Centre for Economic Performance

Discussion Paper ISSN 2042-2695

No.1860
July 2022

Endogenous cross-region human mobility and pandemics

Xiao Chen
Hanwei Huang
Jiandong Ju
Ruoyan Sun
Jialiang Zhang

LSE The London School of Economics and Political Science

UKRI Economic and Social Research Council

Abstract
We study infectious diseases using a Susceptible-Infected-Recovered-Deceased model with endogenous cross-region human mobility. Individuals weigh the risk of infection against economic opportunities when moving across regions. The model predicts that the mobility rate of susceptible individuals declines with a higher infection rate at the destination. With cross-region mobility, a decrease in the transmission rate or an increase in the removal rate of the virus in any region reduces the global basic reproduction number (R0). Global R0 falls between the minimum and maximum of local R0s. A new method of Normalized Hat Algebra is developed to solve the model dynamics. Simulations indicate that a decrease in global R0 does not always imply a lower cumulative infection rate. Local and central governments may prefer different mobility control policies.

Key words: SIRD model, spatial economy, endogenous mobility, basic reproduction number, Normalized Hat Algebra
JEL: C61; C68; I18; J61; R13

This paper was produced as part of the Centre's Trade Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We would like to thank the editors and reviewers for their very helpful and constructive comments. We would like to thank Wenlan Luo, Yang Jiao, Larry Qiu, Chang Sun, and the seminar participants at the CICM 2021, HKU Trade and Development Workshop, and EEA 2021. We are responsible for any remaining errors. Huang is grateful for the funding support from the Research Grant Council of Hong Kong (project No. CityU 11501121).

Xiao Chen, School of International Trade and Economics, University of International Business and Economics. Hanwei Huang, Department of Economics and Finance, City University of Hong Kong and Centre for Economic Performance, London School of Economics. Jiandong Ju, PBC School of Finance, Tsinghua University. Ruoyan Sun, Department of Health Policy and Organization, School of Public Health, University of Alabama at Birmingham. Jialiang Zhang, PBC School of Finance, Tsinghua University.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

The 2019 novel coronavirus (COVID-19) pandemic has caused huge loss of life worldwide and plunged the global economy into a deep recession. Depending on the effect of interventions, there is a real possibility that the pandemic could last until 2025 (Kissler et al., 2020). Although containment policies have slowed the spread of the virus (Hsiang et al., 2020; Maier and Brockmann, 2020), they have also caused substantial disruptions in global air traffic (ICAO, 2020). Reopening national borders for travel while safeguarding lives from the pandemic will remain a critical policy issue in the coming years.

We develop a framework to study infectious diseases with endogenous cross-region mobility. We consider the interactions between human mobility and epidemiological dynamics in a multi-region Susceptible-Infected-Recovered-Deceased (SIRD) model. This model embeds two key mechanisms from the epidemiology and economic geography literature. First, contact between susceptible and infected people with mobility will transmit the virus both within the same region and across regions. Second, individuals consider factors that determine their welfare at destinations when making the mobility decision, including the probability of infection, recovery, and death. We show that these two mechanisms have important implications for the development of the pandemic and human mobility. Human mobility also responds to infectious diseases, which in turn reshapes the spread of the disease across space.

We motivate these two key mechanisms by presenting evidence using data from the US. To measure human mobility, we use anonymous smartphone user location data to construct state-to-state mobility flows at a biweekly frequency (Couture et al., 2021). Combining mobility data with state-level epidemiological and containment policy data for COVID-19, we establish two stylized facts using reduced-form econometric tools. *First, higher mobility inflows from states with more cases of infection are associated with increases in local cases.* We regress the number of new local cases on the average mobility inflows from other states weighted by the number of cases in the states of origin. We find a significant positive association, conditional on the stringency of local containment policies and state and time fixed effects. *Second, a higher local infection rate is associated with lower mobility inflows from other states.* The findings are similar whether we examine state-to-state level mobility flows or state-level aggregate net inflows. We also use a difference-in-differences strategy and find that states with higher cumulative infection rates had significantly lower net inflows after the March 2020 national emergency declaration.

Next, we build an SIRD model with endogenous cross-region mobility. In our model, contact between susceptible and infected individuals can generate new infections regardless of the origin of the infected individuals. Susceptible individuals weigh the risk of infection and economic opportunities across regions when choosing their optimal location for the next period. Using tools developed in the discrete choice and migration literature (Anderson et al., 1992; Artuç, et al., 2010; Caliendo et al., 2019), we derive the time-varying mobility rate for each type of individual. Consistent with the stylized fact, we show that susceptible individuals' mobility rate declines with a higher infection rate at their destination. In other words, they avoid moving to high-risk areas. Furthermore, they are the key drivers of mobility variations over time. We then analyze the properties

of the basic reproduction number, the expected number of susceptible people infected by an infected person when almost everyone is susceptible. Following Antràs et al. (2020), we call this number *global* R_0 in our environment with multiple regions. We show that global R_0 falls between the maximum and minimum of *local* R_0*s*, the basic reproduction number of a region under mobility autarky. We also prove that global R_0 declines whenever the virus has a lower transmission rate or higher removal rate in any region, provided cross-region mobility occurs.

