffusion (cities without quarantine facilities)	N.A	0.351
		(0.115)
Wuhan influence	N.A	0.000
		(0.002)
Abroad Index (log)	0.011	0.010
	(0.009)	(0.009)
Inflow index	-0.005	0.001
	(0.019)	(0.019)
Time effect	Yes	Yes
Weather control	Yes	Yes
Hospital control	Yes	Yes
Provincial effect	Yes	Yes

Note: $n = 283$ excluding Wuhan; The sample period since importation took effect is from March 17th to April 28th; The dependent variable is the total newly confirmed cases of each city. We run a Markov chain of 20000 and 35000 for the linear and spatial panel model, with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

Effect of abroad infection on domestic cases and imported cases. The main results in the paper use the total newly confirmed cases including both domestic and imported cases. The interpretation of the coefficient of the importation index is, given potential exposure to foreign virus infections, how many new cases will arise in China. To gain some idea of how much this increase stemmed from domestic cases versus imported cases, we first run a by-month SDPD model using domestic newly confirmed cases as the dependent variable. The interpretation of the importation risk index is, given potential exposure to foreign virus infections, how many new domestic cases will arise. We also regress province-level aggregated imported cases on importation index using the province by day panel over Feb.29-Apr.30, where Feb 29 is the date the first imported case is reported in our sample, to further support the point that the importation risk contributes to imported cases.

Figure S9 reports the time trend of total newly confirmed cases and domestic newly confirmed cases in China everyday since the imported cases had been reported starting on February 29, 2020. The time trends suggest that there are considerable amount of imported cases detected from mid March to April.

Table S8 reports the by-month SDPD results where the dependent variable is domestic newly confirmed cases. The estimates of abroad infection index is no longer statistically significant in each of the 4 months, including March where the importation risk supposes to take effect. This suggests the potential exposure to importation risk has little influence on domestic newly confirmed cases. The coefficients of the abroad infection index are either insignificant or much smaller compared to the by-month results where the dependent variable is the total confirmed cases (Table S2). Given our main finding has showed that the effect of importation risk on total newly confirmed case is significant between March 17th to April 1st, we may conclude that the importation risk mainly contributes to the imported cases instead of the domestic cases during this period. This implies China's containment efforts are fruitful in controlling the spread of imported infections to local residents. Meanwhile, as the influence of importation risk on domestic newly confirmed cases is negligible for all months, we do not further consider a SDPD model with change-points for the coefficient of the abroad index, when the dependent variables are restricted to be domestic cases.

