Chen, Xi; Qiu, Yun; Shi, Wei; Yu, Pei

Working Paper

Key Links in Network Interactions: Assessing Route-specific Travel Restrictions in China during the Covid-19 Pandemic

GLO Discussion Paper, No. 1030

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Chen, Xi; Qiu, Yun; Shi, Wei; Yu, Pei (2022) : Key Links in Network Interactions: Assessing Route-specific Travel Restrictions in China during the Covid-19 Pandemic, GLO Discussion Paper, No. 1030, Global Labor Organization (GLO), Essen

This Version is available at:
https://hdl.handle.net/10419/249203

Key Links in Network Interactions: Assessing Route-specific Travel Restrictions in China during the Covid-19 Pandemic*

Xi Chen† Yun Qiu‡ Wei Shi§ Pei Yu¶

January 22, 2022

Abstract

We consider a model of network interactions where the outcome of a unit depends on the outcomes of the connected units. We determine the key network link, i.e., the network link whose removal results in the largest reduction in the aggregate outcomes, and provide a measure that quantifies the contribution of a network link to the aggregate outcomes, which complements the intercentrality measure of the key network node proposed by Ballester, Calvó-Armengol, and Zenou (2006). We provide an example examining the spread of Covid-19 in China. Travel restrictions were imposed to limit the spread of infectious diseases. As uniform restrictions can be inefficient and incur unnecessarily high costs, we examine the design of restrictions that target specific travel routes. Our approach may be generalized to multiple countries to guide policies during epidemics ranging from *ex ante* route-specific travel restrictions to *ex post* health measures based on travel histories, and from the initial travel restrictions to the phased reopening.

JEL Classification: C21; I18; D85; H75

Keywords: network interactions; key network links; Covid-19; transmission

*We thank Yiwei Chen, Yuming Fu, Wen-Tai Hsu, Yi Huang, Hanbat Jeong, Bruce Weinberg, Maisy Wong, Junfu Zhang, seminar participants at Jinan University, Singapore Management University, and anonymous reviewers for comments. Chen thanks the following funding sources: US PEPPER Center Scholar Award (P30AG021342) and NIA grant (K01AG053408). Shi thanks the Ministry of Education of China Humanities and Social Science General Program (Grant No.18YJC790138) and the National Natural Science Foundation of China (Grant No.71803062) for financial support. Qiu and Shi acknowledge the support from the 111 Project of China (Grant No.B18026). The GLO Discussion Paper Series serves as a preprint server to deposit latest research for feedback. The views expressed herein and any remaining errors are the authors' and do not represent any official agency. None of the authors have potential conflicts of interests that could bias this work.

†Department of Health Policy and Management, Yale School of Public Health; Department of Economics, Yale University; Global Labor Organization (GLO); Email: xi.chen@yale.edu

‡Institute for Economic and Social Research, Jinan University; Email: yunqiu_yq@jnu.edu.cn

§Institute for Economic and Social Research, Jinan University; Email: wshi16@jnu.edu.cn (*corresponding author*)

¶Department of Economics, Rice University; Email: py14@rice.edu

1 Introduction

The spread of communicable diseases, especially those are transmittable via airborne droplets, depends crucially on the extent of interactions between infectious and susceptible people. Population flows have therefore been shown to strongly predict the spread of Covid-19 (e.g., Fang *et al.*, 2020, Jia *et al.*, 2020, Qiu *et al.*, 2020, Wu *et al.*, 2020) as well as other infectious diseases (e.g., Brockmann and Helbing, 2013) across space. To slow the transmission of Covid-19, many public health measures have been adopted across the world, ranging from mild measures (e.g. social distancing, quarantine and isolation, travel restrictions, testing and contact tracing) to stringent measures (e.g. city lockdown, shelter-in-place). While many of these public health measures are effective in suppressing the spread of Covid-19 (e.g., Tian *et al.*, 2020), they can also bring significant social and economic costs and disruptions (Dai *et al.*, 2021, Duan *et al.*, 2021).

In this paper, we explore the feasibility of imposing travel restrictions on specific origin and destination pairs and examine the optimal designs of such policies. In addition to being less restrictive and therefore more cost-effective than lockdowns of entire cities, route-specific travel restrictions can still be implemented even when complete lockdowns are not (e.g., when the unit under consideration is a major metropolitan area or an entire country). Even in scenarios when route-specific travel restrictions are not possible, our identified targeted travel restrictions, once integrated with advanced mobile technology and specific public health measures[1], can be used to improve risk management for people with certain travel histories. For instance, information on travel histories has been linked to centralized, real-time health insurance databases and electronic health records to allow healthcare facilities to identify high-risk patients for targeted screening, timely quarantine, and aggressive contact tracing (Emanuel *et al.*, 2020, Wang *et al.*, 2020). These data can also guide border checks and surveillance (Whitelaw *et al.*, 2020).

To characterize which route-specific travel restrictions are more effective, we consider a standard model of network interactions. Various studies have highlighted that features of networks can be used to enhance the effectiveness of policy interventions that intend to influence agents' behavior (e.g., Deng and Sun, 2017, Lee *et al.*, 2020). Ballester *et al.* (2006) characterize the key node in a network whose removal has the largest impact on the aggregate outcome. We determine the key link between two nodes, whose removal results in the largest reduction in the aggregate outcome, and show that this can be characterized as a product of the centrality of the origin, the centrality of the destination, and the link intensity. Therefore, an optimal travel restriction policy should take into account the travel intensity of a route, and also the risk of the origin and the risk that the destination can bring to other cities. Our results are applicable in network interaction models where agents' outcomes or actions depend on those of the peers (e.g., Lee, 2007, Patacchini *et al.*, 2017), and a planner can alter the strength of network links. Other potential applications include determining key collaboration relationships in the network of firms (Hsieh *et al.*, 2020),

[1]With the advent of mobile payment applications, social media, security camera footage, facial recognition, and global positioning system (GPS) in vehicles to collect real-time data, travel histories of individuals have become increasingly easily accessible in multiple countries and regions.

strengthening ties between entrepreneurs to improve an entrepreneurial ecosystem (Mellon *et al.*, 2016), and strengthening connections in networks to support interventions on suicide and alcohol use disorders (Philip *et al.*, 2016).

Our paper is not the first to examine key links in network interactions. Our theoretical analysis accommodates features such as directed networks, ex-ante heterogeneity between units and contextual effects, which are motivated by our empirical analysis on key population flow routes in the spread of Covid-19 in China. Building on the key players analysis of Ballester *et al.* (2006), Ballester *et al.* (2010) provide theoretical results on the contributions of specific network links to the aggregate activities. Units are assumed ex-ante homogeneous in their analysis. Sun *et al.* (2021) is a recent paper that comprehensively studies the design of interventions targeting characteristics of network nodes and intensities of network links[2]. Based on a network model of product adoptions, Meng *et al.* (2022) examine key nodes and links. We complement Sun *et al.* (2021) and Meng *et al.* (2022) by considering directed networks and contextual effects.

Applying the theoretical model to the data on the spread of Covid-19 in China between January and February, 2020, we first show that intercity population flows intensify spatial virus spreading. Based on the estimated parameters, we identify the routes in the population flow networks most influential on the total number of Covid-19 cases in China and thus the top candidates where additional public health measures prove necessary. The top routes consist of those closely connected to areas with severe infections and those whose destinations are cities with large population outflows. Accounting for such spillovers saliently alters our perceptions of travel restriction policies. These results can be generalized to guide our responses to other communicable diseases with human-to-human transmission.

We add to the literature on the spatial spread of diseases (Brockmann and Helbing, 2013) by considering the design of policies that affect the rate of spillovers. Our paper also contributes to the growing literature on the optimal designs of various aspects of the public health measures in response to Covid-19. A number of studies embed a Susceptible-Infectious-Recovered (SIR) model (Kermack and McKendrick, 1927) in an optimal control problem, such as lockdowns of certain sections of the population (Acemoglu *et al.*, 2020, Alfaro *et al.*, 2020, Alvarez *et al.*, 2020), testing and quarantine (Berger *et al.*, 2020), along with coordinated and uncoordinated shelter-in-place orders (Holtz *et al.*, 2020), initiatives that help students cope with distance learning (Clark *et al.*, 2021), etc. Fajgelbaum *et al.* (2020) examine optimal restrictions on directional commuting flows, integrating a spatial epidemiology model with a quantitative model of commuting, production, and equilibrium across locations. Their focus is on commuting flows within a metropolitan area, where population flows in the form of commuting affect disease spread and production. We identify the key parameters using a causal inference model, in comparison to the literature, which is dominated by epidemiology models. Our model is more concise and applicable on a larger geographic scale, such as the spread of diseases across many cities or countries.

The paper is organized as follows. Section 2 describes a network interaction model and provides

[2]We thank a referee for bringing our attention to this paper.

measures that identify the key network links. In Section 3, we illustrate the use of the model in determining key population flow routes using the Covid-19 data in China. Section 4 presents conclusions. The appendix contains the proofs of the theorems and supplementary results.

2 The key link in network interactions

2.1 Model setup

There are $\mathcal{N} = \{1, \cdots, n\}$ units interacting through a network described by an $n \times n$ matrix $W = (w_{ij})$ with $w_{ij} \geq 0$ and $w_{ii} = 0$. Let y_i denote the outcome of unit i. Two examples of the model are given in Examples 1 and 2. We assume the following model of network interactions,

$$y_i = \left(\sum_{j=1}^{n} w_{ij} y_j \right) \lambda + x_i' \beta + u_i, \tag{1}$$

where the outcome of unit i is affected by the outcomes of other units with pairwise weights w_{ij} and the scalar parameter λ measuring the intensity of interactions. x_i is a vector of control variables with corresponding coefficients β, which can include time lagged values of y_i. u_i is the error term. Denote $\eta_i = x_i'\beta + u_i$ and $\eta = \begin{pmatrix} \eta_1 & \cdots & \eta_n \end{pmatrix}'$ with $'$ denoting transpose.

Example 1. *In our empirical illustration in Section 3, y_i denotes the number of Covid-19 cases in city i in logarithms. w_{ij}, $i \neq j$ is a measure of the intensity of population flows from city j to i. This specification can be viewed as a static version of the spatial autoregressive model of mobility and disease spread in Brinkman and Mangum (2021). The use of a static model allows us to focus on the analysis of key network links while abstracting away the complicated feedback effects in dynamic models.*

Example 2. *Eq.(1) can be rationalized as best responses in a game where individual utilities depend on linear and quadratic terms of the own and others' actions (e.g., Ballester et al., 2006, Blume et al., 2015). For example, consider the following utility function for individual i:*

$$u_i(y_1, \cdots, y_n) = \eta_i y_i - \frac{1}{2} y_i^2 + \lambda \sum_{j, j \neq i} w_{ij} y_i y_j.$$

y_i can be an effort level in some activities of individual i. The individual utility depends on a liner term of effort level, a concavity component reflecting decreasing marginal utility, and an interaction component between own effort and the levels of effort of others weighted by a connectivity measure w_{ij} and an interaction coefficient λ. In choosing the effort level y_i, an individual maximizes the own utility given the effort levels of others, which gives the best response equation in Eq.(1).

