

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

**Working Paper**

## Optimal Travel Restrictions in Epidemics

IZA Discussion Papers, No. 15290

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Shi, Wei; Qiu, Yun; Yu, Pei; Chen, Xi (2022) : Optimal Travel Restrictions in Epidemics, IZA Discussion Papers, No. 15290, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/263506>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

DISCUSSION PAPER SERIES

IZA DP No. 15290

**Optimal Travel Restrictions in Epidemics**

Wei Shi  
Yun Qiu  
Pei Yu  
Xi Chen

MAY 2022

## DISCUSSION PAPER SERIES

IZA DP No. 15290

# Optimal Travel Restrictions in Epidemics

**Wei Shi**

*Jinan University*

**Yun Qiu**

*Jinan University*

**Pei Yu**

*Rice University*

**Xi Chen**

*Yale School of Public Health, Yale University and IZA*

MAY 2022

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9  
53113 Bonn, Germany

Phone: +49-228-3894-0  
Email: [publications@iza.org](mailto:publications@iza.org)

[www.iza.org](http://www.iza.org)

## ABSTRACT

---

# Optimal Travel Restrictions in Epidemics\*

Travel restrictions are often imposed to limit the spread of infectious diseases. As uniform restrictions can be inefficient and incur unnecessarily high costs, this paper examines the optimal design of restrictions that target specific travel routes. We propose a model with trade-offs between costs of infections and costs of travel restrictions, where decisions are made with or without coordination between local jurisdictions and provide a computational feasible way to solve the optimization problem. We illustrate the model using the COVID-19 data in China. When travel restrictions target key routes, only around 5% of the possible routes need to be closed in order to have the same number of confirmed COVID-19 cases in the initial outbreaks. Uncoordinated travel restrictions ignore policy externalities and therefore are sub-optimal in comparison to coordinated restrictions.

**JEL Classification:** I18, R1, C21, C6

**Keywords:** COVID-19, transmission, public health, economic cost, coordination, externality

**Corresponding author:**

Xi Chen  
Department of Health Policy and Management  
Yale School of Public Health  
60 College St  
New Haven  
CT 06520  
USA  
E-mail: xi.chen@yale.edu

---

\* 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 NIH/NIA grant (K01AG053408). Shi thanks the Ministry of Education of China (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). Declarations of interest: none.

# 1 Introduction

The spread of communicable diseases, especially those that could transmit via airborne droplets, depends crucially on the degree 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](#)) and 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, such as social distancing, quarantine and isolation, travel restrictions, testing and contact tracing, to stringent measures, such as city lockdown, shelter-in-place, etc. While many of these public health measures are effective in suppressing the spread of COVID-19 (e.g., [Tian et al., 2020](#)), they could also bring significant social and economic costs and disruptions.

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, route-specific travel restrictions can still be implementable when complete lockdowns are not, for example, when the unit under consideration is a major metropolitan or an entire country. Even in the scenarios when route-specific travel restrictions are not possible, our identified optimal travel restrictions, once integrated with advanced mobile technology and specific public health measures<sup>1</sup>, 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 database and electronic health records to facilitate health-care facilities to identify high-risk patients for targeted screening, timely quarantine, and aggressive contact tracing ([Emanuel et al., 2020](#), [Wang et al., 2020a](#)). These data also guide border checks and surveillance ([Whitelaw et al., 2020](#)).

To characterize the optimal travel restrictions and offer numerical solutions for policy consideration, we embed a parsimonious model of the spatial spread of the disease in an optimization problem that considers the trade-off between the costs from disease infections and those from travel restriction measures. The optimization problem is solved from the perspective of a social planner that can coordinate policies and dictate the levels of restrictions on the population flows between each pairs of cities. We also solve the optimization problem under the constraint that each city acts independently to maximize its own payoff while taking the actions of other cities as given, i.e., a simultaneous move Nash equilibrium outcome. Due to the externality of policy effects, uncoordinated travel restrictions can be sub-optimal and we illustrate this in an empirical application.

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 top routes in the population flow networks most influential on the total number of COVID-19 cases in China and thus need to

---

<sup>1</sup>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.

be restricted the most, which consist of those closely connected to areas with severe infections and those whose destinations are cities with large population outflows. In comparisons to the scenario without coordination, total social welfare can be improved with often less stringent optimal restrictions imposed due to spillover effects when cities coordinate their travel restriction policies. 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 transmissions.

Our paper 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 an 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). Holtz *et al.* (2020) study coordinated and uncoordinated shelter-in-place orders. Fajgelbaum *et al.* (2020) examine optimal restrictions targeting directional commuting flows. They integrate 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 are related to disease spread and production. We identify the key parameters using a causal inference model, in comparison to this literature dominated by epidemiology models. Our model is more concise and applicable to a larger geographic scale, such as the spread of diseases across many cities or countries, because the number of choice variables is shown to be proportional to the number of units  $N$ , rather than  $N^2$ . We examine local decision making in both coordinated and uncoordinated scenarios.

Our paper is also related to policies on networks. 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. Ballester *et al.* (2006) characterize the node in a network whose removal has the largest impact on the aggregate outcome. We consider the marginal effect of varying the strength of a directed link between two nodes on the individual and aggregate outcomes, and show that this can be characterized as a product between a term similar to the Bonacich centrality of population flows of the origin and a term measuring the contagion risk that the destination poses to other locations through population flows. Therefore, an optimal travel restriction policy should take into account the risk of the origin and also the risk that the destination can bring to other cities. Our results can be extended to spatial interaction models where agents take the network as given (e.g., Lee, 2007) and a planner can alter the strength of network links.

The paper is organized as follows. Section 2 describes a model for the spread of communicable diseases across space and estimates of the model parameters based on the spread of COVID-19 in China are presented in Section 3. Section 4 analyzes the optimal travel restrictions from the perspective of a social planner that can coordinate between cities, and from the perspective of individual cities with no between-city coordination. Section 5 concludes.

## 2 The key link in network interactions

### 2.1 Model setup

There are  $\mathcal{N} = \{1, \dots, n\}$  units interacting through a network described by an  $n \times n$  matrix  $W = (w_{ij})$ . Let  $y_i$  denote the outcome of unit  $i$ . We assume the following model of network interactions,

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

$w_{ij} \geq 0$  and  $w_{ii} = 0$ .  $x_i$  is a vector of control variables which can include time lagged values of  $y_i$ .  $u_i$  is the error term. 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](#)).

Denote  $\eta_i = x_i' \beta + u_i$  and  $\eta = (\eta_1 \dots \eta_n)'$ . Assuming that the matrix  $I - \lambda W$  is invertible, Eq.(1) describes an equilibrium system of  $\{y_i\}_{i=1}^n$  for given  $\eta$  and  $W$ . The reduced form is

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

where  $\ell_i$  an  $n \times 1$  vector with  $i$ -th entry 1 and all other entries 0<sup>2</sup>. We first observe that the increase in the aggregate outcome for a shock to city  $j$  is  $\frac{\partial \sum_i y_i}{\partial \eta_j} = \ell_j' (I - \lambda W)^{-1} \mathbf{1}$ , where  $\mathbf{1}$  is a vector of ones, i.e., units with higher values of Bonacich centralities have larger effects on the aggregate outcome. 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

$$\frac{\partial y_i}{\partial w_{j_0 j_1}} = \lambda \ell_i' (I - \lambda W)^{-1} \ell_{j_0} \ell_{j_1}' (I - \lambda W)^{-1} \eta. \quad (3)$$

From Eq.(3), 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

$$\frac{\partial \sum_i y_i}{\partial w_{j_0 j_1}} = \lambda \mathbf{1}' (I - \lambda W)^{-1} \ell_{j_0} \ell_{j_1}' (I - \lambda W)^{-1} \eta = \lambda \ell_{j_0}' (I - \lambda W')^{-1} \mathbf{1} \ell_{j_1}' (I - \lambda W)^{-1} \eta. \quad (4)$$

In our empirical illustration,  $y_i$  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  $i$ . 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

---

<sup>2</sup>Note that the Bonacich centrality of unit  $i$  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  $\eta$ .

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.(4) 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 ( $\ell'_{j_0} (I - \lambda W')^{-1} \mathbf{1}$ ) and a term similar to the Bonacich centrality of the origin city ( $\ell'_{j_1} (I - \lambda W)^{-1} \eta$ ). 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 the destination city can affect many other cities.

## 2.2 The key network link

Eq.(4) 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.