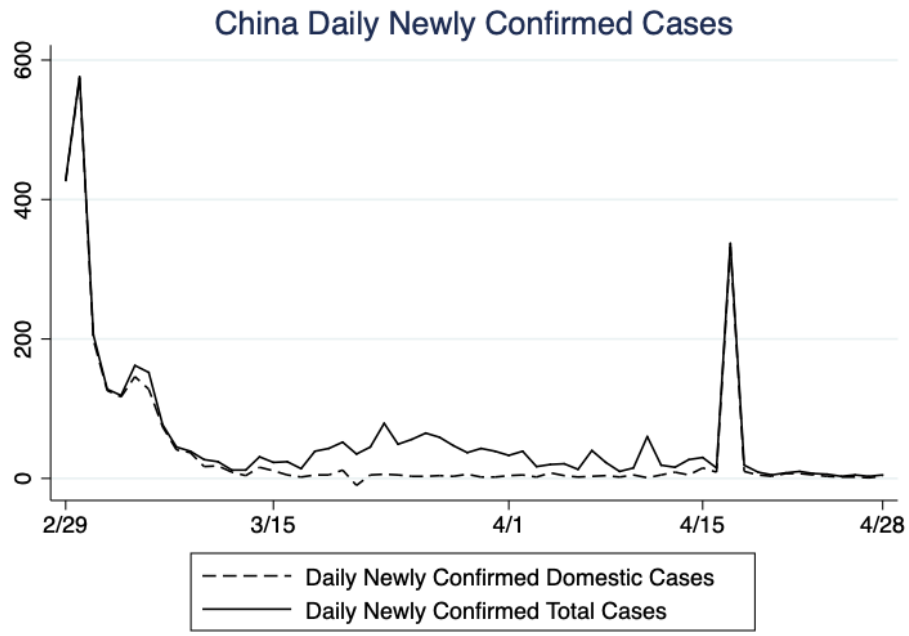


Fig. S9. The daily newly confirmed total cases include domestic cases and imported cases. Mudanjiang and Hulunbeier are excluded from the sample because they are land border cities with imported cases but without the abroad infection index due to lack of international flights. The number of total newly confirmed cases in Wuhan was revised up by 325 on April 16 by the Wuhan municipal headquarters for the novel coronavirus disease epidemic prevention and control, which led to the spikes on the day for both trends depicted.

Table S8. Domestic newly confirmed cases as the dependent variable

	Domestic newly confirmed cases			
	Jan	Feb	Mar	Apr
Cross-city transmission	0.143	0.225	0.041	-0.051
	(0.031)	(0.022)	(0.016)	(0.015)
With-in city transmission	0.584	0.486	0.064	0.524
	(0.015)	(0.009)	(0.011)	(0.010)
Duffision	0.058	0.026	-0.028	-0.065
	(0.037)	(0.025)	(0.022)	(0.015)
Abroad Index (log)	0.827	-0.111	-0.001	0.011
	(0.513)	(0.244)	(0.004)	(0.003)
Wuhan Influence	0.103	0.023	0.002	0.000
	(0.006)	(0.001)	(0.000)	(0.001)
Inflow Index	-0.042	0.297	0.007	0.013
	(0.213)	(0.325)	(0.008)	(0.006)
Time effect	Yes	Yes	Yes	Yes
Weather control	Yes	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes	Yes

Note: $n = 283$ excluding Wuhan; Jan: Jan 21-Feb 3; Feb: Feb 4-Feb 29; Mar: Mar 1-Mar 31, Apr: Apr 1-Apr 28; We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

To further support the point that the importation risk contributes mainly to imported cases, we regress province-level aggregated imported cases on the abroad infection index. The sample includes 27 provinces and four centrally-administered municipalities (CAMs). For provinces with more than one international airports, we add up the importation indices (abroad infection indices) across all airport. We report the estimation results in the following Table S9 using the 5-day smoothed

abroad infection index lagged over the past 5, 7, and 14 days respectively. The results in Columns 1-3 are based on the sample period over Feb.29-Apr.28 when imported case were reported. We find positive correlation between imported cases and abroad infection index at the 5% level and the 10% level in the 5 and 7 lag cases respectively. The correlation turns significantly negative in the 14 lag case. The results in Columns 4-6 are based on the sample period over Mar.17-Apr.1 when the importation risk takes effect. We find larger correlations in both the 5 and 7 lag cases whereas the correlation turns insignificant in the 14 lag case.

Table S9. The effect of importation risk on imported cases

	When imported cases were reported			When importation risk took effect		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	5-day lag	7-day lag	14-day lag	5-day lag	7-day lag	14-day lag
Abroad infection index	0.114**	0.091*	-0.192***	0.335***	0.362***	-0.186
	(0.053)	(0.054)	(0.055)	(0.113)	(0.112)	(0.135)
Observations	1,860	1,860	1,860	496	496	496
R-squared	0.203	0.202	0.206	0.639	0.640	0.633

Note: The estimation results are based on the province by day panel over Feb.29-Apr.28 in Columns 1-3 and over Mar.17-Apr.1 in Columns 4-6. The sample includes 27 provinces and four CAMs. The dependent variable is the number of imported cases. The independent variable is the 5-day smoothed abroad infection index lagged by 5 days, 7 days, and 14 days respectively as shown in the column titles. We control for the province fixed effects and date fixed effects.