Assuming that the matrix $I - \lambda W$ is invertible, Eq.(1) describes an equilibrium system of $\{y_i\}_{i=1}^n$

for given η and W. The reduced form is

$$y_i = \ell_i'(I - \lambda W)^{-1}\eta, \tag{2}$$

where ℓ_i is an $n \times 1$ vector with i-th entry 1 and all other entries 0. Note that the Bonacich centrality of unit i given W is $\ell_i'(I - \lambda W)^{-1}\mathbf{1}$ (Ballester *et al.*, 2006) and Eq.(2) can be viewed as a weighted Bonacich centrality of unit i, with weights given by the vector η. Because

$$\begin{aligned}\ell_i'(I - \lambda W)^{-1}\eta &= \ell_i'\left(I + \lambda W + \lambda^2 W^2 + \cdots\right)\eta \\ &= \eta_i + \lambda \sum_j w_{ij}\eta_j + \lambda^2 \sum_{j,k_1} w_{ik_1}w_{k_1 j}\eta_j + \lambda^3 \sum_{j,k_1,k_2} w_{ik_1}w_{k_1 k_2}w_{k_2 j}\eta_j + \cdots,\end{aligned}$$

the weighted Bonacich centrality of unit i is the discounted sum of network links that end in unit i and originate in other units with weights given by vector η and discount factor λ. Units with higher weighted Bonacich centralities have higher equilibrium outcome levels.

We observe that the increase in the aggregate outcome for a shock to unit j is

$$\frac{\partial \sum_i y_i}{\partial \eta_j} = \ell_j'\left(I - \lambda W'\right)^{-1}\mathbf{1}, \tag{3}$$

where $\mathbf{1}$ is a vector of ones, i.e., units with higher values of Bonacich centralities given network matrix W' have larger effects on the aggregate outcome. The main difference between Eq.(3) and Eq.(2) is that the effect of a unit on *other* units is larger if the discounted sum of network links that *start* in unit i and *end* in other units is larger, hence the Bonacich centralities are based on network matrix W' rather than W. If the network is undirected, $W = W'$ and the two measures are equal. To distinguish the effect of a unit on others from the effect on a unit from others due to network asymmetry, we will refer to Eq.(2) as "origin centrality" and Eq.(3) as "destination centrality".

For a specific network link, the effect of varying $w_{j_0 j_1}$ on the equilibrium y_i, while keeping other network links fixed, is given by

$$\frac{\partial y_i}{\partial w_{j_0 j_1}} = \lambda \ell_i'\left(I - \lambda W\right)^{-1}\ell_{j_0}\ell_{j_1}'\left(I - \lambda W\right)^{-1}\eta. \tag{4}$$

From Eq.(4), the marginal effect of varying the intensity of the network link $w_{j_0 j_1}$ on the aggregate equilibrium outcome, $\sum_i y_i$, is

$$\frac{\partial \sum_i y_i}{\partial w_{j_0 j_1}} = \lambda \mathbf{1}'\left(I - \lambda W\right)^{-1}\ell_{j_0}\ell_{j_1}'\left(I - \lambda W\right)^{-1}\eta = \lambda \ell_{j_0}'\left(I - \lambda W'\right)^{-1}\mathbf{1}\ell_{j_1}'\left(I - \lambda W\right)^{-1}\eta. \tag{5}$$

In a model of network interactions, Ballester *et al.* (2006) show that individual outcomes are proportional to their Bonacich centralities, and that the marginal contribution of a unit to the aggregate outcome is given by its own Bonacich centrality and its contribution to the Bonacich centralities of other units. Conventionally, removing a unit, for instance, through lockdown of

an entire city can be viewed as removing all network links that originate or point to the city. In contrast, in our case the policy targets the values of directed network links, for example, through *ex ante* restriction of route-specific transportation or *ex post* contact tracing and quarantine measures contingent on specific travel histories. Eq.(5) shows that the marginal effect of a directed link on the aggregate outcome depends on the interaction between the Bonacich centrality of the destination city ($\ell'_{j_0}\left(I-\lambda W'\right)^{-1}\mathbf{1}$) and the weighted Bonacich centrality of the origin city ($\ell'_{j_1}\left(I-\lambda W\right)^{-1}\eta$). Intuitively, the intensity of population flows between two cities has a stronger effect on the aggregate outcome if the origin city has a higher infection risk or infections in the destination city can affect other cities more.

2.2 The key link in network interactions

Eq.(5) shows the marginal effect of varying the intensity of a network link on the aggregate outcome. In some circumstances, the policy may be binary. For example, either a travel route is shut down or it is open. Ballester *et al.* (2006) provide results on which node's removal from a network results in the largest reduction in the aggregate outcome. We add to their results by showing which network link's removal leads to the largest reduction in the aggregate outcome. Recall that the network link from node j_1 to j_0 is the j_0j_1 element of matrix W, $w_{j_0j_1}$, and the aggregate outcome is the sum of outcome levels of nodes ($\sum_i y_i$). Denote $W^{-j_0j_1}$ the matrix after replacing the j_0j_1 entry of W by zero.

Theorem 1. *Suppose that the network interactions are described by Eq.(2) and* $|\lambda|\max_i\sum_j|w_{ij}|<1$. *Removing the network link* $w_{j_0j_1}$, *i.e., replacing* $w_{j_0j_1}$ *by* 0, *will reduce the the aggregate outcome by* $\lambda w_{j_0j_1}[\left(I-\lambda W^{-j_0j_1\,\prime}\right)^{-1}\mathbf{1}]_{j_0}[(I-\lambda W)^{-1}\eta]_{j_1}$.

The proof is in the appendix. Theorem 1 provides a geometric characterization of the key network links in terms of their impacts on the aggregate outcome if they are removed, which depend on the link intensity ($w_{j_0j_1}$), the interaction coefficient λ, the origin centrality ($[(I-\lambda W)^{-1}\eta]_{j_1}$)[3], and destination centrality ($[\left(I-\lambda W^{-j_0j_1\,\prime}\right)^{-1}\mathbf{1}]_{j_0}$). We discuss two extensions. The next subsection considers the case where units can be affected by the exogenous characteristics of their connected units, and exogenous changes in more network links after a link removal are discussed in Appendix E.

2.3 Contextual effects

Contextual effects reflect changes in outcomes as a result of exposures to similar factors for those who are close. The identification of causal spillover effects in the presence of contextual effects is the focus in many papers in both econometrics (Manski, 1993, Bramoullé *et al.*, 2009, Lee *et al.*, 2010, Blume *et al.*, 2015) and applied fields (e.g., Christakis and Fowler, 2007, Cohen-Cole and Fletcher, 2008). Units can be influenced by the exogenous characteristics of the connected units (contextual

[3]Note that under the setup of Eq.(2), this equals the equilibrium outcome level of the origin unit, y_{j_1}.

effects), and disentangling contextual effects from the endogenous outcomes of the connected units (endogenous effects) helps us understand the mechanisms underlying social interactions and is policy relevant. For example, a part of the debate on whether obesity can spread via social network involves the impact of contextual effects (Christakis and Fowler, 2007, Cohen-Cole and Fletcher, 2008, Fowler and Christakis, 2008). Cohen-Cole and Fletcher (2008) emphasize that school specific factors, such as the prevalence of fast food restaurants, may have explained the endogenous effects in obesity. Ballester and Zenou (2014) show that ignoring contextual effects can lead to wrong policies that target key players of a network. This section extends the key network link analysis in the previous subsection to allow for contextual effects.

The network interactions model with contextual effects is

$$y_i = \left(\sum_{j=1}^{n} w_{ij} y_j\right) \lambda + \eta_i, \qquad \eta_i = x_i'\beta + \left(\sum_{j=1}^{n} w_{ij} x_j'\right) \beta_w + u_i. \tag{6}$$

Theorem 2. *Suppose that the network interactions are described by Eq.(6) and* $|\lambda| \max_i \sum_j |w_{ij}| < 1$. *Removing the network link* $w_{j_0 j_1}$, *i.e., replacing* $w_{j_0 j_1}$ *by* 0, *will reduce the the aggregate outcome by* $\left[\left(I - \lambda W^{-j_0 j_1\,\prime}\right)^{-1} \mathbf{1}\right]_{j_0} w_{j_0 j_1} \left(\lambda [(I - \lambda W)^{-1} \eta]_{j_1} + x_{j_1}' \beta_w\right)$.

The proof can be found in the appendix. When contextual effects are present, the effect of closing a network link on the aggregate outcome can be decomposed into two components: the direct effects due to spillovers from the origin j_1's characteristics to the destination j_0 ($w_{j_0 j_1} x_{j_1}' \beta_w$), and the effect from equilibrium adjustments in outcomes when the link is removed for fixed residual terms. The second component is similar to Theorem 1 and the first component is new in Theorem 2.

Remark 1. *Theorems 1 and 2 consider the key network link to be removed. They can be slightly modified to determine the key network link to be added. Suppose that the* $j_0 j_1$ *entry of* W *is zero and it is to be replaced by* $w_{j_0 j_1}$ *giving* $W^{+j_0 j_1}$. *The aggregate outcome will be increased by* $\lambda w_{j_0 j_1} \left[\left(I - \lambda W^{+j_0 j_1\,\prime}\right)^{-1} \mathbf{1}\right]_{j_0} [(I - \lambda W)^{-1} \eta]_{j_1}$ *if there are no contextual effects. The aggregate outcome will be increased by* $\left[\left(I - \lambda W^{+j_0 j_1\,\prime}\right)^{-1} \mathbf{1}\right]_{j_0} w_{j_0 j_1} \left(\lambda [(I - \lambda W)^{-1} \eta]_{j_1} + x_{j_1}' \beta_w\right)$ *if there are contextual effects.*

Remark 2. *Theorems 1 and 2 can be extended to characterize the impact on aggregate outcomes when a group of network links are removed. Let* S *denote the set of network links that are removed and* W^{-S} *the network matrix with links in* S *removed. The aggregate outcome will be decreased by* $\lambda \sum_{ij \in S} w_{ij} \left[\left(I - \lambda W^{-S\,\prime}\right)^{-1} \mathbf{1}\right]_i [(I - \lambda W)^{-1} \eta]_j$ *if there are no contextual effects. The aggregate outcome will be decreased by* $\sum_{ij \in S} \left[\left(I - \lambda W^{-S\,\prime}\right)^{-1} \mathbf{1}\right]_i w_{ij} \left(\lambda [(I - \lambda W)^{-1} \eta]_j + x_j' \beta_w\right)$ *if there are contextual effects.*

3 Empirical illustration

In this section, we apply the key network link analysis to determine key population flow routes in the spatial spread of Covid-19, using data on confirmed Covid-19 cases and the intensities of between-city and within-city population flows in China. The empirical model is

$$y_i = \left(\sum_{j=1}^{n} w_{ij} y_j \right) \lambda + t_i \gamma + z'_{1i} \beta_1 + z'_{2i} \beta_2 + u_i. \tag{7}$$

We consider two empirical specifications. In the first specification (**Model A**), we use the measures of population flow during January 1 - February 29, 2020 to construct the between and within city population flows. On January 23, 2020, the city of Wuhan was unexpectedly placed under lockdown that travel out of the city was suspended. Since then, individuals' travel decisions were likely affected by either the perceived risk of infection or the public health measures imposed by the government, which in turn was responding to the infection dynamics. Thus, to mitigate the endogeneity issue of the observed population flows between and within cities in Model A, we use the population flow during the same lunar calendar days in 2019 to construct the instrumental variables[4]. As a comparison, in an alternative specification (**Model B**), we estimate the impacts of between and within city population flows during January 1 - 22, 2020, considering that people's traveling behavior was less likely to be affected by the severity of Covid-19 transmission before January 23, 2020. Model A is the main model and Model B is used as a robustness test where the alternative and exogenous population flow matrix is used. Although the population flow matrix in Model B is unlikely to be affected by the unfold of the Covid-19 pandemic, they do not capture some of the population flows information, and we use this specification as a robustness test.