**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  $w_{j_0 j_1}$ , i.e., replacing  $w_{j_0 j_1}$  by 0, will reduce the the aggregate outcome by  $[(I - \lambda W')^{-1} \mathbf{1}]_{j_0} w_{j_0 j_1} \lambda [(I - \lambda W)^{-1} \eta]_{j_1}$ .*

*Proof.* Let  $-j_0 j_1$  denote the scenario when the network link  $w_{j_0 j_1}$  is removed. Under the assumption that  $|\lambda| \max_i \sum_j |w_{ij}| < 1$ ,  $I - \lambda W$  and  $I - \lambda W^{-j_0 j_1}$  are invertible and the equilibrium outcome described by Eq.(2) exists and is unique. From Eq.(2),

$$\begin{aligned} y_s - y_s^{-j_0 j_1} &= \ell'_s (I - \lambda W)^{-1} \eta - \ell'_s (I - \lambda W^{-j_0 j_1})^{-1} \eta \\ &= \sum_{t=1}^n \sum_{p=1}^{\infty} \lambda^p \left( w_{st}^{[p]} - w_{st(-j_0 j_1)}^{[p]} \right) \eta_t \\ &= \lambda I(s = j_0) w_{j_0 j_1} \eta_{j_1} + \sum_{t=1}^n \sum_{p=2}^{\infty} \lambda^p I(s = j_0) w_{j_0 j_1} w_{j_1 t}^{[p-1]} \eta_t + \sum_{p=2}^{\infty} \lambda^p w_{s j_0}^{[p-1]} w_{j_0 j_1} \eta_{j_1} \\ &\quad + \sum_{t=1}^n \sum_{p=3}^{\infty} \lambda^p \sum_{\substack{a+b=p-1 \\ a \geq 1, b \geq 1}} w_{s j_0}^{[a]} w_{j_0 j_1} w_{j_1 t}^{[b]} \eta_t, \end{aligned} \tag{5}$$

where  $w_{st}^{[p]}$  is the weight of all paths of length  $p$  from  $s$  to  $t$ :  $w_{st}^{[0]} = 1$ ,  $w_{st}^{[1]} = w_{st}$ ,  $w_{st}^{[p]} = \sum_{k_1=1}^n \cdots \sum_{k_{p-1}=1}^n w_{s k_1} w_{k_1 k_2} \cdots w_{k_{p-1} t}$  for  $p \geq 2$ .  $w_{st(-j_0 j_1)}^{[p]}$  is the weight of all  $p$ -length paths from  $s$  to  $t$  that does not contain  $w_{j_0 j_1}$ .  $I()$  is the indicator function.

Summing over  $s$ , the impact of removing  $w_{j_0 j_1}$  on the aggregate outcome is

$$\sum_{s=1}^n (y_s - y_s^{-j_0 j_1}) = \lambda \left( 1 + \sum_{s=1}^n \lambda w_{s j_0}^{[1]} + \sum_{s=1}^n \lambda^2 w_{s j_0}^{[2]} + \cdots \right) w_{j_0 j_1}$$



$$\begin{aligned} & \left( \eta_{j_1} + \sum_{t=1}^n \lambda w_{j_1 t}^{[1]} \eta_t + \sum_{t=1}^n \lambda^2 w_{j_1 t}^{[2]} \eta_t + \dots \right) \\ &= [(I - \lambda W')^{-1} \mathbf{1}]_{j_0} w_{j_0 j_1} \lambda [(I - \lambda W)^{-1} \eta]_{j_1}. \end{aligned}$$

□

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 ( $\lambda w_{j_0 j_1}$ ), the centrality of the origin unit ( $[(I - \lambda W)^{-1} \eta]_{j_1}$ ), and the centrality of the destination unit ( $[(I - \lambda W')^{-1} \mathbf{1}]_{j_0}$ ).

### 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 econometric (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). Ballester and Zenou (2014) extend the key player analysis of Ballester *et al.* (2006) to allow for contextual effects and show that the identification of the unit whose removal results in the largest reduction in the aggregate outcome can be different when contextual effects are taken into account. This section extends the key network link analysis to allow for contextual effects.

The network interactions model is

$$y_i = \left( \sum_{j=1}^n w_{ij} y_j \right) \lambda + \eta_i, \quad \eta_i = x'_i \beta + \left( \sum_{j=1}^n w_{ij} x'_j \right) \beta_w + u_i. \quad (6)$$

Let  $\eta^{-j_0 j_1}$  denote the  $n \times 1$  vector  $\eta$  with the  $w_{j_0 j_1}$  term replaced by zero, i.e.,  $\eta^{-j_0 j_1} = \eta - \ell_{j_0} w_{j_0 j_1} x'_{j_1} \beta_w$ .

**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  $[(I - \lambda W')^{-1} \mathbf{1}]_{j_0} w_{j_0 j_1} \left( \lambda [(I - \lambda W)^{-1} \eta^{-j_0 j_1}]_{j_1} + x'_{j_1} \beta_w \right)$ .

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

$$\begin{aligned} y_s - y_s^{-j_0 j_1} &= \ell'_s (I - \lambda W)^{-1} \eta - \ell'_s (I - \lambda W^{-j_0 j_1})^{-1} \eta^{-j_0 j_1} \\ &= \ell'_s (I - \lambda W)^{-1} (\eta - \eta^{-j_0 j_1}) + \ell'_s (I - \lambda W)^{-1} \eta^{-j_0 j_1} - \ell'_s (I - \lambda W^{-j_0 j_1})^{-1} \eta^{-j_0 j_1} \\ &= \ell'_s (I - \lambda W)^{-1} \ell_{j_0} w_{j_0 j_1} x'_{j_1} \beta_w + \ell'_s (I - \lambda W)^{-1} \eta^{-j_0 j_1} - \ell'_s (I - \lambda W^{-j_0 j_1})^{-1} \eta^{-j_0 j_1}. \end{aligned}$$

The proof is completed by using the result in Eq.(5) and summing over  $s$ . □

Table 1: Summary of Empirical Model Specifications

	Model A	Model B
$y_i$	$\log(1 + \# \text{ of confirmed COVID-19 cases by February 29})$	
$z_i$	local weather variables that affect infection rates	
$w_{ij}, t_i$	average pop. flows, Jan 1-Feb 29	average pop. flows, Jan 1-Jan 22
$\tilde{w}_{ij}, \tilde{t}_i$	average pop. flows, 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	$\tilde{t}_i, \sum_{j=1}^n \tilde{w}_{ij}z_j$	$\sum_{j=1}^n w_{ij}z_j$

When contextual effects are present, the additional term in the effect of closing a network link on the aggregate outcome is due to spillovers from the origin  $j_1$ 's characteristics ( $w_{j_0j_1}x'_{j_1}\beta_w$ ) to the destination  $j_0$  and further spillover effects from  $j_0$ .

### 3 Estimates of Model Parameters

In this section, we estimate the model parameters in Eq.(1) using data on confirmed COVID-19 cases and the intensities of between and within city population flows in China. The estimated coefficient of between city spread ( $\lambda$ ) will be the basis for the model of optimal travel restrictions in Section 4.

We consider two empirical specifications (Table 1). 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, the city of Wuhan was placed under lockdown and travelling out of Wuhan was mostly stopped. Since then, people's travel decisions were likely affected by either the perceived risk of infection or the public health measures imposed by the government, which in turn were 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 as the instrumental variables. 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 is less likely to be affected by the severity of COVID-19 transmission before January 23. 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 variable in other cities following Qiu *et al.* (2020). We control weather conditions in the own city, including precipitation, the interaction between precipitation and wind speed, 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 ( $\sum_{j=1}^n w_{ij}y_j$ ).

### 3.1 Data

We collect the numbers of cumulative confirmed COVID-19 cases of 360 prefecture level cities by February 29, 2020 from 32 provincial-level Health Commissions in China. The National Oceanic and Atmospheric Administration (NOAA) provides precipitation, visibility, wind speed, an indicator for bad weather (fog, rain or drizzle, snow or ice pellets, hail, thunder, tornado or funnel cloud), 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 calculate daily weather variables for each city from station-level weather records following the inverse distance weighting method. Specifically, for each city, we draw a circle of 100 km from the city's centroid and calculate the weighted average daily weather variables using stations within the 100-km circle. We use the inverse of the distance between the city's centroid and each station as the weight. Second, we calculate the average weather characteristics of each city for each specification, which are then matched with the number of COVID-19 cases based on the city identifier.

We obtain the data on population movement between and within cities from Baidu Migration<sup>3</sup>, which is based on mobile phone location data to track population flow. From the Baidu Migration data, we collect the daily inflow index and outflow index for 369 cities between January 1 and February 29 in 2020, and on the same lunar calendar days in 2019. For each of the 369 cities, Baidu Migration also records the shares of the top 100 origination 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 assume that the population flow is zero for destination or origin cities outside the top 100 lists. Then the between city population flow intensities are calculated by multiplying the daily migration index of the population flows with the share of the flows<sup>4</sup>. Regarding the within city population flow intensities, Baidu also provides the daily within city migration index for January 1 - February 29 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.