Table S10. The effect of imported cases on domestic cases

	Domestic newly confirmed cases		
	(I)	(II)	(III)
Cross-city transmission	-0.000	0.019	-0.047
	(0.009)	(0.021)	(0.016)
With-in city transmission	0.283	0.084	0.518
	(0.006)	(0.013)	(0.010)
Diffusion	-0.058	-0.098	-0.065
	(0.011)	(0.027)	(0.014)
Imported case	0.009	0.014	-0.002
	(0.003)	(0.005)	(0.003)
Wuhan Influence	0.002	0.031	0.000
	(0.000)	(0.215)	(0.001)
Inflow Index	0.017	0.016	0.009
	(0.005)	(0.010)	(0.006)
Time effect	Yes	Yes	Yes
Weather control	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes

Note: $n = 283$ excluding Wuhan; We run a Markov chain of 20000 with a 20% burn-in ratio. The sample periods for (I), (II), (III) are, respectively, when imported cases were reported (Feb.29-Apr.28), and when importation risk took effect (Mar.17-Apr.1) and died out (Apr.2-Apr.28). We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

We also directly test whether there is a significant correlation between imported cases and domestic cases (Table S10). Specifically, we use domestic newly confirmed cases as the dependent variable and replace the abroad infection index by the number of imported cases. We investigate the following three sub-sample periods: the period when imported cases were reported (Feb. 29-Apr.28), the period when importation risk took effect (Mar.17- Apr.1) and the period when importation risk died out (Apr.2-Apr.28). We find a weak but statistically significant correlation between imported cases and domestic cases when the sample period contains the period that importation took effect (Mar.17-Apr.1, Column II). But after the centralized quarantine policy was implemented on Apr.1 (Column III), the correlation between the two became insignificant. These findings imply that the the effect of imported cases in seeding new domestic cases is somehow limited and brief. The results further justify that the centralized quarantine measures (effective starting April 1) have been effective to deter the importation risk.

Incorporating city inflow intensity indices in the travel flow network. When we performed our original analyses, as an alternative to our baseline model, we did consider a spatial weighting matrix that incorporates city inflow intensity indices in the travel flow network. In such a model specification, we no longer control inflow intensity as a regressor. Instead, we multiply the flow percentage by the daily inflow intensity to obtain "flow rates" between cities, and use the flow rates to construct an intensity adjusted travel flow network through the following steps:

Step 1: For each date t and city i, we construct the flow rate from city j to city i as follows:

$$w_{ijt} = p_{ijt} * MI_{it},$$

where p_{ijt} is the daily percentage of the inflow population in city i from city j on date t; MI_{it} is the inflow intensity index in city i on date t.

Step 2: For each date t and city i, we construct the lagged 5-day moving average as follows:

$$\tilde{w}_{ijt} = \frac{\sum_{l=3}^{7} w_{ij,t-l}}{5} = \frac{\sum_{l=3}^{7} p_{ij,t-l} * MI_{i,t-l}}{5}, \quad \text{Lagged 5-day average flow rate}$$

Step 3: we row normalise $\tilde{w}_{ijt}$'s to derive the intensity adjusted travel flow network among 283 cities (excluding Wuhan):

$$\bar{w}_{ijt} = \frac{\tilde{w}_{ijt}}{\sum_j \tilde{w}_{ijt}}, \quad \text{Normalized 5-day average intensity adjusted travel flow network}$$

Note that the daily inflow intensity MI_{it} would not be cancelled out by the row-normalization, due to the 5 day average smoothing in Step 2. Hence, it still presents in the spatial weights of the SDPD model and is able to capture the influence of move-in index on the cross-city transmission that may vary by origin cities.

Tables S11 compares the estimation results of the SDPD model using the average flow rate spatial weights matrix, with our baseline model using average flow percentage spatial weights matrix. We find that under the average flow rate network, the estimated coefficients for the cross and within-city transmission, the influence of the abroad index, as well as the identified change-points are very similar to their corresponding counterparts under the average flow percentage network, which suggests that the incorporation of inflow indices in the spatial weights does not affect the main result.