For both models, the population flow weighted average number of infections in other cities may correlate with the error term, because infections can spread in both directions as cities are interconnected through the population flow network. We construct instrumental variables using population flow weighted meteorological variables in other cities similar to Qiu *et al.* (2020). Based on weather characteristics in the existing literature that may strongly predict virus transmission, we control for weather conditions, including temperature, sea level pressure, station pressure, visibility, wind speed, snow depth, precipitation, and a dummy for bad weather. The same set of weather characteristics in other cities weighted by the between city population flow intensities are used as IVs for the spatially lagged dependent variables. In both models, we additionally control for some city level social and economic variables, including population density, GDP per capita, the employment share of primary industry, and the employment share of tertiary industry. The details of the model specifications and variable definitions are summarized in Table 1.

[4]We assume that the shares of destination and origin cities for travel routes in 2019 are the same as those in 2020, i.e., only the travel intensities of people leaving and arriving at cities differ between the years. The travel shares in 2019 were not released by Baidu Migration.

Table 1: Summary of Empirical Model Specifications

	Model A (main model)	Model B
y_i	$log(1 + \#$ of confirmed Covid-19 cases by February 29, 2020)	
z_{1i}	local weather variables$^{\diamond}$ that affect infection rates	
z_{2i}	city characteristics variables$^{\diamond}$	
t_i	average within city pop. flows intensity, Jan 1-Feb 29, 2020	average within city pop. flows intensity, Jan 1-Jan 22, 2020
w_{ij}	average between city pop. flows intensity, Jan 1-Feb 29, 2020	average between city pop. flows intensity, Jan 1-Jan 22, 2020
$\tilde{w}_{ij}$, $\tilde{t}_i$	pop. flows variables, same lunar calendar days in 2019 as Jan 1-Feb 29, 2020	
Endogenous	w_{ij}, t_i, $\sum_{j=1}^{n} w_{ij}y_j$	$\sum_{j=1}^{n} w_{ij}y_j$
IV†	$\tilde{t}_i$, $\sum_{j=1}^{n} \tilde{w}_{ij}z_{1j}$	$\sum_{j=1}^{n} w_{ij}z_{1j}$

†: in addition to the exogenous variables in the model. $^{\diamond}$: the variables are standardized by subtracting their sample averages.

3.1 Data

We collected the numbers of cumulative confirmed Covid-19 cases of 360 cities by February 29, 2020, using data from 32 provincial-level Health Commissions in China. The National Oceanic and Atmospheric Administration (NOAA) provides precipitation, visibility, wind speed, indicators for bad weather (fog, rain or drizzle, snow or ice pellets, hail, thunder, tornados or funnel clouds), average temperature, etc. at the daily level for 362 weather stations in China. To merge the meteorological variables with the cumulative number of Covid-19 cases, we first calculated daily weather variables for each city from station-level weather records, following the inverse distance weighting method. Specifically, for each city, we drew a circle of 100 km from the city's centroid and calculated the weighted average daily weather variables using stations within the 100-km circle. We used the inverse of the distance between the city's centroid and each station as the weight. Second, we calculated the average weather characteristics of each city for each specification, which were then matched with the number of Covid-19 cases, based on the city identifier. The city characteristics variables were collected form the most recent China city statistical yearbooks.

We obtained the data on population movement between and within cities from Baidu Migration[5], which tracks population flows based on mobile phone location data. From the Baidu Migration data, we collected the daily inflow index and outflow index for 360 cities between January 1st and February 29th in 2020, and on the same lunar calendar days in 2019. For each of the 360 cities, Baidu Migration also records the shares of the top 100 origin cities for the population inflow to the city and the shares of the top 100 destination cities for the population outflow from the city. We assumed that the population flow is zero for destination or origin cities outside the top 100 lists. Then the between city population flow intensities were calculated by multiplying the daily

[5] http://qianxi.baidu.com/

migration index of the population flows with the share of the flows[6]. Regarding the within city population flow intensities, Baidu also provided the daily within city migration index for January 1st - February 29th of 2020 and the same lunar calendar days in 2019. Summary statistics are presented in Table 2. The average intensities of within city population flows are smaller in cities without confirmed cases than those in cities with confirmed cases.

An implication of Theorem 1 in the context of the spread of Covid-19 is that cities which receive more population inflows from high risk areas and have high destination centralities could seed more infections in a region. To test this hypothesis, we construct a provincial level data on the number of infections, the intensity of population flows, risks of the origin cities and centralities of the destination cities. We calculate the destination centrality of cities based on the average population flow subnetwork within the province that the city is in between January 1 and February 29, 2020, assuming that the discount factor λ is 0.3. Figure 1 shows a clear positive correlation between the total number of confirmed cases in a province by February 29, 2020 and the average destination centralities of cities within the province. Given this empirical relationship between the number of Covid-19 cases and destination centralities, we proceed to estimate the model parameters and examine the policy implications.

3.2 Estimation results

In Table 3, columns (1), (3) and (5) report the OLS estimates and columns (2), (4) and (6) report the IV estimates from Eq.(7). The estimates from the first stage regressions are reported in Table C.1 in the appendix. Columns (3)-(6) include socioeconomic controls, while columns (1) and (2) do not. The socioeconomic variables are not available in some cities and the sample sizes in columns (3)-(6) are therefore smaller. We control for province fixed effects in all columns. The estimation results in Table 3 show a significantly positive spillover effect of infections in other cities mediated by population flows. The IV estimate $\lambda = 0.180$ (column (2)) implies that a 1% increase in infections in a city where 100,000 people travel to the focal city causes a 0.198% increase in the number of cases in the focal city[7]. The magnitudes of IV estimates on λ are smaller than those of OLS estimates. Population movements between cities lead to the spread of the virus, which can be reduced by travel restrictions.

For the effect of the within city population flow intensity on infections, the OLS coefficients are significantly negative, while the IV estimated coefficients are smaller and statistically insignificant. The negative estimates can be ascribed to the issue of reverse causality because people may avoid going outside when the risk of catching the virus is high. Similar phenomena have been documented in, for example, Fang *et al.* (2020) that the panic effect of covid-19 substantially reduced human

[6]In the event of a slight discrepancy between the population flow intensity calculated by the inflow index of destination cites and that by the outflow index of origination cities for a city pair, we take average of the two intensities.

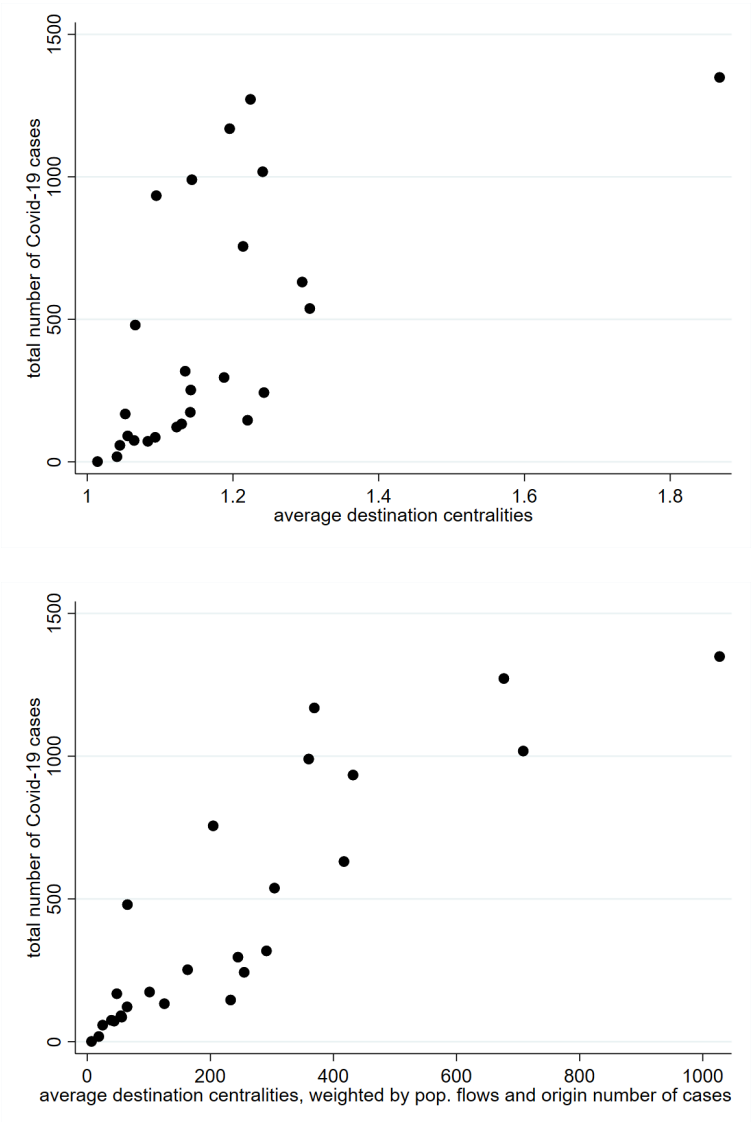
[7]In Fang *et al.* (2020), one migration index unit represents 90,848 people movements. Based on this estimate, we find that the destination city infections increase by 0.198% ($0.180 \times 100000 \div 90848 \times 1\%$) for a 1% increase in the infections in the origin city from where on average 100,000 people per day travel to the destination city from the origin city.