### 3.2 Estimation Results

In Table 3, columns (1) and (3) report the OLS estimates and columns (2) and (4) display the IV estimates from Eq.(1). Columns (3) and (4) include province fixed effects, while columns (1) and (2) do not. All columns in Table 3 show a significantly positive spillover effect of infections in other cities mediated by population flows. The IV estimate  $\lambda = 0.258$  (column (2)) implies that a 1% increase in infections in a city where 100,000 people travel to the focal city causes a 0.284% increase in the number of cases in the focal city<sup>5</sup>. The magnitudes of OLS and IV estimates on  $\lambda$  are very

---

<sup>3</sup><http://qianxi.baidu.com/>

<sup>4</sup>In the event of a slight gap between the population flow intensity calculated by the inflow index of destination cities and that by the outflow index of origination cities for a city pair, we take average of the two intensities.

<sup>5</sup>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 0.284% ( $0.258 \times 100000 \div 90848 \times 1\%$ ) for a 1% increase in the infections in the origin city from where 100,000 people travel to the destination city on average daily.

Table 2: Summary Statistics

Variables	N	Mean	SD	Min	Median	Max
<b>Cities with confirmed cases</b>						
<i>Jan1-Feb29, 2020</i>						
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 precipitation,mm	324	0.218	0.415	0	0.122	4.528
Average wind speed,m/s	324	2.278	0.761	0.958	2.163	5.050
Bad weather	324	0.393	0.185	0	0.381	0.805
<i>Jan1-Jan22, 2020</i>						
Average within city population flows	324	5.230	0.562	3.096	5.322	6.553
Average precipitation,mm	324	0.169	0.369	0	0.065	4.233
Average wind speed,m/s	324	2.122	0.795	0.732	1.966	6.077
Bad weather	324	0.399	0.223	0	0.364	0.909
<b>Cities without confirmed cases</b>						
<i>Jan1-Feb29, 2020</i>						
Average within city population flows	36	3.105	0.788	1.761	3.027	4.776
Average precipitation,mm	36	0.385	1.023	0	0.034	3.961
Average wind speed,m/s	36	1.966	0.752	0.887	1.806	3.745
Bad weather	36	0.187	0.164	0	0.159	0.583
<i>Jan1-Jan22, 2020</i>						
Average within city population flows	36	4.138	1.039	1.968	4.264	6.366
Average precipitation,mm	36	0.395	1.074	0	0.016	4.365
Average wind speed,m/s	36	1.887	0.812	0.797	1.677	3.680
Bad weather	36	0.232	0.181	0	0.224	0.818

Please see the text for variable sources.

Table 3: Model A: Estimation Results

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
$\lambda$	0.252*** (0.0406) [0.277]	0.258*** (0.0269) [0.284]	0.240*** (0.0373) [0.264]	0.245*** (0.0329) [0.270]
$\gamma$	-0.365 (0.365)	0.294 (0.209)	-0.387** (0.143)	-0.144 (0.136)
Precipitation	-0.987*** (0.253)	-0.614*** (0.207)	-0.249*** (0.0731)	-0.237*** (0.0660)
Precipitation $\times$ wind speed	0.329*** (0.0709)	0.245*** (0.0702)	0.0594* (0.0346)	0.0547* (0.0322)
Bad weather	3.434*** (1.131)	2.959** (1.175)	0.562 (0.592)	0.619 (0.605)
Observations	360	360	360	360
Province FE	NO	NO	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. Control variables are precipitation, the interaction between precipitation and wind speed, and a dummy for bad weather in own cities. The set of these 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. In Column (3)-(4), province fixed effects are included. Elasticity of infection spillovers per 100000 daily population movements are reported in brackets. Standard errors in parentheses are clustered by provinces. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

similar. Population movements between cities lead to the spread of the virus, which can be reduced by travel restrictions. The IV estimated coefficient on the within city population flow intensity is positive in column (2) and negative in column (4), and both are insignificant. Notice that the OLS estimate of  $\gamma$  is negative, which can be ascribed to the issue of reverse causality because people may avoid going outside when the risk of catching the virus is high. After controlling for province fixed effects, the estimated spillover effects remain stable and significant although the magnitudes are somewhat smaller.

Table 4 presents the first-stage estimates for the IV regressions 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. We also report the R-squared and  $F$ -test statistics for the joint significance of excluded instruments in the first stage.

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-

Table 4: Model A: First Stage Results

VARIABLES	(1) W <sub>y</sub>	(2) t	(3) W <sub>y</sub>	(4) t
Own city				
Within-city population flow intensities in 2019	0.184 (0.136)	0.733*** (0.0960)	0.255* (0.143)	0.752*** (0.138)
Precipitation	0.102 (0.151)	-0.293*** (0.0472)	0.0649 (0.0462)	-0.0817*** (0.0174)
Precipitation $\times$ wind speed	-0.0750 (0.0546)	0.0581*** (0.0195)	-0.0436 (0.0277)	0.0337*** (0.00733)
Bad weather	-0.362 (0.377)	0.219 (0.311)	-0.283 (0.451)	-0.108 (0.120)
Other cities, weight = population flow				
Precipitation	-4.042** (1.645)	-0.981* (0.543)	-2.342** (1.083)	-0.145 (0.283)
Precipitation $\times$ wind speed	1.516** (0.631)	0.215 (0.184)	1.006** (0.379)	0.0929 (0.0934)
Bad weather	4.649*** (0.409)	0.0523 (0.131)	4.424*** (0.261)	-0.253*** (0.0698)
First-stage $R^2$	0.937	0.587	0.969	0.840
$F$ -test of excluded instruments	138.7	30.30	238.3	78.31
$F$ -test $p$ -value	0.000	0.000	0.000	0.000
Observations	360	360	360	360
Province FE	NO	NO	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. Standard errors in parentheses are clustered by provinces. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 5: Model B: Estimation Results

	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
$\lambda$	0.179*** (0.0165) [0.197]	0.163*** (0.0126) [0.179]	0.170*** (0.0136) [0.187]	0.170*** (0.0125) [0.187]
$\gamma$	0.288* (0.148)	0.294** (0.141)	0.131 (0.164)	0.131 (0.153)
Precipitation	-0.606*** (0.140)	-0.638*** (0.152)	-0.229*** (0.0731)	-0.228*** (0.0685)
Precipitation $\times$ wind speed	0.234*** (0.0568)	0.242*** (0.0610)	0.0513 (0.0360)	0.0513 (0.0337)
Bad weather	2.387** (0.973)	2.531** (1.052)	0.439 (0.752)	0.436 (0.704)
Observations	360	360	360	360
Province FE	NO	NO	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 between 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. Control variables are precipitation, the interaction between precipitation and wind speed, a dummy for bad weather in own cities. The sum of these variables in other cities weighted by the the population flow intensities between cities in 2020 are used as instrumental variables in the IV regressions. In Column (3)-(4), province fixed effects are included. Elasticity of infection spillovers per 100000 daily population movements are reported in brackets. Standard errors in parentheses are clustered by provinces. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

scale public health measures seed COVID-19 across space. Results are reported in Table 5 and are similar to the baseline results in Table 3 except for the OLS estimates of  $\gamma$  in column (1), which is significantly positive. This is expected, because the within city population flows before the lockdown of Wuhan are not likely to be affected by the unobservables that drive the infection dynamics, hence smaller endogeneity biases. The IV estimate of  $\gamma$  is positive in column (4), and is positive and significant in column (2). With more within city population movement, people may be more likely to be in contact with the infected in the same city. The estimated cross city spillover effects are smaller than those in Table 3, which is possibly because that the 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 6. If population flows are exogenous, the OLS estimate of  $\lambda$  is overestimated due to positive spatial spillovers between cities, which can be seen from Table 5. It is also of interest to note that in Model A, people's travel intensities after the virus outbreak were likely negatively correlated with the severity of local infections, which could lead to a negative bias of the OLS estimate, and the direction of the bias is therefore ambiguous for the OLS estimates of Model A (Table 3).