Relaxing row normalization of travel flow network

The row-normalized spatial weights matrices in our model enable us to identify the average cross city transmission effect that one city faces, originated from its "neighbors" through the travel flow network. Additionally, the city connection (network edge) in our network is travel flow, which is continuous rather than binary (0-1). Hence, the row-normalized travel flow network put more weights on stronger connections, which complements some recent literature (18), that studies social interaction through binary networks and ignores the heterogeneity in connection intensities.

We also follow the spirit of (18) to run our baseline model with non-row normalized (i.e., aggregate) flow percentage weights and flow rate weights, to see how row standardization affect our parameter estimates. Let W_t be the non row-normalized travel flow network and $\Delta_w = \max_{t=0,1,\cdots,T} ||W_t||_\infty$. To estimate the SDPD model with aggregate spatial weights matrices, we need to modify the stability conditions on λ, ρ and μ as $\max\{|\lambda_1|, |\lambda_2|\}\Delta_w < 1$ and $\max\{|\lambda_1|, |\lambda_2|\}\Delta_w + \max\{|\rho_1|, |\rho_2|\} + |\mu|\Delta_w < 1$. For the case of flow percentage weights, Δ_w is close to 1 because it is the sum of percentage connections with the top 100 origins. Thus, the restrictions on λ do not change much compared with the row-normalized case. Then the identification of change-point τ_λ for cross-city transmission effect λ would not be affected when implementing the MCMC estimation procedure. But for flow rate weights where the inflow intensity is incorporated, Δ_w could be larger than 8, which would shrink the magnitude of λ dramatically from the stability condition and make τ_λ much harder to identify than the row-normalized case. To tackle this issue, we follow (19) to consider a scalar normalization on W_t, i.e., set $\mathbf{W}_t = W_t/\Delta_w$ in the model, and have the cross-city transmission scaled up as $\lambda\Delta_w$. We also impose a more informative prior on τ_λ to improve identification. We keep its lower bound at Feb 3, when all provinces had already launched the highest level response to COVID-19 and some cities had adopted partial shutdown strategy. But we further set its upper bound to be the date of February 26, when 13 provinces downgraded their response level and 26 out of 31 provinces had resumed shuttle services across provinces. The estimation results are summarized in Table S12. We indeed find that the parameter and change-point estimates produced by the model with non row-normalized percentage weights (aggregate percentage share, Table S12) are very similar to the baseline result under row-standardization. For the case of non row-normalized flow rate weights (aggregate flow rate, Table S12), the result shows a delayed change point for cross-city transmission λ – February 21 vs. February 5 in the baseline case, but magnitudes and change-points of other transmission mechanisms remain similar to the baseline row-normalized case. This result implies that cross-city transmission lasted longer if we impose a different city network interpretation. Meanwhile, other transmission mechanisms remain robust to this alternative network interpretation.