Table 2: Summary Statistics

Variables	N	Mean	SD	Min	Median	Max
Cities with confirmed cases						
Time varying variables, Jan1-Feb29, 2020						
Average confirmed cases	324	246.170	2741.927	1	19	49122
Average within city population flows	324	3.709	0.607	1.808	3.724	5.534
Average temperature, °C	324	3.972	9.341	-25.228	5.110	22.187
Average sea level pressure, kPa	324	102.476	0.422	101.212	102.58.2	103.578
Average station pressure, kPa	324	96.767	6.855	70.370	100.221	102.901
Average visibility, m	324	7.585	3.445	1.790	6.851	18.381
Average wind speed, m/s	324	2.278	0.761	0.958	2.163	5.050
Average snow depth, mm	324	6.434	20.605	0	0.635	188.860
Average precipitation, mm	324	0.218	0.415	0	0.122	4.528
Bad weather	324	0.393	0.185	0	0.381	0.805
Time varying variables, Jan1-Jan22, 2020						
Average within city population flows	324	5.230	0.562	3.096	5.322	6.553
Average temperature, °C	324	3.631	10.093	-26.096	4.381	23.276
Average sea level pressure, kPa	324	102.446	0.452	101.095	102.551	103.656
Average station pressure, kPa	324	96.746	6.838	70.325	100.192	102.710
Average visibility, m	324	6.516	3.741	1.132	5.645	18.400
Average wind speed, m/s	324	2.122	0.795	0.732	1.966	6.077
Average snow depth, mm	324	7.315	22.577	0	0.462	204.816
Average precipitation, mm	324	0.169	0.369	0	0.065	4.233
Bad weather	324	0.399	0.223	0	0.364	0.909
City characteristics						
Population density, 1000 per km2	272	0.433	0.321	0.010	0.363	2.524
Per capita GDP, 10,000RMB	272	5.273	2.992	1.189	4.447	21.549
Primary industry employment share	272	0.021	0.055	0.000	0.005	0.543
Tertiary industry employment share	272	0.527	0.132	0.179	0.533	0.870
Cities without confirmed cases						
Time varying variables, Jan1-Feb29, 2020						
Average within city population flows	36	3.105	0.788	1.761	3.027	4.776
Average temperature, °C	36	-1.025	9.938	-13.691	-4.120	21.666
Average sea level pressure, kPa	36	102.560	0.680	101.212	102.704	103.468
Average station pressure, kPa	36	86.992	10.798	71.031	86.665	102.590
Average visibility, m	36	10.011	4.801	3.765	8.942	18.035
Average wind speed, m/s	36	1.966	0.752	0.887	1.806	3.745
Average snow depth, mm	36	10.795	44.961	0	0.106	260.297
Average precipitation, mm	36	0.385	1.023	0	0.0341	3.961
Bad weather	36	0.187	0.165	0	0.159	0.583
Time varying variables, Jan1-Jan22, 2020						
Average within city population flows	36	4.138	1.039	1.968	4.264	6.366
Average temperature, °C	36	-1.563	10.767	-15.144	-4.811	22.841
Average sea level pressure, kPa	36	102.537	0.712	101.141	102.722	103.545
Average station pressure, kPa	36	86.956	10.826	70.902	86.613	102.611
Average visibility, m	36	9.000	5.324	2.723	7.286	18.250
Average wind speed, m/s	36	1.887	0.812	0.797	1.677	3.680
Average snow depth, mm	36	11.110	49.217	0	0	293.716
Average precipitation, mm	36	0.395	1.074	0	0.0156	4.365
Bad weather	36	0.232	0.181	0	0.224	0.818
City characteristics						
Population density, 1000 per km2	9	0.136	0.118	0.006	0.082	0.385
Per capita GDP, 10,000RMB	9	6.754	4.904	2.540	5.172	16.402
Primary industry employment share	9	0.044	0.047	0.000	0.020	0.112
Tertiary industry employment share	9	0.494	0.110	0.355	0.512	0.673

Please see the text for variable sources.

Figure 1: Total number of Covid-19 cases and destination centralities



Each dot represents a province. The vertical axis represents the total number of confirmed Covid-19 cases in a province by February 29, 2020. The horizontal axis represents the average destination centralities of cities within the province, computed based on the average population flow subnetwork within the province that the city is in between January 1 and February 29, 2020, assuming that the discount factor λ is 0.3. For the lower figure, the destination centrality of a city is weighted by the intensity of population inflows and the number of Covid-19 cases in the origin cities. The province of Hubei and four centrally administered municipalities (Beijing, Chongqing, Shanghai, Tianjin) are not included.

mobility within Wuhan, and in Leeson and Rouanet (2021) that self-limiting behavior attenuated localized social interactions in the context of infectious diseases. Controlling for city characteristics variables, the estimated spillover effects remain stable and significant.

$$y_i = \left(\sum_{j=1}^{n} w_{ij} y_j\right) \lambda + t_i \gamma + z'_{1i} \beta_1 + z'_{2i} \beta_2 + \left(\sum_{j=1}^{n} w_{ij} z'_{2j}\right) \beta_{2w} + u_i. \tag{8}$$

We examine whether the spatial spread of Covid-19 from a city is influenced by its socioeconomic characteristics by testing the presence of contextual effects. The regression equation is Eq.(8) where variable definitions follow Eq.(7). The contextual effects are captured by the term $\sum_{j=1}^{n} w_{ij} z'_{2j}$. The results are reported in columns (5) and (6) of Table 3. Given the same risk of importing infections from other cities, cities that are more closely connected to cities of higher population densities report fewer infections. Because cities with higher population densities report more infections, cities that are more closely connected to them may adopt more stringent public health measures which reduce the disease spread, which also implies that without considering these compensating policy behavior, the risk of importing infections will be understated. Comparing the between city transmission parameters in columns (4) and (6), we indeed find that the between city transmission parameter is underestimated when socioeconomic contextual effects are not controlled for.

To illustrate to what extent our estimates are sensitive to the specification of the population flow matrix, we consider a different specification where the between and within city population flow intensities are averages of the population flows between January 1 and January 22, 2020 (Model B in Table 1). The city of Wuhan was placed under lockdown on January 23, 2020. As is in Jia *et al.* (2020), this specification examines how population flows before the adoption of the large-scale public health measures seed Covid-19 across space. Results are reported in Table 4 and the signs of estimates are similar to the baseline results in Table 3. The estimated cross-city spillover effects are smaller than those in Table 3, which may be due to the fact that average population flows before January 22 only measure part of the population flows that generate spatial spillovers in infections by February 29. The first-stage results for these IV regressions are reported in Table C.2 in the appendix.

3.3 Key link analysis

We illustrate the use of Theorem 1 by calculating the contribution of each population flow route to the total number of infections, based on the specification of column 2 of Table 3. The network links average population flow intensities between January 1st, 2020 and February 29th, 2020. Because travel out of areas with serious risks of infection, such as the city of Wuhan, were already severely restricted during this time period, the key link analysis demonstrates, *given the population flow restrictions already in place* such as the lockdown of Wuhan, the routes whose closure would have the greatest additional effects in reducing infections. The benefit of closing a route in terms of

Table 3: Model A (Main Model): Estimation Results

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Between city transmission (λ)	0.224***	0.180***	0.213***	0.172***	0.361***	0.310***
	(0.0359)	(0.0353)	(0.0322)	(0.0221)	(0.040)	(0.0580)
	[0.247]	[0.198]	[0.234]	[0.189]	[0.397]	[0.341]
Within city transmission (γ)	-0.393***	-0.143	-0.566**	-0.158	-0.587***	-0.167
	(0.141)	(0.136)	(0.187)	(0.190)	(0.173)	(0.182)
Weather controls						
Temperature	-0.00721	-0.0214	-0.0282	-0.0411	-0.0035	-0.0196
	(0.0244)	(0.0248)	(0.0431)	(0.0370)	(0.0445)	(0.0416)
Sea level pressure	-0.700*	-0.889**	-0.976	-1.276*	-0.876	-1.177*
	(0.379)	(0.377)	(0.762)	(0.695)	(0.742)	(0.698)
Station pressure	0.0371*	0.0418**	0.0422	0.0450**	0.0316	0.0347
	(0.0195)	(0.0205)	(0.0252)	(0.0227)	(0.0243)	(0.0219)
Visibility	-0.0183	-0.0379	0.0198	0.00178	0.0213	0.0070
	(0.0286)	(0.0248)	(0.0316)	(0.0279)	(0.0306)	(0.0267)
Wind speed	0.150*	0.188***	0.156	0.170**	0.129	0.146***
	(0.0752)	(0.0709)	(0.0974)	(0.0821)	(0.0884)	(0.0729)
Snow depth	0.000176	0.000349	-0.00196	-0.00176	-0.00065	-0.00084
	(0.00425)	(0.00407)	(0.00248)	(0.00203)	(0.00223)	(0.00175)
Precipitation	-0.157***	-0.181***	-0.289**	-0.329***	-0.239**	-0.288***
	(0.0550)	(0.0503)	(0.124)	(0.118)	(0.106)	(0.106)
Bad weather	0.562	0.517	0.257	0.268	0.110	0.143
	(0.531)	(0.548)	(0.534)	(0.552)	(0.547)	(0.537)
Socioeconomic controls						
Population density			0.439*	0.727**	0.438**	0.619**
			(0.215)	(0.308)	(0.205)	(0.249)
Per capita GDP			-0.0363	-0.00356	-0.0715*	-0.0403
			(0.0402)	(0.0385)	(0.0386)	(0.0411)
Primary industry employment share			-3.626	-3.448*	-2.603	-2.808
			(2.457)	(2.027)	(2.517)	(2.105)
Tertiary industry employment share			0.301	0.345	-0.0613	-0.0690
			(0.471)	(0.402)	(0.425)	(0.424)
Contextual effects						
Population density					-2.097***	-1.523***
					(0.393)	(0.395)
Per capita GDP					0.0734**	0.0304
					(0.0322)	(0.0336)
Primary industry employment share					14.58***	9.062
					(4.589)	(6.248)
Tertiary industry employment share					-0.816	-0.383
					(0.844)	(0.866)
Observations	360	360	281	281	281	281
Province FE	YES	YES	YES	YES	YES	YES

The dependent variable is the log of the number of cumulative confirmed cases by February 29, 2020. The endogenous explanatory variables include the log of cumulative number of confirmed cases in other cities and the intensity of population flows between and within cities. Weather controls are temperature, sea level pressure, station pressure, visibility, wind speed, snow depth, precipitation, and a dummy for bad weather in own cities. Socioeconomic controls are population density, GDP per capita, primary industry employment share, and tertiary industry employment share in own cities. Socioeconomic controls are included in the last four columns, while the first two columns only control for weather variables. The set of these weather variables in other cities weighted by the population flow intensities between cities in 2019, and the within-city population flow intensities in 2019 are used as instrumental variables in the IV regressions. Columns (5)-(6) include the contextual effects of the socioeconomic variables, and use these socioeconomic variables in other cities weighted by the population flow intensities between cities in 2019 as IV. In all models, province fixed effects are included. Elasticity of infection spillovers per 100,000 daily population movements are reported in brackets. Standard errors in parentheses are clustered by provinces. *** p<0.01,** p<0.05, * p<0.1.