Table 6: Model B: First Stage Results

VARIABLES	(1) Wy	(2) Wy
Own city		
Within-city population flow intensities in 2020	-0.182 (0.213)	-0.169 (0.205)
Precipitation	0.198 (0.189)	0.0537 (0.106)
Precipitation $\times$ wind speed	-0.183** (0.0894)	-0.0739 (0.0636)
Bad weather	-1.932** (0.918)	-0.605 (0.790)
Other cities, weight = population flow		
Precipitation	-8.266*** (2.963)	-5.542*** (1.892)
Precipitation $\times$ wind speed	3.822*** (1.267)	3.011** (1.126)
Bad weather	8.068*** (0.809)	7.714*** (0.794)
First-stage $R^2$	0.931	0.968
$F$ -test of excluded instruments	219.3	171.3
$F$ -test $p$ -value	0.000	0.000
Observations	360	360
Province FE	NO	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. Standard errors in parentheses are clustered by provinces. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



## 4 Optimal Travel Restrictions

Travel restrictions such as city lockdowns and suspensions of intercity transportation significantly disrupted population flows between cities in China since the outbreak of COVID-19 in January. While studies have shown that public health interventions such as city lockdowns and shelter-in-place orders have reduced the number of infections (Fang *et al.*, 2020, Tian *et al.*, 2020, Qiu *et al.*, 2020), it remains an open question whether the set of policies adopted can be improved, especially in light of the large number of available policy tools and the socioeconomic costs that these policies inevitably incur. In Section 2, we characterize travel routes on which the population flows contribute the most to the spread of viral diseases. The presence of policy effect externality calls for coordination between local jurisdictions in the adoption of travel restrictions, and aggregate social welfare could be lower if local decisions do not account for their spillover effects (Holtz *et al.*, 2020). In this section, we propose a model that considers the trade-off between costs from viral infections and those from the public health measures, and solve for the set of optimal travel restrictions when cities can coordinate (4.1) or act independently (4.2). The theoretical model is then applied to the data on COVID-19 infections in China (4.3).

### 4.1 Social Planner’s Problem

Suppose that the social planner can impose restrictions on between city travels so that  $w_{ij}$  can be reduced to  $\phi_{ij}w_{ij}$  with  $0 \leq \phi_{ij} \leq 1$ . The restrictions we discuss in this paper aim to limit the spread of infections across cities that arise from the interactions between infected and susceptible people. There are many policy instruments that can achieve this goal, such as lockdowns which stop intercity travels (Qiu *et al.*, 2020) or imposing proactive testing and quarantine on people with certain travel histories after their arrival (Wang *et al.*, 2020b), and the feasibility of the policy tools can vary in different places. In the following, we refer to the policies simply as travel restrictions.

The marginal cost of such restrictions is assumed to be increasing with the amount of travel that could have happened without the restrictions but did not happen with the restrictions imposed. The cost of the restriction on route  $w_{ij}$  is modeled by  $\left(\phi_{ij}\varphi_1 + \frac{1}{2}\phi_{ij}^2\varphi_2\right)w_{ij}$ , with parameters  $\varphi_1 < 0$  and  $\varphi_2 > 0$ . Summarize travel restrictions by matrix  $\Phi$ ,  $(\Phi)_{ij} = \phi_{ij}$  and the diagonal elements of  $\Phi$  are restricted to zeros<sup>6</sup>. The social planner applies weight  $g_i$  to city  $i$  and  $g = (g_1 \ \cdots \ g_n)'$ .

The planner’s optimization problem is to minimize the cost from COVID-19 infections and the travel restrictions imposed,

$$\min_{0 \leq \phi_{ij} \leq 1, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j} \sum_i g_i y_i + \sum_{i,j,i \neq j} \left(\phi_{ij}\varphi_1 + \frac{1}{2}\phi_{ij}^2\varphi_2\right)w_{ij}, \quad (7)$$

---

<sup>6</sup>This is without loss of generality. In the first order conditions of the optimization problem, terms such as  $\Phi \odot W$  are not affected by this restriction, because the diagonal entries of  $W$  are zeros.

subject to

$$y_i = \left( \sum_{j=1}^n \phi_{ij} w_{ij} y_j \right) \lambda + \eta_i, \quad \text{for all } i. \quad (8)$$

**Remark 1.** The welfare weights  $g_i$  in (7) are determined by the social planner and could accommodate heterogeneous costs of infections across space. As one example,  $g_i = \frac{1}{\text{population size}_i}$  if the cost of infections is proportional to the number of cases in logs per capita. The special case that  $g_i = 1$  for all  $i$  is also included.

**Remark 2.** The planner's problem is a static one-period model. We have assumed that the planner's information set is  $\sigma(W, \{\eta_i\}_{i=1}^n, \lambda)$  where  $\sigma()$  denotes the  $\sigma$ -field generated by the argument inside. In this parsimonious model, spatial distribution of infections is determined by the initial viral spread ( $\eta$ ) and the exogenous population flows between cities which are the targets of travel restrictions.

In reality,  $\lambda$  can be estimated empirically as in Section 3 or based on prior studies on similar infectious diseases. For example, when  $\lambda w_{ij} y_j$  is specified as  $\lambda \frac{\text{pop flow}_{j \rightarrow i}}{\text{pop}_j} y_j$ ,  $\lambda$  can be calculated based on the basic reproduction number in an SIR model of the disease (Kermack and McKendrick, 1927). In addition, the planner can adjust its decisions as new information become available with the passage of time, and the future evolution of the variables may depend on the decisions made in the past. Jeong and Lee (2020) consider a dynamic optimization problem where forward looking agents interact with others over time on a fixed network. It is an area for future research that the network is among the choice variables as well in the settings of Jeong and Lee (2020). We have assumed fixed population flows, but people can adjust their behavior in response to factors such as risks of infections, government policies, etc., as has been shown in many studies such as Wong (2008) and Fang et al. (2020). Fajgelbaum et al. (2020) endogenize the costs of lockdowns by modeling individuals' commuting decisions, which provide labor supply to firms, generate incomes, and also spread diseases. While the main objective of our model is to show what types of routes should be restricted the first and how this relates to the network structure, these are substantive future research topics.

**Remark 3.** Our model can accommodate uncertainties in the initial viral spread. Suppose that  $W$  and the distribution of  $\eta$  are known to the decision maker. The planner's problem is

$$\min_{0 \leq \phi_{ij} \leq 1, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j} \mathbb{E} \sum_i g_i y_i + \sum_{i,j, i \neq j} (\phi_{ij} \varphi_1 + \frac{1}{2} \phi_{ij}^2 \varphi_2) w_{ij}, \quad (9)$$

subject to (8). Notice that  $\sum_i g_i y_i$  is linear in  $\eta$ ,  $\sum_i g_i y_i = g'(I - \lambda \Phi \odot W)^{-1} \eta$ . Let  $\bar{\eta} = \mathbb{E} \eta$ . We can rewrite Eq.(9) as

$$\min_{0 \leq \phi_{ij} \leq 1, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j} \sum_i g_i \bar{y}_i + \sum_{i,j, i \neq j} (\phi_{ij} \varphi_1 + \frac{1}{2} \phi_{ij}^2 \varphi_2) w_{ij},$$

subject to

$$\bar{y}_i = \left( \sum_{j=1}^n \phi_{ij} w_{ij} \bar{y}_j \right) \lambda + \bar{\eta}_i, \quad \text{for all } i.$$

In the empirical illustration in Section (4.3), we consider two specifications for the expectations of the initial viral spread. In the first scenario, we assume that  $\bar{\eta}_{Wuhan} = \log(10000)$ ,  $\bar{\eta}_i = 0$  for  $i \neq Wuhan$ . In the second scenario, we construct  $\bar{\eta}$  from the estimated parameters and the residual,  $\bar{\eta}_i = t_i \hat{\gamma} + x_i' \hat{\beta} + \hat{u}_i$ .

By the extreme value theorem, a continuous function attains a minimum on a compact set, and therefore the optimization problem in (7) always has a solution. The first order condition of (7) with respect to  $\phi_{ij}$  for  $w_{ij} \neq 0$  and  $i \neq j$  is

$$\lambda w_{ij} g' (I - \lambda \Phi \odot W)^{-1} \ell_i \ell_j' (I - \lambda \Phi \odot W)^{-1} \eta + (\varphi_1 + \phi_{ij} \varphi_2) w_{ij} = 0$$

for interior  $0 < \phi_{ij} < 1$ . For the boundary solutions,

$$\begin{aligned} \lambda w_{ij} g' (I - \lambda \Phi \odot W)^{-1} \ell_i \ell_j' (I - \lambda \Phi \odot W)^{-1} \eta + (\varphi_1 + \varphi_2) w_{ij} &\leq 0, & \text{if } \phi_{ij} = 1, \\ \lambda w_{ij} g' (I - \lambda \Phi \odot W)^{-1} \ell_i \ell_j' (I - \lambda \Phi \odot W)^{-1} \eta + \varphi_1 w_{ij} &\geq 0, & \text{if } \phi_{ij} = 0. \end{aligned}$$

Let  $\Gamma$  be a function from an  $n \times n$  matrix to an  $n \times n$  matrix. Define the following functions,

$$\begin{aligned} f(x) &= \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < 1, \\ 1 & \text{if } x \geq 1 \end{cases} \\ (\Gamma(\Phi))_{ij} &= \begin{cases} f \left( -\varphi_2^{-1} \left( \lambda g' (I - \lambda \Phi \odot W)^{-1} \ell_i \ell_j' (I - \lambda \Phi \odot W)^{-1} \eta + \varphi_1 \right) \right) & \text{if } i \neq j, w_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \quad (10)$$

The first order necessary conditions above can be summarized by  $\Gamma(\Phi) = \Phi$  for  $\Phi \in [0, 1]^{n \times n}$ . Notice that  $\Gamma(\cdot)$  is a continuous function from a convex and compact set to itself. By Brouwer's fixed-point theorem,  $\Gamma(\Phi) = \Phi$  has a solution. Therefore the optimal solution to (7) can be obtained by solving  $\Gamma(\Phi) = \Phi$ .