Table S11. SDPD model with alternative spatial weights

		Average flow Percentage	Average flow rate
Cross-city transmission	λ_1	0.279	0.261
	S.D	(0.009)	(0.009)
	95% CI	$[0.259,\ 0.295]$	$[0.243,\ 0.278]$
	λ_2	0.034	0.022
	S.D	(0.011)	(0.011)
	95% CI	$[0.012,\ 0.056]$	$[0.001,\ 0.041]$
Wuhan influence	γ_1	0.029	0.024
	S.D	(0.003)	(0.000)
	95% CI	$[0.028,\ 0.030]$	$[0.023,\ 0.024]$
	γ_2	-0.009	-0.011
	S.D	(0.002)	(0.002)
	95% CI	$[-0.014,\ -0.005]$	$[-0.016,\ -0.006]$
Within-city transmission	ρ_1	0.708	0.723
	S.D	(0.005)	(0.005)
	95% CI	$[0.698,\ 0.717]$	$[0.713,\ 0.734]$
	ρ_2	0.163	0.181
	S.D	(0.007)	(0.006)
	95% CI	$[0.149,\ 0.176]$	$[0.168,\ 0.195]$
Abroad index (log)	δ_1	-0.087	-0.096
	S.D	(0.063)	(0.063)
	95% CI	$[-0.211,\ 0.037]$	$[-0.220,\ 0.028]$
	δ_2	0.298	0.245
	S.D	(0.087)	(0.088)
	95% CI	$[0.120,\ 0.459]$	$[0.073,\ 0.413]$
	δ_3	-0.058	-0.086
	S.D	(0.052)	(0.053)
	95% CI	$[-0.167,\ 0.039]$	$[-0.201,\ 0.009]$
Inflow index	ϕ_1	0.228	N.A
	S.D	(0.080)	
	95% CI	$[0.066,\ 0.382]$	N.A
	ϕ_2	-0.003	N.A
	S.D	(0.077)	
	95% CI	$[-0.156,\ 0.144]$	N.A
Change-points	τ_λ	16 (Feb 5)	16 (Feb 5)
	τ_ρ	22 (Feb 11)	23 (Feb 12)
	τ_γ	28 (Feb 17)	28 (Feb 17)
	$\tau_{\delta\|1}$	57 (Mar 17)	57 (Mar 17)
	$\tau_{\delta\|2}$	72 (Apr 1)	72 (Apr 1)
	τ_ϕ	30 (Feb 19)	N.A

Note: $T = 99$, $n = 283$ excluding Wuhan; The dependent variable is the total newly confirmed cases for cities. We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates. For change-point estimation, we treat the posterior modes of change-points as their Bayesian point estimates.

Table S12. SDPD model without row-normalization

		Aggregate Percentage Share	Aggregate Inflow Intensity
Cross-city transmission	λ_1	0.244	0.150
		(0.011)	(0.033)
	95% CI	[0.220, 0.264]	[0.067, 0.194]
	λ_2	-0.006	0.036
		(0.014)	(0.077)
	95% CI	[−0.037, 0.019]	[−0.137, 0.167]
Wuhan influence	γ_1	0.025	0.098
		(0.000)	(0.002)
	95% CI	[0.024, 0.026]	[0.094, 0.102]
	γ_2	-0.013	-0.116
		(0.003)	(0.030)
	95% CI	[−0.018, −0.007]	[−0.171, −0.054]
Within-city transmission	ρ_1	0.729	0.800
		(0.005)	(0.005)
	95% CI	[0.720, 0.738]	[0.791, 0.809]
	ρ_2	0.184	0.238
		(0.007)	(0.007)
	95% CI	[0.171, 0.197]	[0.224, 0.252]
Abroad index (log)	δ_1	-0.086	-0.080
		(0.064)	(0.067)
	95% CI	[−0.213, 0.038]	[−0.213, 0.052]
	δ_2	0.267	0.254
		(0.090)	(0.089)
	95% CI	[0.010, 0.441]	[0.094, 0.431]
	δ_3	-0.063	-0.041
		(0.055)	(0.059)
	95% CI	[−0.180, 0.038]	[−0.166, 0.066]
Inflow index	ϕ_1	0.188	N.A
		(0.081)	
	95% CI	[0.029, 0.346]	N.A
	ϕ_2	0.018	N.A
		(0.099)	
	95% CI	[−0.189, 0.185]	N.A
Change-points	τ_λ	16 (Feb 5)	32 (Feb 21)
	τ_ρ	23 (Feb 12)	23 (Feb 12)
	τ_γ	28 (Feb 17)	29 (Feb 18)
	$\tau_{\delta\|1}$	57 (Mar 17)	55 (Mar 15)
	$\tau_{\delta\|2}$	72 (Arp 1)	72 (Arp 1)
	τ_ϕ	26 (Feb 15)	N.A

Note: $T = 99$, $n = 283$ excluding Wuhan; The dependent variable is the total newly confirmed cases for cities. We run a Markov chain of 30000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates. For change-point estimation, we treat the posterior modes of change-points as their Bayesian point estimates.