Table 4: Model B (Alternative Model): Estimation Results

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Between city transmission (λ)	0.161***	0.133***	0.155***	0.127***	0.241***	0.198***
	(0.0140)	(0.0163)	(0.0149)	(0.0120)	(0.0242)	(0.0317)
	[0.177]	[0.146]	[0.171]	[0.140]	[0.265]	[0.218]
Within city transmission (γ)	0.127	0.121	0.00624	-0.0159	-0.106	-0.0952
	(0.159)	(0.151)	(0.177)	(0.164)	(0.156)	(0.152)
	Weather controls					
Temperature	-0.0376	-0.0378	-0.0408	-0.0430	-0.0134	-0.0241
	(0.0284)	(0.0266)	(0.0398)	(0.0362)	(0.0443)	(0.0427)
Sea level pressure	-0.965**	-1.002**	-1.387*	-1.382**	-1.243	-1.309*
	(0.432)	(0.402)	(0.775)	(0.700)	(0.777)	(0.711)
Station pressure	0.0344	0.0383*	0.0361	0.0399*	0.0258	0.0322
	(0.0226)	(0.0221)	(0.0234)	(0.0214)	(0.0234)	(0.0210)
Visibility	-0.0331	-0.0424*	-0.00088	-0.00519	0.000694	-0.00108
	(0.0245)	(0.0223)	(0.0287)	(0.0260)	(0.0279)	(0.0254)
Wind speed	0.186**	0.203***	0.168*	0.176**	0.154*	0.161**
	(0.0728)	(0.0709)	(0.0869)	(0.0798)	(0.0821)	(0.0743)
Snow depth	0.00105	0.000934	-0.00154	-0.00160	-0.000537	-0.000939
	(0.00450)	(0.00416)	(0.00219)	(0.00197)	(0.00197)	(0.00171)
Precipitation	-0.176***	-0.188***	-0.345**	-0.348***	-0.294**	-0.312***
	(0.0502)	(0.0483)	(0.136)	(0.124)	(0.120)	(0.115)
Bad weather	0.237	0.275	0.0347	0.0742	-0.181	-0.0865
	(0.633)	(0.622)	(0.647)	(0.604)	(0.657)	(0.612)
	Socioeconomic controls					
Population density			0.551**	0.682**	0.487**	0.593***
			(0.217)	(0.268)	(0.200)	(0.224)
Per capita GDP			-0.00758	0.00634	-0.0323	-0.0199
			(0.0349)	(0.0324)	(0.0372)	(0.0348)
Primary industry employment share			-3.442	-3.388*	-2.519	-2.794
			(2.126)	(1.926)	(2.253)	(2.040)
Tertiary industry employment share			-0.217	0.0344	-0.495	-0.288
			(0.376)	(0.384)	(0.350)	(0.390)
	Contextual effects					
Population density					-1.141***	-0.979***
					(0.315)	(0.245)
Per capita GDP					0.0421**	0.0482**
					(0.0202)	(0.0197)
Primary industry employment share					7.923**	4.081
					(3.283)	(4.290)
Tertiary industry employment share					-0.0353	-0.0399
					(0.466)	(0.398)
Observations	360	360	281	281	281	281
Province FE	YES	YES	YES	YES	YES	YES

The dependent variable is the log of the number of cumulative confirmed cases by February 29, 2020. The average intensities of population flows between and within cities are calculated based on data from January 1 and January 22, 2020, which are treated as exogenous. The endogenous explanatory variables include the log of cumulative number of confirmed cases in other cities. Weather controls are temperature, sea level pressure, station pressure, visibility, wind speed, snow depth, precipitation, and a dummy for bad weather in the own cities. Socioeconomic controls are population density, GDP per capita, primary industry employment share, and tertiary industry employment share in own cities. Socioeconomic controls are included in the last four columns, while the first two columns only control for weather variables. The sum of these weather variables in other cities weighted by the population flow intensities between cities in 2020 are used as instrumental variables in the IV regressions. Columns (5)-(6) include the contextual effects of the socioeconomic variables. In all models, province fixed effects are included. Elasticity of infection spillovers per 100,000 daily population movements are reported in brackets. Standard errors in parentheses are clustered by provinces. *** p<0.01,** p<0.05, * p<0.1.

Table 5: Key Network Links, Top 25

Rank	Origin	Destination	Total Effect†	Orig. Centrality‡	Dest. Centrality*	Link Intensity*	Cum. % Reductions Infections	Cum. % Restrictions Pop. Flows
1	Shenzhen	Dongguan	2.62	6.04	2.65	0.94	0.80	0.37
2	Foshan	Guangzhou	2.59	4.44	3.07	1.10	1.71	0.80
3	Guangzhou	Foshan	2.49	5.85	2.32	1.07	1.96	1.21
4	Dongguan	Shenzhen	2.09	4.61	2.92	0.89	2.56	1.56
5	Shanghai	Suzhou	1.60	5.82	2.18	0.71	2.97	1.84
6	Beijing	Langfang	1.49	6.03	1.67	0.85	3.22	2.17
7	Langfang	Beijing	1.46	3.43	2.66	0.91	3.63	2.52
8	Suzhou	Shanghai	1.44	4.48	2.54	0.72	4.09	2.80
9	Dongguan	Guangzhou	1.22	4.61	3.07	0.49	4.33	2.99
10	Shenzhen	Guangzhou	1.21	6.04	3.07	0.37	4.59	3.13
11	Shenzhen	Huizhou	1.16	6.04	1.79	0.61	4.74	3.37
12	Huizhou	Shenzhen	1.15	4.14	2.92	0.54	5.01	3.58
13	Guangzhou	Dongguan	1.11	5.85	2.65	0.40	5.12	3.73
14	Xi'an	Xianyang	1.07	4.80	1.36	0.94	5.19	4.10
15	Guangzhou	Shenzhen	0.95	5.85	2.92	0.31	5.36	4.22
16	Wuhan	Xiaogan	0.84	10.80	1.12	0.39	9.48	4.37
17	Tianjin	Beijing	0.83	4.92	2.66	0.35	9.74	4.51
18	Wuhan	Huanggang	0.83	10.80	1.13	0.38	13.05	4.66
19	Chongqing	Chengdu	0.77	6.36	2.34	0.29	13.19	4.77
20	Beijing	Baoding	0.76	6.03	1.45	0.49	13.25	4.96
21	Xianyang	Xi'an	0.74	2.89	1.84	0.79	13.33	5.27
22	Baoding	Beijing	0.72	3.50	2.66	0.44	13.50	5.44
23	Beijing	Tianjin	0.63	6.03	1.68	0.35	13.57	5.57
24	Huizhou	Dongguan	0.59	4.14	2.65	0.30	13.63	5.69
25	Guangzhou	Qingyuan	0.58	5.85	1.32	0.42	13.65	5.85

This table lists the top 25 population flow routes, ranked by their contributions to the aggregate outcome ($\sum_i y_i$ with y_i the number of Covid-19 cases in logarithms) as given in Theorem 1. The estimates use column 2 of Table 3. †: in log points. ‡: weighted Bonacich centrality of the origin city, $(I - \lambda W)^{-1}\eta$. *: Centrality of the destination city j_0, $(I - \lambda W')^{-1}\mathbf{1}$. *: $w_{\text{destination,origin}}$. Cities in blue/red/purple are in the Pearl River Delta/Jingjinji Metropolitan Region/Yangtze River Delta, respectively. The second last column shows the cumulative percentage reductions in the total number of cases, and the last column shows the cumulative percentage of population flows that are stopped, if the routes with equal or higher rankings are all closed and population flows on unaffected routes do not change.

reduced infections depends on the centrality of the origin city, the centrality of the destination city, and the intensity of population flow on the route. Given the population flow patterns observed in the data, Table 5 lists the top 25 most consequential population flow routes, which are those in the regions with large population outflows to many distant cities in China, such as the regions of the Pearl River Delta, the Jingjinji Metropolitan Region, and the Yangtze River Delta, as indicated by their high destination centralities. For example, the top ranked route is from Shenzhen to Dongguan, which features a relatively high destination centrality as Dongguan is a city with a large population of migrant workers who return to their hometowns in this time period around the Chinese New Year and some can travel long distances. Infections in these cities can affect more cities in the country and pose higher risks for virus spread. The total number of cases is predicted to be 13.65% lower if the top 25 network links are all closed, which constitute only 0.02% of all network links, or 5.85% of the amount of population flows, confirming the value of identifying key network links for targeted policy interventions.

It is interesting to observe that the most influential routes are not necessarily those with the highest population flow intensities, because the centralities of the origin and the destination are also relevant in the determination of the key network links. The most influential routes also may not be the routes whose origins have highest infection risks. While infection risks in the city of Wuhan were high during the sample period as indicated by its high origin centrality, only two routes from Wuhan are ranked in the top 25, one from Wuhan to Xiaogan and the other from Wuhan to Huanggang with both destinations in Hubei province. Given the strict travel restrictions such as the lockdown of Wuhan in place, the intensity of population flows from Wuhan were low already for many destinations, and this lowers the marginal benefits of even more stringent travel controls from Wuhan. Note that this does not imply that the lockdown of Wuhan was not consequential in reducing infections, but only that travel restrictions around Wuhan *in addition* to those imposed would have relatively smaller effects.

The contextual effects can affect the key link analysis. We show in earlier analysis that cities that are closely connected to other cities with high population densities report fewer cases, given the same case importation risks, which may be due to compensating public health measures. The contextual effects partly mitigate the risk of receiving population inflows from high risk cities. The key link ranking taking into account the contextual effects (Table 6) attaches smaller values to the risk of the origin cities as seen in the smaller origin centralities than those in Table 5. Population flow routes that end in cities in the Pearl River Delta, especially Shenzhen and Guangzhou, are still among the high risk routes due to their large population outflows to wider areas, i.e., high destination centralities.

A parameter in the key link analysis is the spatial interaction coefficient λ whose value is empirically estimated. We assess to what extent the key link rankings are sensitive to its value by performing the analysis for alternative values of λ. We consider two scenarios, one with low spatial interaction intensity (i.e., λ taking half the value of the estimate), and the other with high spatial interaction intensity (i.e., λ taking 1.5$\times$ the value of the estimate). The results are reported in

Table 6: Key Network Links with Contextual Effects, Top 25

Rank	Origin	Destination	Total Effect†	Orig. Centrality‡	Dest. Centrality⋆	Link Intensity⋆	Cum. % Reductions Infections	Cum. % Restrictions Pop. Flows
1	Huizhou	Shenzhen	4.69	4.14	6.56	0.54	2.33	0.27
2	Dongguan	Shenzhen	4.52	4.61	6.56	0.89	3.99	0.71
3	Foshan	Guangzhou	3.50	4.44	6.97	1.10	5.00	1.26
4	Guangzhou	Foshan	3.36	5.85	5.17	1.07	5.23	1.79
5	Dongguan	Guangzhou	2.83	4.61	6.97	0.49	5.74	2.03
6	Huizhou	Dongguan	2.49	4.14	6.04	0.30	6.03	2.18
7	Qingyuan	Guangzhou	1.81	2.56	6.97	0.30	6.39	2.33
8	Huizhou	Guangzhou	1.68	4.14	6.97	0.17	6.71	2.42
9	Guangzhou	Dongguan	1.66	5.85	6.04	0.40	6.75	2.62
10	Yueyang	Changsha	1.45	5.06	2.76	0.34	7.36	2.79
11	Guangzhou	Shenzhen	1.41	5.85	6.56	0.31	7.43	2.94
12	Zhuhai	Zhongshan	1.29	4.60	2.72	0.38	7.61	3.13
13	Wuhan	Huanggang	1.22	10.80	1.29	0.38	13.41	3.32
14	Wuhan	Xiaogan	1.21	10.80	1.24	0.39	19.92	3.51
15	Xi'an	Xianyang	1.17	4.80	1.85	0.94	20.03	3.98
16	Xianyang	Xi'an	1.15	2.89	2.67	0.79	20.15	4.37
17	Huanggang	Wuhan	1.08	7.97	2.12	0.20	39.34	4.47
18	Zhaoqing	Foshan	1.06	3.00	5.17	0.21	39.39	4.58
19	Dongguan	Huizhou	1.05	4.61	3.63	0.34	39.42	4.75
20	Xiaogan	Wuhan	1.02	8.17	2.12	0.23	50.89	4.86
21	Kunming	Qujing	0.99	3.99	1.40	0.55	50.93	5.13
22	Mianyang	Chengdu	0.98	3.14	3.37	0.30	51.04	5.28
23	Yiyang	Changsha	0.96	4.11	2.76	0.33	51.22	5.45
24	Changsha	Yueyang	0.94	5.49	1.64	0.39	51.34	5.64
25	Zhaoqing	Guangzhou	0.94	3.00	6.97	0.14	51.40	5.71

This table lists the top 25 population flow routes, ranked by their contributions to the aggregate outcome as given in Theorem 2. The estimates use column 6 of Table 3. †: in log points. ‡: weighted Bonacich centrality of the origin city, $(I - \lambda W)^{-1}\eta$. ⋆: Centrality of the destination city, $(I - \lambda W')^{-1}\mathbf{1}$. ⋆: $w_{\text{destination,origin}}$. Cities in blue are in the Pearl River Delta. The second last column shows the cumulative percentage reductions in the total number of cases, and the last column shows the cumulative percentage of population flows that are stopped, if the routes with equal or higher rankings are all closed and population flows on unaffected routes do not change.