Taking the network structure of population flows into account, an optimal policy on between city mobility not only considers the severity of the outbreak in the immediate origin city, but also the outbreak intensity in cities that are connected to the origin city and the risk that infections in the destination city poses to other connected cities. From Eq.(10), assuming positive between city spread ( $\lambda > 0$ ), the travel from city  $j$  to  $i$  will be more severely restricted (i.e., smaller  $\phi_{ij}$ ) if the infection risk is high for travelers from city  $j$  ( $\ell_j' (I - \lambda \Phi \odot W)^{-1} \eta$ ), or city  $i$  is influencing many

other cities  $((I - \lambda\Phi \odot W)^{-1} \ell_i)$ .

$\Phi$  contains  $n^2 - n$  unknown parameters and directly solving for them can be infeasible. However, we note that Eq.(10) imposes restrictions on  $\Phi$  which can greatly reduce the number of free parameters, as we can write

$$\Phi = \Gamma(\Phi) = \Lambda \left( -\varphi_2^{-1} \lambda (I - \lambda(\Phi \odot W)')^{-1} g\eta' (I - \lambda(\Phi \odot W)')^{-1} - \frac{\varphi_1}{\varphi_2} \mathbf{1}_{n \times n} \right), \quad (11)$$

where  $\Lambda()$  is a function from  $\mathbb{R}^{n \times n} \rightarrow [0, 1]^{n \times n}$  with  $(\Lambda(M))_{ij} = f(M_{ij})$  if  $i \neq j, w_{ij} \neq 0$  and  $(\Lambda(M))_{ij} = 0$  otherwise. Because the rank of matrix  $(I - \lambda(\Phi \odot W)')^{-1} g\eta' (I - \lambda(\Phi \odot W)')^{-1}$  is only 1, we can parameterize  $\Phi$  as

$$\Phi = \Lambda \left( ab' - \frac{\varphi_1}{\varphi_2} \mathbf{1}_{n \times n} \right), \quad (12)$$

with unknown parameters  $a = \begin{pmatrix} 1 & a_2 & \dots & a_n \end{pmatrix}'$  and  $b = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}'$ <sup>7</sup>. The number of unknown parameters is now  $2n - 1$ , which can be much smaller than  $n^2 - n$ .

**Remark 4. Heterogeneities.** *The extent that cross city spread of the disease is mediated by population flows may depend on the nature of interactions between travellers and the local population, and may be heterogeneous across city pairs. This can be modelled by assuming a route specific nonzero spillover coefficient  $\lambda_{ij}$ , and rewrite Eq.(8) as  $y_i = \sum_{j=1}^n \phi_{ij} \tilde{w}_{ij} y_j + \eta_i$ , with  $\tilde{w}_{ij} = \lambda_{ij} w_{ij}$ . The equation that characterizes the optimal solutions is*

$$\Phi = \Lambda \left( -\varphi_2^{-1} L \odot \left( (I - (\Phi \odot \tilde{W})')^{-1} g\eta' (I - (\Phi \odot \tilde{W})')^{-1} \right) - \frac{\varphi_1}{\varphi_2} \mathbf{1}_{n \times n} \right), \quad (13)$$

with  $(L)_{ij} = \lambda_{ij}$  and  $(\tilde{W})_{ij} = \tilde{w}_{ij}$ . If the cost of travel restrictions is not linear in the population flow intensity, such as when routes have different compositions of business and leisure travels, we can replace the term by  $\left( \phi_{ij} \varphi_1 + \frac{1}{2} \phi_{ij}^2 \varphi_2 \right) \theta_{ij}$  for some route specific cost measure  $\theta_{ij}$ .

In both cases, the rest of the analysis can be modified easily, with the exception that we may not be able to parameterize the set of optimal policies using  $2n - 1$  parameters, because the matrix as an argument to  $\Lambda()$  in Eq.(11) may not be rank 1 due to heterogeneities in  $\lambda_{ij}$  or  $\theta_{ij}$ , as Eq.(13) shows. We opt for the simpler specification because it is feasible to solve for the optimality numerically when  $n$  is large.

**Remark 5. Risk Aversion.** *The planner's objective function can be nonlinear in  $y_i$ , which can allow for risk aversion when there is uncertainty in  $\eta$ . Let  $u()$  be a concave function. The planner's objective function is now*

$$\min_{0 \leq \phi_{ij} \leq 1, 1 \leq i \leq n, 1 \leq j \leq n, i \neq j} \mathbb{E} \sum_i g_i u(y_i) + \sum_{i,j,i \neq j} \left( \phi_{ij} \varphi_1 + \frac{1}{2} \phi_{ij}^2 \varphi_2 \right) w_{ij},$$

---

<sup>7</sup> $a_1$  is normalized to 1, because  $ab' = (ra)(r^{-1}b')$  for  $r \neq 0$ .

subject to (8). Taking derivatives with respect to  $\phi_{ij}$ , the optimal solution can be characterized by

$$\Phi = \mathbb{E}\Lambda \left( -\varphi_2^{-1} \lambda \left( (I - \lambda(\Phi \odot W)')^{-1} g(y) \eta' (I - \lambda(\Phi \odot W)')^{-1} \right) - \frac{\varphi_1}{\varphi_2} \mathbf{1}_{n \times n} \right),$$

where  $g(y) = \begin{pmatrix} g_1 u'(y_1) & \cdots & g_n u'(y_n) \end{pmatrix}'$ . Note when there is no uncertainty,  $\Phi$  can still be parameterized using (12) which allows for easier computation.

## 4.2 Travel Restrictions without Coordination

Restricting population flows has both benefits and costs. The costs from lost economic activities may be mostly borne by the cities that impose restrictions, at least in the short run. On the other hand, the benefits from reduced contagion may be enjoyed by other cities. If each city is making decisions on travel restrictions independently with no coordination between cities, this becomes a classic example of positive externalities and the level of travel restrictions may be suboptimal. In the case of compliance with social distancing orders in the United States, [Holtz et al. \(2020\)](#) show that social distancing behaviors in a state can be influenced by people in other states that are geographically or socially close, and because of the presence of spillovers, uncoordinated resumption of economic activities is not optimal. Similar results also hold in our context. To see this, assume that each city can choose the level of population inflows to the city. The optimization problem of city  $i$  is to choose  $\{\phi_{ij}\}_{j=1, j \neq i}^n$  while taking the choices of other cities as given,

$$\min_{0 \leq \phi_{ij} \leq 1, 1 \leq j \leq n, j \neq i} g_i y_i + \sum_{j, i \neq j} (\phi_{ij} \varphi_1 + \frac{1}{2} \phi_{ij}^2 \varphi_2) w_{ij}, \quad (14)$$

subject to

$$y_i = \left( \sum_{j=1}^n \phi_{ij} w_{ij} y_j \right) \lambda + \eta_i.$$

The necessary conditions for individual optimalities can be summarized by  $\tilde{\Gamma}(\tilde{\Phi}) = \tilde{\Phi}$  for  $\tilde{\Phi} \in [0, 1]^{n \times n}$ . The function  $\tilde{\Gamma}()$  is defined as

$$\left( \tilde{\Gamma}(\tilde{\Phi}) \right)_{ij} = \begin{cases} f \left( -\varphi_2^{-1} \left( \lambda g_i \ell'_i \left( I - \lambda \tilde{\Phi} \odot W \right)^{-1} \ell_i \ell'_j \left( I - \lambda \tilde{\Phi} \odot W \right)^{-1} \eta + \varphi_1 \right) \right) & \text{if } i \neq j, w_{ij} \neq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

where  $f()$  is the same as in Eq.(10).