Relaxing the stability condition on the within-city transmission coefficient

In our baseline model, we impose the stability condition that requires the with-city transmission (AR coefficients) ρ to be less than 1. One may worry that these stability conditions would rule out rapidly growing infection cases in cities at the early phase of the epidemic in China. To alleviate this concern, we re-estimate our baseline model without the stability condition for ρ. Specifically, we drop the condition that $\max\{|\lambda_1|, |\lambda_2|\} + \max\{|\rho_1|, |\rho_2|\} + |\mu| < 1$, and still maintain the stability condition on λ, i.e., $\max\{|\lambda_1|, |\lambda_2|\} < 1$. We further modify the upper and lower bounds of the uniform priors on ρ_1 and ρ_2 to be respectively, 5 and -5, to accommodate non-stationary exponential growth of infection cases. Thus, the stability conditions left in the M-H step for λ, ρ and μ are modified as $\max\{|\lambda_1|, |\lambda_2|\} < 1$, $\max\{|\rho_1|, |\rho_2|\} < 5$ (from priors), and $|\mu| < 1$ (from priors). The estimation results for the baseline model without stability conditions on ρ are provided in the following Table S13. We find that the results are similar to baseline results– the within-city transmission coefficient is 0.709 [95% CI: 0.699,0.718]. We also ran a constant coefficient SDPD model without the stability condition for two short panels (first 14 days and 22 days, just before within city transmission decreased), to examine the earlier spread in Table S14, yet the within-city transmission parameter is still less than 1 and none of the corresponding Bayesian 95% CIs exceed 1. This is evident that the exponential growth of confirmed cases at the outbreak in China was driven by cross-city transmission, as suggested by the marginal effect derivation. Our work highlights the interconnection of regions and the spillover effect.

Table S13. SDPD model without stability condition on ρ

		Average Percentage Share
Cross-city transmission	λ_1	0.293
	S.D	(0.013)
	95% CI	[0.270, 0.319]
	λ_2	0.038
	S.D	(0.012)
	95% CI	[0.012, 0.056]
Wuhan influence	γ_1	0.029
	S.D	(0.003)
	95% CI	[0.028, 0.030]
	γ_2	-0.009
	S.D	(0.002)
	95% CI	[−0.014, −0.005]
Within-city transmission	ρ_1	0.709
	S.D	(0.005)
	95% CI	[0.699, 0.718]
	ρ_2	0.165
	S.D	(0.008)
	95% CI	[0.149, 0.181]
Abroad index (log)	δ_1	-0.091
	S.D	(0.063)
	95% CI	[−0.213, 0.033]
	δ_2	0.278
	S.D	(0.091)
	95% CI	[0.103, 0.448]
	δ_3	-0.065
	S.D	(0.054)
	95% CI	[−0.179, 0.034]
Inflow index	ϕ_1	0.219
	S.D	(0.078)
	95% CI	[0.062, 0.371]
	ϕ_2	-0.011
	S.D	(0.077)
	95% CI	[−0.162, 0.132]
Change-points	τ_λ	16 (Feb 5)
	τ_ρ	22 (Feb 11)
	τ_γ	28 (Feb 17)
	$\tau_{\delta\|1}$	57 (Mar 17)
	$\tau_{\delta\|2}$	72 (Apr 1)
	τ_ϕ	26 (Feb 15)

Note: $T = 99$, $n = 283$ excluding Wuhan; The dependent variable is the total newly confirmed cases for cities. We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mean of parameters as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates. For the change-point estimation, we treat the posterior modes of change-points as their Bayesian point estimates.