Tables D.1 and D.2 in the appendix. The rankings are similar in both scenarios, and the key link targeting is more effective in reducing infections when the spatial interaction intensity is higher.

4 Conclusion and discussion

Network interactions are ubiquitous. The outcomes or actions of individuals can depend on those of other connected individuals. The literature has characterized the key network node whose removal can influence the aggregate outcome the most. Besides network nodes, the network links between nodes can also be targets of policy interventions (Mellon *et al.*, 2016, Philip *et al.*, 2016), with the objective of reshaping aggregate outcome levels. In this paper, we provide a geometric interpretation of the factors affecting the contribution of a network link to the aggregate outcome, and show that the importance of a network link depends on the centralities of the origin and the destination nodes, along with their link intensity. We discuss how the theoretical results can accommodate contextual effects and networks that can change in response to the interventions.

We apply the model to examine the spread of Covid-19 in China. Restrictions on travel are frequently imposed as part of governments' responses to the Covid-19 pandemic. Blanket bans or lockdowns of entire regions could incur significant social and economic costs, which may outweigh the benefits from reduced infections, and erode public support for the epidemic control and prevention measures. In this paper, we show that the marginal effect of decreasing intercity population flows in reducing total infections is not homogeneous, but rather depends on the positions of the origin and destination cities in the network of population flows. The benefit of restricting population flows on a route is larger if the origin city is closely connected to areas with severe infections, or the infections in the destination city can spill over to many other cities. Population flow restrictions that target these links could be more cost-effective. Importantly, the key routes may not necessarily be those with highest link intensities.

These findings may have rich implications for virus mitigation strategies that go beyond imposing *ex ante* route-specific travel restrictions to optimize *ex post* management. For regions at the beginning of an epidemic [8] or with inadequate resources, systematic infection screening and personnel training may take time and are demanding. The incubation period and high prevalence of asymptomatic infections may also limit the effectiveness of screening vital signs or self-reporting of symptoms (World Health Organization, 2020). Therefore, a number of economies have adopted innovative approaches in their strategies to effectively curb spread of the virus. For instance, tools such as migration maps, which collect real-time data on the location of people via mobile phones, mobile payment applications and social media, allow mainland China to track the movement of people who flowed out of Wuhan or other high risk areas. These data also guide border checks and surveillance (Wu *et al.*, 2020, Liu, 2020). Taiwan initiated health checks for airline travelers

[8] Our findings may also apply to later stages of a pandemic when the original strain of virus spreading is under control but more transmissible new variants, such as Delta and Omicron, continue to emerge locally or to be introduced through cross-border transportation to domestic areas, which leaves less time to respond due to the high contagiousness.

from Wuhan and surrounding cities, integrating data from immigration records with its centralized, real-time health insurance database. This integration allowed healthcare facilities to access patients' travel histories and identify high-risk individuals for testing and tracking (Wang *et al.*, 2020). South Korea's aggressive contact tracing using security camera footage, facial recognition technology, bank card records and GPS data from vehicles and mobile phones provides real-time data and detailed timelines of people's travel, which facilitates targeted screening and timely quarantine (Fisher and Sang-Hun, 2020). Such mobile technology will continue to help advance policies on travel restrictions while striking a balance between privacy concerns and public welfare.

Finally, large negative infectious disease externalities are often inherent to pandemics, given the fact that first movers often bear the largest costs and therefore have insufficient incentive to internalize these externalities. While governments have largely responded to COVID-19 with costly public health interventions, most notably lockdowns of cities or travel routes that restrict human interactions, many missed the best window of opportunity. To motivate local governments to impose the route-specific travel restrictions contingent on their heterogeneous risks at the very beginning of a pandemic, targeted support and subsidies from the central government to first movers are warranted. For instance, preferential policies aimed at reviving and sustaining business operations of Small and Micro Enterprises (SMEs) have been implemented by both central and provincial governments, including reliefs on taxation, finance, social security, subsidies, and rent reduction (Okyere *et al.*, 2020). Overall, the incurred costs are more than offset by benefits accrue to the aggregate level. Moreover, coordinated restrictions among governments may also reduce the required restrictions and associated costs to achieve the same objective of combating a pandemic, which generates additional benefits to all parties (Holtz *et al.*, 2020).

References

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, M. D. (2020) A multi-risk SIR model with optimally targeted lockdown, *NBER Working Paper 27102*.

Alfaro, L., Faia, E., Lamersdorf, N. and Saidi, F. (2020) Social interactions in pandemics: Fear, altruism, and reciprocity, *NBER Working Paper 27134*.

Alvarez, F. E., Argente, D. and Lippi, F. (2020) A simple planning problem for COVID-19 lockdown, *NBER Working Paper 26981*.

Ballester, C., Calvó-Armengol, A. and Zenou, Y. (2006) Who's who in networks. Wanted: The key player, *Econometrica*, **74**, 1403–1417.

Ballester, C. and Zenou, Y. (2014) Key player policies when contextual effects matter, *Journal of Mathematical Sociology*, **38**, 233–248.

Ballester, C., Zenou, Y. and Calvó-Armengol, A. (2010) Delinquent networks, *Journal of the European Economic Association*, **8**, 34–61.

Berger, D. W., Herkenhoff, K. F. and Mongey, S. (2020) An seir infectious disease model with testing and conditional quarantine, *NBER Working Paper 26901*.

Blume, L. E., Brock, W. A., Durlauf, S. N. and Jayaraman, R. (2015) Linear social interactions models, *Journal of Political Economy*, **123**, 444–496.

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009) Identification of peer effects through social networks, *Journal of Econometrics*, **150**, 41–55.

Brinkman, J. and Mangum, K. (2021) JUE insight: The geography of travel behavior in the early phase of the COVID-19 pandemic, *Journal of Urban Economics*.

Brockmann, D. and Helbing, D. (2013) The hidden geometry of complex, network-driven contagion phenomena, *Science*, **342**, 1337–1342.

Christakis, N. A. and Fowler, J. H. (2007) The spread of obesity in a large social network over 32 years, *New England Journal of Medicine*, **357**, 370–379.

Clark, A. E., Nong, H., Zhu, H. and Zhu, R. (2021) Compensating for academic loss: Online learning and student performance during the COVID-10 pandemic, *China Economic Review*, **68**.

Cohen-Cole, E. and Fletcher, J. M. (2008) Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic, *Journal of Health Economics*, **27**, 1382–1387.

Comola, M. and Prina, S. (2021) Treatment effect accounting for network changes, *Review of Economics and Statistics*, **103**, 597–604.

Dai, R., Feng, H., Hu, J., Jin, Q., Li, H., Wang, R., Wang, R., Xu, L. and Zhang, X. (2021) The impact of COVID-19 on small and medium-sized enterprises (SMEs): Evidence from two-wave phone surveys in China, *China Economic Review*, **67**.

Deng, L. and Sun, Y. (2017) Criminal network formation and optimal detection policy: The role of cascade of detection, *Journal of Economic Behavior & Organization*, **141**, 43–63.

Duan, H., Bao, Q., Tian, K., Wang, S., Yang, C. and Cai, Z. (2021) The hit of the novel coronavirus outbreak to China's economy, *China Economic Review*, **67**.

Emanuel, E. J., Zhang, C. and Glickman, A. (2020) Learning from Taiwan about responding to Covid-19 - and using electronic health records, https://www.statnews.com/2020/06/30/taiwan-lessons-fighting-covid-19-using-electronic-health-records/.

Fajgelbaum, P. D., Khandelwal, A., Kim, W., Mantovani, C. and Schaal, E. (2020) Optimal lockdown in a commuting network, *Working Paper*.

Fang, H., Wang, L. and Yang, Y. (2020) Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China, *Journal of Public Economics*, **191**.

Fisher, M. and Sang-Hun, C. (2020) How South Korea flattened the curve, https://www.nytimes.com/2020/03/23/world/asia/coronavirus-south-korea-flatten-curve.html.

Fowler, J. H. and Christakis, N. A. (2008) Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais, *Journal of Health Economics*, **27**, 1400–1405.

Griffith, A. (2021) Random assignment with non-random peers: A structural approach to counterfactual treatment assessment, *Working Paper*.

Holtz, D., Zhao, M., Benzell, S. G., Cao, C. Y., Rahimian, M. A., Yang, J., Allen, J., Collis, A., Moehring, A., Sowrirajan, T., Ghosh, D., Zhang, Y., Dhillon, P., Nicolaides, C., Eckles, D. and Aral, S. (2020) Interdependence and the cost of uncoordinated response to COVID-19, *Proceedings of the National Academy of Sciences*, **117**, 19837–19843.

Hsieh, C.-S., König, M. D. and Liu, X. (2020) A structural model for the coevolution of networks and behavior, *Review of Economics and Statistics*.

Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jia, J. and Christakis, N. A. (2020) Population flow drives spatio-temporal distribution of COVID-19 in China, *Nature*.

Kermack, W. O. and McKendrick, A. G. (1927) A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society A, Mathematical, Physical and Engineering Sciences*, **115**, 700–721.

Lee, L.-f. (2007) Identification and estimation of econometric models with group interactions, contextual factors and fixed effects, *Journal of Econometrics*, **140**, 333–374.

Lee, L.-F., Liu, X. and Lin, X. (2010) Specification and estimation of social interaction models with network structures, *Econometrics Journal*, **13**, 145–176.

Lee, L.-F., Liu, X., Patacchini, E. and Zenou, Y. (2020) Who is the key player? a network analysis of juvenile delinquency, *Journal of Business & Economic Statistics*.

Leeson, P. T. and Rouanet, L. (2021) Externality and COVID-19, *Southern Economic Journal*, **87**, 1107–1118.

Liu, J. (2020) Deployment of health IT in China's fight against the COVID-19 pandemic, https://www.itnonline.com/article/deployment-health-it-china\OT1\textquoterights-fight-against-covid-19-pandemic.

Manski, C. (1993) Identification of endogenous social effects: The reflection problem, *Review of Economic Studies*, **60**, 531–542.

Manski, C. F. (2004) Statistical treatment rules for heterogeneous population, *Econometrica*, **72**, 1221–1246.

Manski, C. F. and Molinari, F. (2021) Estimating the covid-19 infection rate: Anatomy of an inference problem, *Journal of Econometrics*, **220**, 181–192.

Mele, A. (2017) A structural model of dense network formation, *Econometrica*, **85**, 825–850.

Mellon, J., Yoder, J. and Evans, D. (2016) Undermining and strengthening social networks through network modification, *Scientific Reports*, **6**.

Meng, D., Sun, L. and Tian, G. (2022) Dynamic mechanism design on social networks, *Games and Economic Behavior*, **131**, 84–120.