Comparing (15) with the coordinated travel restriction policies (10), we can see that without coordination between cities, a city only takes into account the effect of travel restrictions on its own, which is less than the total effect because of external benefits to other cities, and this results

in travel restrictions that may not be optimal. Note that theoretically the set of travel restrictions with coordination may not be more stringent than the restrictions without coordination across all city pairs. With coordination, it may be optimal that with certain origin and destination pairs strictly controlled, travels on other routes can be restricted less.

To solve for the optimal  $\tilde{\Phi}$ , as in the previous subsection, we observe that

$$\tilde{\Phi} = \tilde{\Gamma}(\tilde{\Phi}) = \Lambda \left( -\varphi_2^{-1} \lambda \begin{pmatrix} g_1 \ell'_1 \left( I - \lambda \tilde{\Phi} \odot W \right)^{-1} \ell_1 \\ \vdots \\ g_n \ell'_n \left( I - \lambda \tilde{\Phi} \odot W \right)^{-1} \ell_n \end{pmatrix} \eta' \left( I - \lambda (\tilde{\Phi} \odot W)' \right)^{-1} - \frac{\varphi_1}{\varphi_2} \mathbf{1}_{n \times n} \right), \quad (16)$$

where function  $\Lambda()$  is the same as in Eq.(11) and the term involving  $\tilde{\Phi}$  on the right hand side is an  $n \times n$  matrix of rank 1. Therefore  $\tilde{\Phi}$  can be parameterized using Eq.(12) and Eq.(16) can be solved for these  $2n - 1$  unknowns.

### 4.3 Empirical Illustration

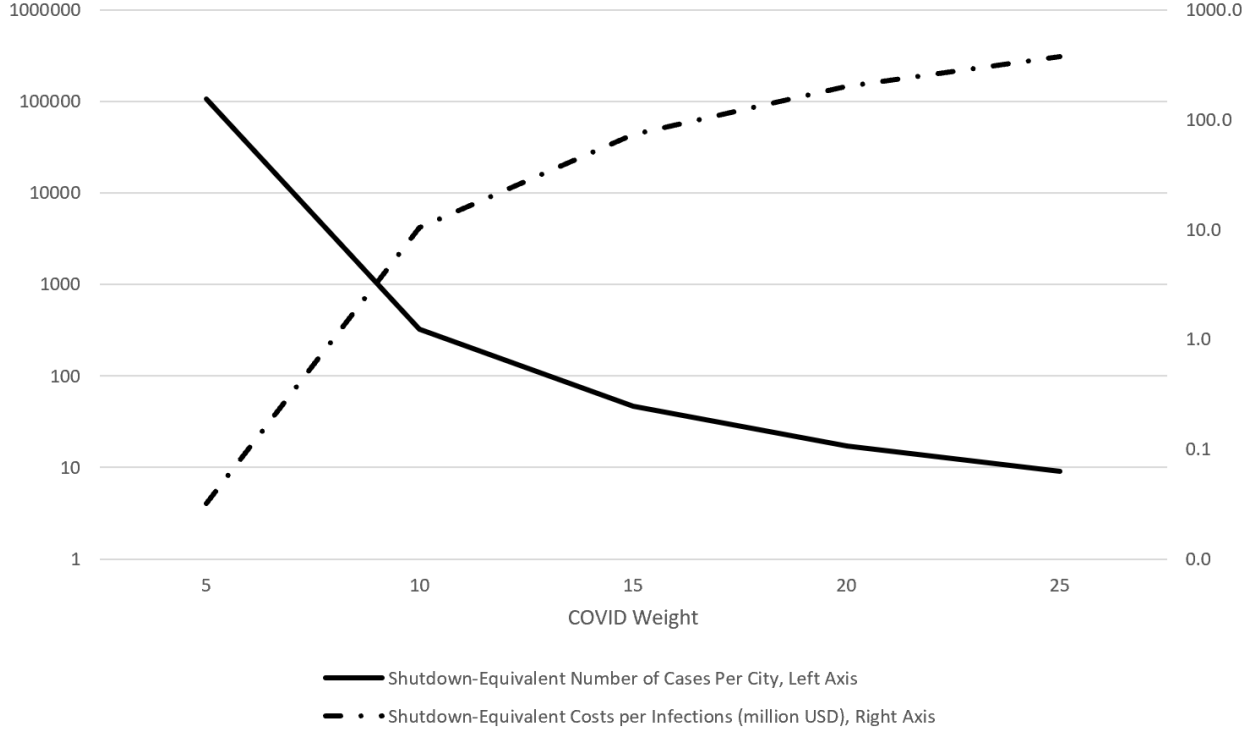
Section 2 shows that the marginal effects of adjusting the population flow intensity on a route on the total number of infections are heterogeneous, and depend on the product between the Bonacich centralities of the origin and the destination. Suppose that travels on routes whose marginal effects are larger are stopped first and we can then calculate the number of infections in the counterfactual scenario using Eq.(2). A few additional assumptions are needed. As a proxy of population flow patterns without any government intervention, the population flow matrix  $W$  is constructed from the averages of intercity population flows between January 1, 2020 and January 22, 2020 before the implementation of major public health measures such as the lockdown of Wuhan. For the coefficient of the spatial spillover effect ( $\lambda$ ) and the initial viral spread of COVID-19 infections, we use the estimate from Column 4 of Table 3. In the ranking of travel routes by their marginal effects (4), we consider two rankings which correspond to different information sets of the decision maker. In the first ranking, the decision maker has complete information and the initial viral spread ( $\eta$ ) is based on the estimated residuals. In the second ranking, the decision maker believes that there is a log(10000) shock to Wuhan and none to other cities. The choice of the size of the shock is arbitrary because official estimates on the size of infections at the time when large scale public health measures were being implemented are not available<sup>8</sup>.

In Figure ??, travel routes are either closed or open. We now examine the case that the intensity of restrictions is variable, there is a trade-off between reduced infections and levels of restrictions, and numerically solve the optimization problems of (7) and (14). The costs of travel restrictions are parameterized by  $\varphi_1 = -1$  and  $\varphi_2 = 1$ , under which the marginal cost of new travel restrictions starting at no restrictions is zero, and the marginal cost is increasing with the levels of restrictions

---

<sup>8</sup>As a reference, there were 444 confirmed cases in Wuhan by January 22, 2020, one day before the lockdown of Wuhan. Wu *et al.* (2020) estimate that 75,815 individuals had been infected in Wuhan by January 25. Many people infected with COVID-19 do not show symptoms, however.

Figure 1: Selection of  $\tilde{g}$



This figure shows the number of COVID-19 cases per city and the cost of a COVID-19 confirmed case in equivalent monetary terms, that have the equal costs as a total shutdown of intercity travels across the whole country in Eq.(7), for different values of the weight  $\tilde{g}$ . It is assumed that without any restrictions, the population flow patterns in 2020 would be the same as those observed in 2019. For the monetary values, we assume that the cost of total shutdown of intercity travel equals to 40% of GDP in the first quarter of 2019 (1.246 trillion USD). The choice of 40% is because the GDP of Hubei province, which was subject to almost province-wide lockdowns since the outbreak, contracted by 39.18% year on year in the first quarter of 2020.

imposed, until it reaches  $w_{ij}$  at the maximum level of restrictions for the  $j \rightarrow i$  route (i.e.,  $\phi_{ij} = 0$ ). The relative costs of COVID-19 infections and travel restrictions are measured by the vector  $g$ , and we consider  $g = \tilde{g}\mathbf{1}$  with  $\tilde{g} \in \{1, 10, 25\}$ .

To provide economic meanings to the values of  $\tilde{g}$ , Figure 1 plots the number of confirmed COVID-19 cases per city and the corresponding monetary costs per confirmed case, whose costs are equivalent to a complete shutdown of intercity travels in the social planner's objective function (7). While this only serves an illustrative purpose because the values depend on assumptions such as the counterfactual population flow patterns in 2020 without government interventions and the economic costs of restrictions on intercity population flows which are topics in separate analysis (e.g., Fang *et al.*, 2020, Baker *et al.*, 2020), a value of  $\tilde{g}$  that is between 5 and 10 seems reasonable.