Table S14. Sub-sample analysis for SDPD model without stability condition on ρ

	Total newly confirmed cases			
	Linear panel		SDPD	
	First 14 days	First 22 days	First 14 days	First 22 days
Cross-city transmission	N.A	N.A	0.147	0.183
			(0.029)	(0.024)
With-in city transmission	0.599	0.646	0.584	0.637
	(0.015)	(0.010)	(0.014)	(0.011)
Diffusion	N.A	N.A	0.056	-0.097
			(0.035)	(0.026)
Abroad index (log)	0.819	0.093	0.824	0.108
	(0.507)	(0.378)	(0.511)	(0.372)
Wuhan Influence	0.109	0.045	0.103	0.045
	(0.006)	(0.003)	(0.006)	(0.003)
Inflow index	-0.091	0.010	-0.032	-0.016
	(0.213)	(0.175)	(0.218)	(0.174)
Time effect	Yes	Yes	Yes	Yes
Weather control	Yes	Yes	Yes	Yes
Hospital control	Yes	Yes	Yes	Yes
Provincial effect	Yes	Yes	Yes	Yes

Note: The dependent variable is the total newly confirmed cases for cities. The sample period is from Jan 21 to Feb 3 (First 14 days), or from Jan 21 to Feb 11 (First 22 days). We run a Markov chain of 20000 with a 20% burn-in ratio. We treat the posterior mode of change-points as their Bayesian point estimates. We also report the standard deviation of the posterior samples of parameters in the parentheses. We rely on the Bayesian 95% credible interval (CI) to judge the significance of parameter estimates.

References

1. SA Lauer, et al., The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals Intern. Medicine* **172**, 577–582 (2020).
2. Y Qiu, X Chen, W Shi, Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China. *J. Popul. Econ.* (2020).
3. H Gibbs, et al., Changing travel patterns in China during the early stages of the COVID-19 pandemic. *Nat. Commun.* **11**, 5012 (2020).
4. C Cheng, et al., The coupled impact of emergency responses and population flows on the covid-19 pandemic in china. *GeoHealth*, e2020GH000332 (2020).
5. S Lai, et al., Assessing spread risk of wuhan novel coronavirus within and beyond china, january-april 2020: a travel network-based modelling study. *medRxiv* (2020).
6. European Centre for Disease Prevention and Control, Download historical data (to 14 december 2020) on the daily number of new reported covid-19 cases and deaths worldwide (2020) https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide.
7. LF Lee, J Yu, Some recent developments in spatial panel data models. *Reg. Sci. Urban Econ.* **40**, 255–271 (2010).
8. K Li, Spatial panel data models with structural change. *Available at MPRA Pap. No. 85388* (2018).
9. R Horn, C Johnson, *Matrix Analysis*. (Cambridge University Press), (1985).
10. J LeSage, K Pace, *Introduction to Spatial Econometrics*. (Chapman and Hall/CRC Press), (2009).
11. X Han, CS Hsieh, LF Lee, Estimation and Model selection of Higher-order Spatial Autoregressive Model: An Efficient Bayesian Approach. *Reg. Sci. Urban Econ.* **63**, 97–120 (2017).
12. H Haario, E Saksman, J Tamminen, An adaptive metropolis algorithm. *Bernoulli* **7**, 223–242 (2001).
13. G Roberts, J Rosenthal, Examples of adaptive MCMC. *J. Comput. Graph. Stat.* **18**, 349–367 (2009).
14. A Gelman, G Roberts, W Gilks, Efficient metropolis jumping rules. *Bayesian Stat.* **5**, 599–607 (1996).
15. Y Mundlak, On the pooling of time series and cross section data. *Econometrica* **46**, 69–85 (1978).
16. H Fang, L Wang, Y Yang, Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China. *J. Public Econ.* **191**, 104272 (2020).
17. M Tanner, W Wong, The Calculation of Posterior Distributions by Data Augmentation. *J. Am. Stat. Assoc.* **82**, 528–540 (1987).
18. X Liu, E Patacchini, Y Zenou, Endogenous peer effects: Local aggregate or local average? *J. Econ. Behav. Organ.* **103**, 39–59 (2014).
19. H Kelejian, I Prucha, Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *J. Econom.* **157**, 53–67 (2010).