Okyere, M. A., Forson, R. and Essel-Gaisey, F. (2020) Positive externalities of an epidemic: The case of the coronavirus (COVID-19) in China, *Journal of Medical Virology*, **92**, 1376–1379.

Patacchini, E., Rainone, E. and Zenou, Y. (2017) Heterogeneous peer effects in education, *Journal of Economic Behavior & Organization*, **134**, 190–227.

Philip, J., Ford, T., Henry, D., Rasmus, S. and Allen, J. (2016) Relationship of social network to protective factors in suicide and alcohol use disorder intervention for rural Yup'ik Alaska native youth, *Psychosocial Intervention*, **25**.

Qiu, Y., Chen, X. and Shi, W. (2020) Impacts of social and economic factors on the transmission of Coronavirus Disease 2019 (COVID-19) in China, *Journal of Population Economics*.

Sun, Y., Zhao, W. and Zhou, J. (2021) Structural interventions in networks, *arXiv 2101.12420*.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U. G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O. N., Yang, R., Grenfell, B. T., Pybus, O. G. and Dye, C. (2020) An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China, *Science*.

Wang, C. J., Ng, C. Y. and Brook, R. H. (2020) Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing, *JAMA*, **323**, 1341–1342.

Whitelaw, S., Mamas, M. A., Topol, E. and Van Spall, H. G. C. (2020) Applications of digital technology in COVID-19 pandemic planning and response, *The Lancet Digital Health*.

World Health Organization (2020) Key considerations for repatriation and quarantine of travellers in relation to the outbreak of novel coronavirus 2019-nCoV.

Wu, J. T., Leung, K. and Leung, G. M. (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study, *Lancet*.

A Proof of Theorem 1

Let $-j_0j_1$ denote the scenario when the network link $w_{j_0j_1}$ is removed. Under the assumption that $|\lambda| \max_i \sum_j |w_{ij}| < 1$, $I - \lambda W$ and $I - \lambda W^{-j_0j_1}$ are invertible and the equilibrium outcome described by Eq.(2) exists and is unique. From Eq.(2),

$$\begin{aligned} y_s^{-j_0j_1} - y_s &= \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \eta - \ell_s' \left(I - \lambda W\right)^{-1} \eta \qquad \text{(A.1)} \\ &= -\lambda \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \left(W - W^{-j_0j_1}\right) \left(I - \lambda W\right)^{-1} \eta \\ &= -\lambda w_{j_0j_1} \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \ell_{j_0} \ell_{j_1}' \left(I - \lambda W\right)^{-1} \eta \end{aligned}$$

where in Eq.(A.1) we have used the identity $A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}$ for invertible matrices A and B.

Summing over s, the impact of removing $w_{j_0j_1}$ on the aggregate outcome is

$$\begin{aligned} \sum_{s=1}^{n} \left(y_s - y_s^{-j_0j_1}\right) &= \lambda w_{j_0j_1} \mathbf{1}' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \ell_{j_0} \ell_{j_1}' \left(I - \lambda W\right)^{-1} \eta \\ &= \lambda w_{j_0j_1} [\left(I - \lambda W^{-j_0j_1\,\prime}\right)^{-1} \mathbf{1}]_{j_0} [(I - \lambda W)^{-1} \eta]_{j_1}. \end{aligned}$$

B Proof of Theorem 2

The proof mirrors the proof of Theorem 1 with some modifications.

$$\begin{aligned} y_s^{-j_0j_1} - y_s &= \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \eta^{-j_0j_1} - \ell_s' \left(I - \lambda W\right)^{-1} \eta \\ &= \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \left(\eta^{-j_0j_1} - \eta\right) + \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \eta - \ell_s' \left(I - \lambda W\right)^{-1} \eta \\ &= -\ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \ell_{j_0} w_{j_0j_1} x_{j_1}' \beta_w + \ell_s' \left(I - \lambda W^{-j_0j_1}\right)^{-1} \eta - \ell_s' \left(I - \lambda W\right)^{-1} \eta. \end{aligned}$$

The proof is completed by using the result in Eq.(A.1) and summing over s.

C First stage results

Table C.1 presents the first stage estimates for the IV regressions for Model A as described in Table 3. Columns (1) and (3) display the coefficients on the sum of cumulative confirmed cases in other cities weighted by population flows. Columns (2) and (4) present the coefficients on within city population flow intensity. The estimates of the first stage regressions for Model B are reported in Table C.2. We also report the R-squared and F-test statistics for the joint significance of excluded instruments in the first stage.

D Key network links under alternative specifications

Please see Tables D.1 and D.2.

Table C.1: Model A: First Stage Results

VARIABLES	(1) Wy	(2) t	(3) Wy	(4) t	(5) Wy	(6) t
	Own city					
Within-city population flow intensities in 2019	-0.245	0.760***	-0.435	0.859***	-0.274	0.862***
	(0.219)	(0.122)	(0.294)	(0.0905)	(0.232)	(0.0976)
Temperature	-0.0463	0.0285	-0.0167	-0.00329	-0.00025	-0.00362
	(0.0911)	(0.0182)	(0.0614)	(0.00827)	(0.0554)	(0.00851)
Sea level pressure	-0.520	0.273*	1.017	0.148	1.222	0.150
	(1.033)	(0.159)	(1.072)	(0.190)	(1.083)	(0.203)
Station pressure	0.0230	-0.00022	0.00566	0.00241	0.00530	0.00299
	(0.0504)	(0.00933)	(0.0309)	(0.00673)	(0.0277)	(0.00656)
Visibility	-0.110***	0.0195***	-0.0663**	0.0243***	-0.0332	0.0231***
	(0.0357)	(0.00668)	(0.0296)	(0.00558)	(0.0273)	(0.00552)
Wind speed	0.0956	0.025	0.0662	-0.0149	0.0608	-0.0142
	(0.122)	(0.0272)	(0.0969)	(0.0267)	(0.0828)	(0.0273)
Snow depth	0.000742	0.00135	-0.0046	0.000786**	-0.00255	0.00082**
	(0.00467)	(0.00128)	(0.00320)	(0.000301)	(0.00202)	(0.000338)
Precipitation	0.0401	0.0170	0.234	0.00583	0.128	0.00212
	(0.246)	(0.0343)	(0.184)	(0.0216)	(0.134)	(0.0220)
Bad weather	0.0145	-0.0435	-0.241	0.0109	-0.415	0.00990
	(0.660)	(0.148)	(0.669)	(0.159)	(0.495)	(0.164)
Population density			0.700	-0.0660	0.804	-0.0866
			(0.549)	(0.141)	(0.538)	(0.143)
Per capita GDP			0.0624	-0.00542	0.0611	-0.00476
			(0.0520)	(0.0128)	(0.0527)	(0.0130)
Primary industry employment share			-2.733	0.406*	-2.770	0335
			(2.362)	(0.207)	(2.256)	(0.218)
Tertiary industry employment share			2.145**	0.133	2.389**	0.0769
			(1.059)	(0.167)	(0.988)	(0.173)
	Other cities, weight = population flow					
Temperature	0.115	-0.0194**	0.0730	-0.00448	-0.105	-0.0100
	(0.137)	(0.00829)	(0.0908)	(0.00547)	(0.0905)	(0.00620)
Sea level pressure	-1.632	-0.245	-1.862	-0.0179	-3.883*	-0.0343
	(3.047)	(0.178)	(2.275)	(0.148)	(2.002)	(0.152)
Station pressure	0.0803	0.00498	0.108**	-0.00167	0.190***	-0.00793
	(0.0747)	(0.00926)	(0.0474)	(0.00784)	(0.0487)	(0.00728)
Visibility	-0.394*	0.00818	-0.368*	0.00871	-0.244	0.0218
	(0.214)	(0.0337)	(0.178)	(0.0281)	(0.178)	(0.0365)
Wind speed	1.025	-0.00915	1.405***	-0.00542	1.197***	-0.00544
	(0.729)	(0.0490)	(0.459)	(0.0686)	(0.399)	(0.0592))
Snow depth	0.0264	-0.00173	0.0112	0.000542	0.0132	0.00107
	(0.0235)	(0.00239)	(0.0166)	(0.00141)	(0.0128)	(0.00126)
Precipitation	-6.209**	0.314*	-6.337**	0.321	-2.891	0.321
	(2.761)	(0.157)	(2.334)	(0.233)	(2.095)	(0.224)
Bad weather	2.894	-0.128	2.681	-0.0962	1.396	0.139
	(3.016)	(0.282)	(2.759)	(0.301)	(2.408)	(0.373)
Population density					-0.0212	0.228
					(0.657)	(0.137)
Per capita GDP					-0.00913	-0.00868
					(0.0664)	(0.00977)
Primary industry employment share					-63.88***	0.302
					(17.56)	(1.450)
Tertiary industry employment share					4.069***	-0.424
					(0.998)	(0.289)
First-stage R^2	0.862	0.846	0.889	0.865	0.915	0.869
F-test of excluded instruments	15.73	42.95	37.89	28.11	214.5	23.96
F-test p-value	0.000	0.000	0.000	0.000	0.000	0.000
Observations	360	360	281	281	281	281
Province FE	YES	YES	YES	YES	YES	YES

This table reports the first stage results for the weighted sum of cumulative confirmed cases in other cities and the intensities of population flows within cities. The first-stage R-squared and F-tests for the joint significance of excluded instruments in the first stages are reported. Weather controls and socioeconomic controls are included in the last two columns, while the first two columns only control weather variables. Columns (5)-(6) include the contextual effects of the socioeconomic variables.Standard errors in parentheses are clustered by provinces. *** p<0.01,** p<0.05, * p<0.1.

Table C.2: Model B: First Stage Results

VARIABLES	(1) Wy	(2) Wy	(3) Wy
Own city			
Within-city population flow intensities in 2020	0.116	-0.449	-0.130
	(0.376)	(0.498)	(0.247)
Temperature	-0.0992	-0.0811	-0.0741
	(0.0693)	(0.0987)	(0.0880)
Sea level pressure	-0.686	0.438	0.571
	(1.234)	(1.627)	(1.541)
Station pressure	-0.0398	0.0263	0.0364
	(0.0574)	(0.0514)	(0.0461)
Visibility	-0.0775	-0.102**	-0.0536
	(0.0533)	(0.0450)	(0.0436)
Wind speed	-0.151	0.179	0.162
	(0.198)	(0.195)	(0.154)
Snow depth	-0.000556	-0.00511	-0.00259
	(0.00985)	(0.00455)	(0.00307)
Precipitation	0.409	0.223	0.115
	(0.539)	(0.307)	(0.241)
Bad weather	1.372	0.703	0.255
	(1.396)	(0.885)	(0.694)
Population density		1.136	1.358*
		(0.769)	(0.787)
Per capita GDP		0.166**	0.181**
		(0.0624)	(0.0671)
Primary industry employment share		-3.998	-4.046
		(3.349)	(3.519)
Tertiary industry employment share		3.077*	3.434**
		(1.527)	(1.302)
Other cities, weight = population flow			
Temperature	0.204	0.191	-0.107
	(0.208)	(0.156)	(0.168)
Sea level pressure	-1.158	-1.200	-4.965
	(4.318)	(3.699)	(3.218)
Station pressure	0.177	0.129*	0.233**
	(0.114)	(0.0736)	(0.104)
Visibility	-0.803**	-0.855***	-0.635*
	(0.376)	(0.297)	(0.351)
Wind speed	1.393	2.096**	1.485**
	(0.992)	(0.755)	(0.719)
Snow depth	0.00239	0.0268	0.0293
	(0.0463)	(0.0282)	(0.0244)
Precipitation	-10.76**	-8.692**	-4.259
	(4.152)	(4.090)	(3.609)
Bad weather	1.178	2.953	1.493
	(3.964)	(3.840)	(3.264)
Population density			-0.166
			(0.971)
Per capita GDP			0.117
			(0.105)
Primary industry employment share			-96.46**
			(38.08)
Tertiary industry employment share			3.760*
			(2.033)
First-stage R^2	0.790	0.903	0.922
F-test of excluded instruments	39.44	76.03	21.43
F-test p-value	0.000	0.000	0.000
Observations	360	281	281
Province FE	YES	YES	YES

This table reports the first stage results for the weighted sum of cumulative confirmed cases in other cities. The first-stage R-squared and F-tests for the joint significance of excluded instruments in the first stages are reported. Weather controls and socioeconomic controls are included in the last two columns, while the first column only controls weather variables. Column (3) includes the contextual effects of the socioeconomic variables. Standard errors in parentheses are clustered by provinces. *** p<0.01,** p<0.05, * p<0.1.