The set of solutions of the optimization problem is the off-diagonal elements of the  $n \times n$  matrix  $\Phi$ . Figure 2 shows the average levels of optimal outbound restrictions for the top 20 cities, according to the average levels of restrictions on outbound population flows  $(1 - \frac{1}{n-1} \sum_{i=1}^n \phi_{ij})$ . The

viral spread before spatial interactions is measured by  $\eta$ , and we consider two specifications which reflect different information sets of the decision makers, as in Figure ?? . In the first specification (upper panel of Figure 2), we assume that a  $\log(10000)$  unit shock is observed in Wuhan, and other elements of  $\eta$  are zeros. In the second specification (lower panel of Figure 2), we use the estimated  $\hat{\eta}$  based on the estimated coefficients from Column 2 of Table 3 in the appendix. Figure 3 shows the travel routes from Wuhan where the optimal travel restrictions are the most stringent.

In this illustrative example, the outbreak started in Wuhan. The optimal policies indicate that travels out of the outbreak city and cities that are closely connected to the outbreak city in terms of population flows should be strictly controlled, and this allows for less control on other cities. Among the travel routes from the outbreak city, those whose destination cities have large population outflows should be more strictly controlled, such as those in the Pearl River Delta, Jingjinji Metropolitan Region which includes Beijing, and the Yangtze River Delta (Figure 3). Travel routes with destinations in the Pearl River Delta should be more strictly restricted, because of the large population outflows from this area during this period of the year. The time window includes the Lunar New Year and there is a large number of migrant workers in cities in the Pearl River Delta who traditionally return to their hometowns before the Lunar New Year<sup>9</sup>. When cities act independently, they do not take into account the external benefit of their travel restrictions in reducing infections in other cities, and the levels of travel restrictions are not optimal, and in many cases, lower than those imposed when cities can coordinate. It is also interesting to note that when the infections are sufficiently costly ( $\tilde{g} = 25$ ), the social planner imposes stringent restrictions on population outflows from the outbreak city, which in turn may actually enable a less stringent control on population flows from other cities than the policies that are optimal when cities do not coordinate, as seen for the city of Xiaogan in Hubei province (upper panel of Figure 2) or Beijing (lower panel of Figure 2). This indicates that quick and stringent restrictions on cities with outbreak of costly infectious diseases could optimally allow for less restrictive public health measures in other places.

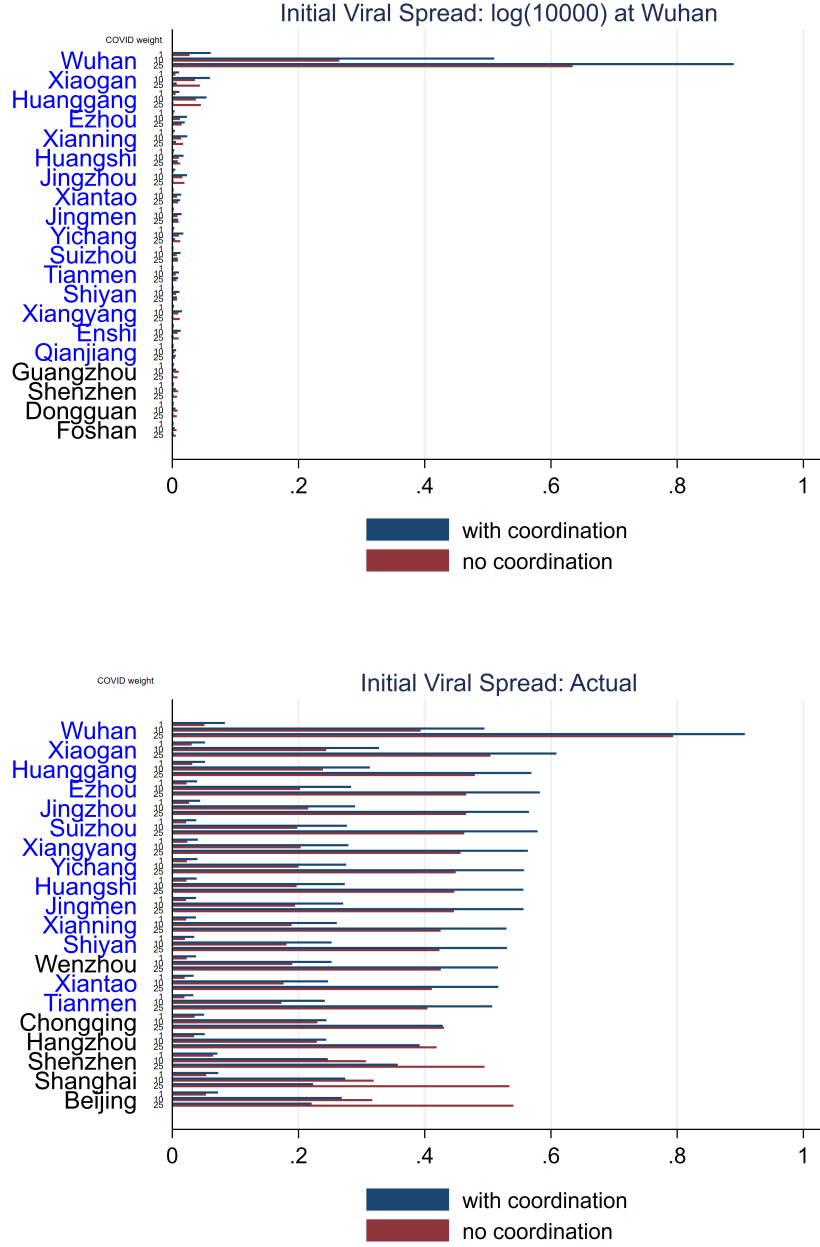
While the empirical illustration examines the spread of COVID-19 in China, our modelling framework can be applied to improve public health policy responses in different contexts, such as to a different infectious disease or in different geographic locations. With coordination between cities that takes into account the external effects of local policies, the optimal set of travel restrictions tends to place greater restrictions on routes whose destinations have large population outflows, and less restrictions on routes that have a lower likelihood of generating large spillovers, which is confirmed in Figure 4 that compares the actual population flow patterns in 2020 after government interventions with those that are predicted to be optimal according to the model.

---

<sup>9</sup>According to the 2010 population census of China, there were 85.9 million people who lived outside their home provinces in China, among whom 21.5 million lived in Guangdong province, the highest number among provinces. The Pearl River Delta includes Hong Kong, Macau and cities in Guangdong province such as Guangzhou, Shenzhen and Dongguan.

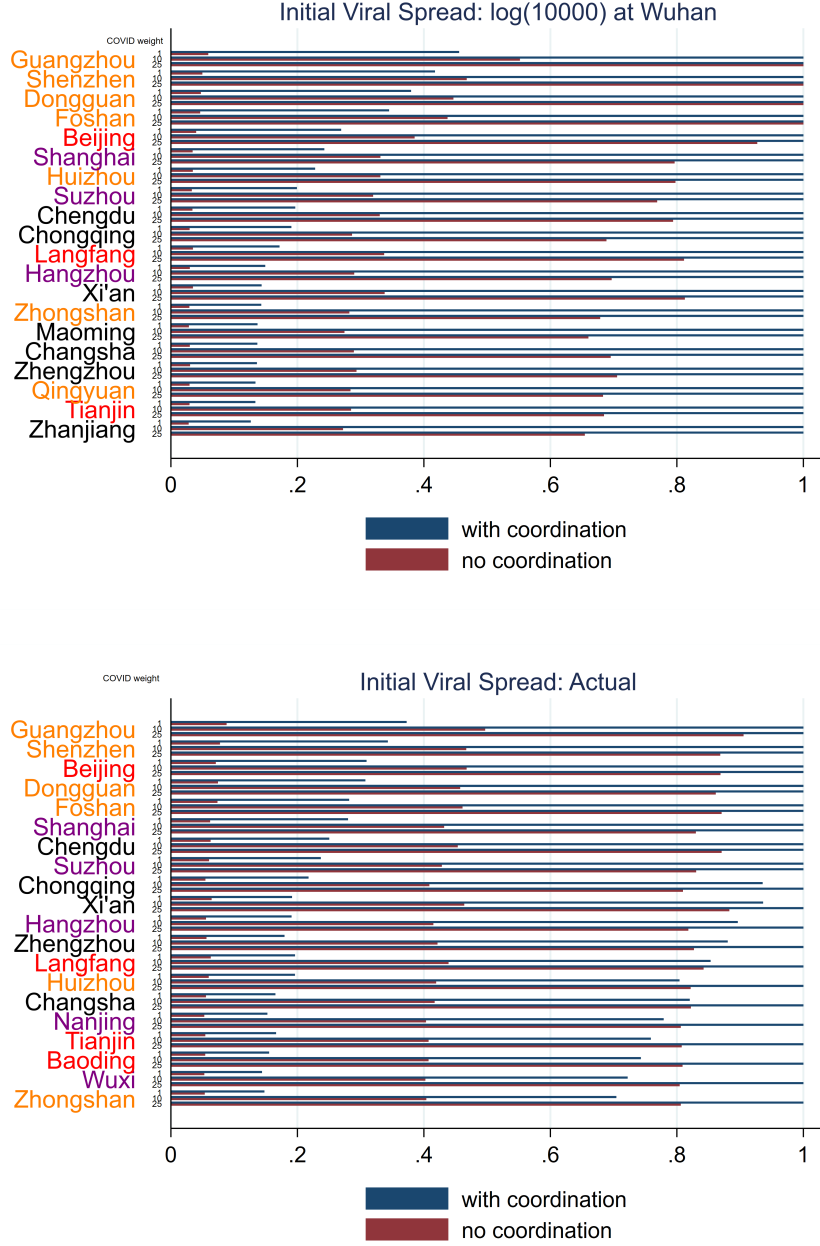


Figure 2: Average Optimal Levels of Outbound Restrictions, Top 20 Cities



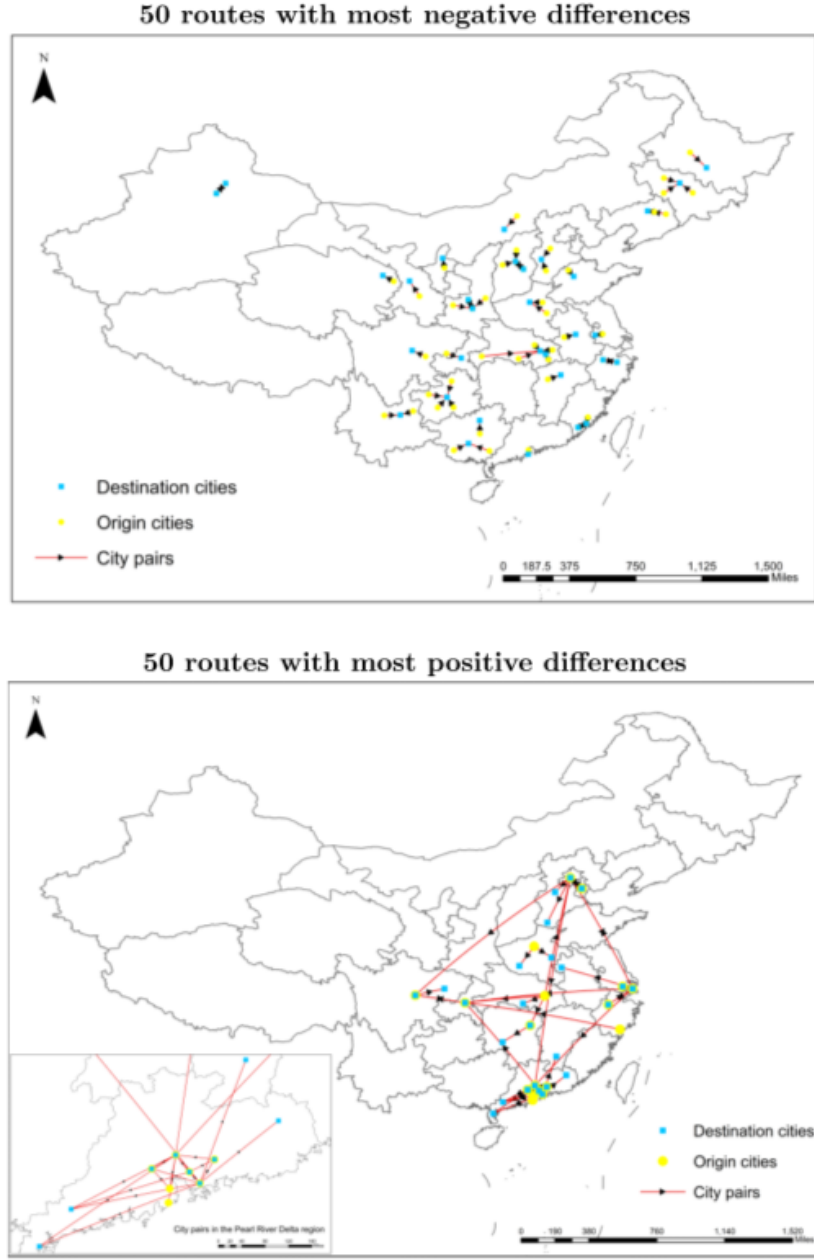
This figure shows the average levels of restrictions on outbound population flows for the top 20 cities ( $1 - \frac{1}{n-1} \sum_{i=1}^n \phi_{ij}$ ) with  $\tilde{g} = 10$  and coordination between cities.  $\tilde{g}$  is the value of the weight on COVID-19 infections, and is given in the “COVID weight” column. “With coordination” denotes solutions of the optimization problem of the social planner, and “no coordination” denotes solutions of the optimization problem where cities act independently. Cities in blue are in Hubei province.

Figure 3: Optimal Intensities of Outbound Restrictions from Wuhan, Top 20 Destinations



This figure shows the intensities of restrictions on outbound population flows from Wuhan for the top 20 destinations in terms of the intensity of restrictions  $(1 - \phi_{i, \text{Wuhan}})$  with  $\tilde{g} = 1$  and coordination between cities.  $\tilde{g}$  is the value of the weight on COVID-19 infections, and is given in the “COVID weight” column. “With coordination” denotes solutions of the optimization problem of the social planner, and “no coordination” denotes solutions of the optimization problem where cities act independently. Cities in orange/red/purple are in the Pearl River Delta/Jingjinji Metropolitan Region/Yangtze River Delta, respectively.

Figure 4: Actual and Optimal Levels of Intercity Population Flows



For each route, we compute the difference between the average population flow intensities in 2020 and the optimal level predicted by the model. Panel A shows the 50 routes with most negative differences, i.e., the optimal level of travel restriction is less restrictive than the actual policy implemented, and Panel B shows the 50 routes with most positive differences, i.e., the optimal policy is more restrictive than the actual one. The model assumes the following: (1) Travel patterns would be the same as those on the same lunar calendar days in 2019 in the absence of government interventions; (2) There are coordination between cities; (3) The weight  $\tilde{g}$  in Eq.(7) is 10. (4) Initial viral spread is based on the estimated coefficients from Column 2 of Table 3 in the appendix.

## 5 Conclusions and Discussion

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 the public support in the pandemic control measures. In this paper, we show that the marginal effect of decreasing intercity population flows in reducing total infections is not homogeneous but depends on the positions of the origin and destination cities in the network of population flows. The marginal benefit of population flow restrictions 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. Based on the dynamics of the spatial spread of diseases, we propose a model of optimal travel restrictions on each origin and destination city pairs, which considers the trade-off between the public health costs of infections and the economic costs from restricting population flows, and whether cities can coordinate their policies. We characterize the optimal solutions and show how they can be solved numerically.

We apply the theoretical model to the data on the spread of COVID-19 in China. Intercity population flows lead to spatial spread of the disease. Based on the estimated parameters, we identify routes in the population flow network that are most influential on the total number of COVID-19 cases in China. Travels out of the city experiencing virus outbreak or those whose destination cities having large population outflows will be restricted the most according to the optimal solution of travel restriction policies, which in turn may loosen a large portion of existing travel restrictions while substantially improve social welfare. Comparing the scenario when cities can coordinate their policies, the set of travel restriction policies with no coordination are suboptimal with lower total social welfare. The results of this paper could also be applied to optimize the public health measures in response to other infectious diseases where population flows significantly contribute to their spread.

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 countries at the beginning of an epidemic 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 to screen 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 virus spreading. 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. These data also guide border checks and surveillance (Wu *et al.*, 2020, Liu, 2020). Taiwan initiated health checks for airline travelers 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.*, 2020a). 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 that 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.

## 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*.
- Baker, S. R., Bloom, N., Davis, S. J. and Terry, S. J. (2020) COVID-induced economic uncertainty, *NBER Working Paper 26983*.
- 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.
- 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.
- 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.
- 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.

- 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 lock-down 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>.
- 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, *Working Paper*.
- Jeong, H. and Lee, L.-f. (2020) Spatial dynamic models with intertemporal optimization: Specification and estimation, *Journal of Econometrics*.
- 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.
- 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.
- 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*.

- 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. (2020a) Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing, *JAMA*, **323**, 1341–1342.
- Wang, M., Jiang, A., Gong, L., Luo, L., Guo, W., Li, C., Zheng, J., Li, C., Yang, B., Zeng, J., Chen, Y., Zheng, K. and Li, H. (2020b) Temperature significant change COVID-19 transmission in 429 cities, *medRxiv*.
- 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*.
- Wong, G. (2008) Has SARS infected the property market? Evidence from Hong Kong, *Journal of Urban Economics*, **63**, 74–95.
- 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*.