Table D.1: Key Network Links, Low Interaction Intensity, Top 25

Rank	Origin	Destination	Total Effect†	Orig. Centrality‡	Dest. Centrality*	Link Intensity*	Cum. % Reductions Infections	Cum. % Restrictions Pop. Flows
1	Shenzhen	Dongguan	0.53	3.91	1.60	0.94	0.07	0.37
2	Foshan	Guangzhou	0.51	2.89	1.78	1.10	0.17	0.80
3	Guangzhou	Foshan	0.47	3.39	1.47	1.07	0.21	1.21
4	Beijing	Langfang	0.37	3.91	1.23	0.85	0.23	1.54
5	Dongguan	Shenzhen	0.37	2.66	1.72	0.89	0.31	1.89
6	Xi'an	Xianyang	0.37	3.83	1.14	0.94	0.32	2.25
7	Wuhan	Xiaogan	0.35	9.68	1.05	0.39	2.23	2.41
8	Shanghai	Suzhou	0.35	3.79	1.47	0.71	2.28	2.68
9	Wuhan	Huanggang	0.35	9.68	1.05	0.38	3.64	2.83
10	Langfang	Beijing	0.34	2.51	1.69	0.91	3.70	3.18
11	Suzhou	Shanghai	0.30	2.93	1.62	0.72	3.77	3.46
12	Shenzhen	Huizhou	0.27	3.91	1.27	0.61	3.79	3.70
13	Huizhou	Shenzhen	0.26	3.09	1.72	0.54	3.84	3.91
14	Shenzhen	Guangzhou	0.23	3.91	1.78	0.37	3.88	4.05
15	Xianyang	Xi'an	0.22	2.26	1.36	0.79	3.90	4.36
16	Tianjin	Beijing	0.22	4.05	1.69	0.35	3.95	4.50
17	Dongguan	Guangzhou	0.21	2.66	1.78	0.49	3.98	4.69
18	Beijing	Baoding	0.20	3.91	1.17	0.49	3.99	4.88
19	Xiaogan	Wuhan	0.20	7.55	1.28	0.23	10.18	4.97
20	Guangzhou	Dongguan	0.20	3.39	1.60	0.40	10.20	5.12
21	Chongqing	Chengdu	0.19	4.75	1.58	0.29	10.22	5.23
22	Yueyang	Changsha	0.19	4.52	1.34	0.34	10.27	5.37
23	Xi'an	Weinan	0.19	3.83	1.08	0.50	10.27	5.56
24	Baoding	Beijing	0.18	2.80	1.69	0.44	10.31	5.73
25	Kunming	Qujing	0.18	3.32	1.10	0.55	10.31	5.95

This table lists the top 25 population flow routes, ranked by their contributions to the aggregate outcome as given in Theorem 1. The estimates use column 2 of Table 3 with the exception that λ is half the estimated value. †: in log points. ‡: weighted Bonacich centrality of the origin city, $(I-\lambda W)^{-1}\eta$. *: Centrality of the destination city, $(I-\lambda W')^{-1}\mathbf{1}$. *: $w_{\text{destination,origin}}$. Cities in blue/red/purple are in the Pearl River Delta/Jingjinji Metropolitan Region/Yangtze River Delta, respectively. The second last column shows the cumulative percentage reductions in the total number of cases, and the last column shows the cumulative percentage of population flows that are stopped, if the routes with equal or higher rankings are all closed and population flows on unaffected routes do not change.

Table D.2: Key Network Links, High Interaction Intensity, Top 25

Rank	Origin	Destination	Total Effect†	Orig. Centrality‡	Dest. Centrality*	Link Intensity⋆	Cum. % Reductions Infections	Cum. % Restrictions Pop. Flows
1	Foshan	Guangzhou	11.61	7.80	5.60	1.10	18.43	0.43
2	Shenzhen	Dongguan	11.52	10.29	4.81	0.94	24.81	0.80
3	Guangzhou	Foshan	11.41	10.84	4.10	1.07	26.03	1.21
4	Dongguan	Shenzhen	10.06	8.57	5.31	0.89	29.04	1.56
5	Dongguan	Guangzhou	6.11	8.57	5.60	0.49	30.05	1.75
6	Shanghai	Suzhou	5.65	9.12	3.37	0.71	32.07	2.03
7	Shenzhen	Guangzhou	5.58	10.29	5.60	0.37	32.95	2.17
8	Guangzhou	Dongguan	5.48	10.84	4.81	0.40	33.30	2.32
9	Suzhou	Shanghai	5.28	7.05	4.05	0.72	35.10	2.60
10	Beijing	Langfang	5.00	9.39	2.48	0.85	36.75	2.93
11	Langfang	Beijing	4.98	5.12	4.22	0.91	38.04	3.29
12	Huizhou	Shenzhen	4.81	6.48	5.31	0.54	38.90	3.50
13	Shenzhen	Huizhou	4.77	10.29	2.93	0.61	39.16	3.73
14	Guangzhou	Shenzhen	4.72	10.84	5.31	0.31	39.63	3.86
15	Xi'an	Xianyang	2.59	6.30	1.73	0.94	39.81	4.22
16	Tianjin	Beijing	2.53	6.35	4.22	0.35	40.52	4.36
17	Huizhou	Dongguan	2.49	6.48	4.81	0.30	40.67	4.48
18	Chongqing	Chengdu	2.45	9.05	3.51	0.29	41.15	4.59
19	Beijing	Baoding	2.37	9.39	1.95	0.49	41.34	4.78
20	Baoding	Beijing	2.30	4.73	4.22	0.44	41.73	4.95
21	Dongguan	Huizhou	2.25	8.57	2.93	0.34	41.79	5.08
22	Guangzhou	Qingyuan	2.20	10.84	1.82	0.42	41.87	5.24
23	Beijing	Tianjin	2.04	9.39	2.36	0.35	42.01	5.38
24	Xianyang	Xi'an	2.00	3.90	2.54	0.79	42.15	5.69
25	Chengdu	Chongqing	1.85	7.60	2.69	0.34	42.80	5.82

This table lists the top 25 population flow routes, ranked by their contributions to the aggregate outcome as given in Theorem 1. The estimates use column 2 of Table 3 with the exception that λ is 1.5× the estimated value. †: in log points. ‡: weighted Bonacich centrality of the origin city, $(I - \lambda W)^{-1}\eta$. *: Centrality of the destination city, $(I - \lambda W')^{-1}\mathbf{1}$. ⋆: $w_{\text{destination,origin}}$. Cities in blue/red/purple are in the Pearl River Delta/Jingjinji Metropolitan Region/Yangtze River Delta, respectively. The second last column shows the cumulative percentage reductions in the total number of cases, and the last column shows the cumulative percentage of population flows that are stopped, if the routes with equal or higher rankings are all closed and population flows on unaffected routes do not change.

E Exogenous changes in multiple links

Networks may rewire after interventions that target components of the networks. Lee *et al.* (2020) extend the key player analysis of Ballester *et al.* (2006) by accounting for network changes after the removal of network nodes and show that the key player rankings can be different if networks are assumed fixed but in fact may change after the interventions. Some recent research on policy evaluations in the presence of between unit interactions have also used various ways to take into account how networks change in response to the interventions (Comola and Prina, 2021, Griffith, 2021). We show how the key link analysis can be extended to allow for network changes after the removal of a link. As before, the pre-intervention network is represented by matrix W and $W^{-j_0j_1}$ is the same as W except for the link j_0j_1 which is replaced by zero. Suppose that after the link j_0j_1 is replaced by zero, the network matrix changes to $G^{-j_0j_1}$. Denote the ij entry of matrix $G^{-j_0j_1}$ by g_{ij}. We assume that $G^{-j_0j_1}$ differs from $W^{-j_0j_1}$ for a set of links $S(j_0j_1)$:

$$G^{-j_0j_1} - W^{-j_0j_1} = \sum_{ij \in S(j_0j_1)} \xi_{ij} \ell_i \ell'_j, \tag{E.1}$$

where ξ_{ij} measures the network changes for the ij link.

Theorem 3. *Suppose that the network interactions are described by Eq.(1) and the network changes according to Eq.(E.1) when removing the network link $w_{j_0j_1}$, i.e., replacing $w_{j_0j_1}$ by 0. Also assume that $|\lambda| \max_i \sum_j |w_{ij}| < 1$ and $|\lambda| \max_i \sum_j |g_{ij}| < 1$. Removing the network link $w_{j_0j_1}$, will reduce the aggregate outcome by*

$$\begin{aligned} &\lambda w_{j_0j_1} [\left(I - \lambda W^{-j_0j_1\prime}\right)^{-1} \mathbf{1}]_{j_0} [(I - \lambda W)^{-1} \eta]_{j_1} \\ &\quad - \lambda \sum_{ij \in S(j_0j_1)} \xi_{ij} [\left(I - \lambda G^{-j_0j_1\prime}\right)^{-1} \mathbf{1}]_i [\left(I - \lambda W^{-j_0j_1}\right)^{-1} \eta]_j. \end{aligned} \tag{E.2}$$

The proof is similar to that of Theorem 2 and is omitted. Comparing Theorems 1 and 3, whether changes in other parts of the network as a result of removing a link magnify or dampen the direct effects of removing the link depends on the term in Eq.(E.2), specifically the sign and size of ξ_{ij} and the centralities of the affected routes. For example, in the context of population flows and disease spread, restrictions on a travel route may increase the flows to other destinations, which may attenuate the direct effects of restrictions in reducing aggregate outcomes. On the other hand, closing a travel route may discourage travel on other routes due to awareness of increased disease risks, which may enhance the direct effects of the route closure.

A more substantive treatment on changes in multiple network links for undirected networks is given by Sun *et al.* (2021). We have assumed that how the network rewires after interventions is deterministic and is known to the decision maker. Allowing for uncertainty requires a model for network formation such as Mele (2017) and Lee *et al.* (2020) and methods from statistical treatment assignment (Manski, 2004). One may also consider bounds on the indirect effects from network changes (Manski and Molinari, 2021).