We also use the model to study the effect of mobility costs on global R_0 and the cumulative infection rate using numerical simulations. Unfortunately, existing quantitative methods to solve mobility models, such as "Exact Hat Algebra" (Deckle et al., 2007) and "Dynamic Hat Algebra" (Caliendo et al., 2019), do not work for our model. The key hurdle is that the rate at which susceptible individuals become infected is endogenous to the distribution of individuals in the area and varies over time. To overcome this difficulty, we develop a new method called "Normalized Hat Algebra," which normalizes the model equations according to their corresponding steady-state equations. Although this method requires us to know the steady-state values, it enables us to compute the transitional dynamics and conduct counterfactual simulations along the transitional path.

Using Normalized Hat Algebra, we simulate a calibrated three-region economy and obtain the following findings. *First,* we find that susceptible individuals are the main drivers of mobility dynamics. This is because that other individuals (infected, recovered, or deceased) do not respond to changes in infection risk. *Second*, a central government prioritizing the global cumulative infection rate should adopt contingent mobility control. A universal decrease in cross-region mobility costs does not necessarily increase the cumulative infection rate, which depends on the origin of the pandemic. For example, if the pandemic originates from a region with the highest local R_0, decreasing mobility costs induces more individuals to move to other regions with lower local R_0, which reduces the global cumulative infection rate. In contrast, if the pandemic originates in the region with the lowest local R_0, decreasing cross-region mobility costs increases the global cumulative infection rate. *Third,* local governments prioritizing local cumulative infection rates may prefer different mobility policies than those preferred by the central government. At the local level, increasing cross-region mobility costs reduce the local cumulative infection rate. When the central government wants to lower mobility friction, its interests can conflict with those of local governments.

The first contribution of our paper is to the field of epidemiological models (Kermack and McKendrick, 1927). Studies show that human mobility contributed to the spread of SARS-CoV-2 across Chinese cities (Fang et al., 2020), US counties (Xiong et al., 2020) and large cities (Glaeser et al., 2022), and Italian provinces (Valsecchi and Durante, 2021). We confirm these findings using smartphone user data capturing US interstate mobility. Spatial friction also has important implications in SIRD models (Adda, 2016; Bartlett, 1956; Fajgelbaum et al., 2020; Muroya et al., 2013). However, human mobility is typically not included or is assumed to follow an exogenous process in these models.[1] Therefore, they do not capture behavioral responses in mobility in the face of infection risk, which can reduce infections even without containment policies (Fang et al., 2020; Farboodi

[1]For example, Fajgelbaum et al. (2021) take prepandemic commuting flows as given when studying optimal lockdowns.

et al., 2021; Goolsbee and Syverson, 2021). Antràs et al. (2020) and Bisin and Moro (2022) also consider individual behavioral responses in spatial-SIR (Susceptible-Infected-Recovered) models and highlight the role of geographic friction in shaping the spread of the virus. However, Antràs et al. (2020) consider interactions via international trade without modeling the movements of people across space. Bisin and Moro (2022) focus on within-city mobility, whereas we focus on cross-region mobility. They assume that individuals follow exogenous behavioral rules, whereas our modeled individuals optimize their mobility choices. To the best of our knowledge, this is the first paper to jointly consider endogenous human mobility in a multi-region SIRD model and analytically characterize the relationship between global and local basic reproduction numbers.

Second, we contribute to the literature on non-pharmaceutical interventions during a pandemic. By now, there is substantial evidence that these policies are effective in flattening the epidemic curve of COVID-19 (Fang et al., 2020; Hsiang et al., 2020; Maier and Brockmann, 2020). There are normative studies on optimal policies along various dimensions, including quarantine and testing (Berger et al., 2020; Piguillem and Shi, 2020), lockdowns (Acemoglu et al., 2021; Alvarez et al., 2021; Fajgelbaum et al., 2021), and general economic policies (Eichenbaum et al., 2021). Our theoretical results imply that all local regions should implement disease control policies, and disease control in the region with the highest local R_0 should be prioritized. Provided the highest local R_0 is less than 1, global R_0 is less than 1 regardless of any government-imposed mobility restraints. Our simulation results also highlight potential policy disagreements between local and central governments, assuming that local governments aim to decrease local cumulative infections and the central government aims to decrease global cumulative infections. We find that although local governments prefer to increase mobility friction, the central government prefers contingent mobility control.

Our final contribution is to the literature on quantitative economic geography (Allen and Arkolakis, 2014; Caliendo et al., 2019; Monte et al., 2018; Redding and Rossi-Hansberg, 2017; Tombe and Zhu, 2019). The key innovation of our model is that individuals can switch types and face different mobility problems in following periods, while the probability of switching is endogenous and depends on the distribution of individuals.[2] This makes existing quantitative methods dealing with dynamic mobility models, e.g., Caliendo et al. (2019), unsuitable for us.[3] We thus make a methodological contribution by developing the Normalized Hat Algebra approach to solve models similar to ours.

The rest of the paper is arranged as follows. Section 2 presents the stylized facts that motivate our model. Section 3 sets up the model and discusses the analytical results. Section 4 introduces our Normalized Hat Algebra approach. Section 5 presents the findings of our simulations using a calibrated three-region economy. Section 6 presents our conclusions and discusses extensions of our model.

[2]For example, the infection rate depends on the densities of susceptible and infected individuals in the local population (Acemoglu et al., 2021).

[3]Caliendo et al. (2021) extend the study by Caliendo et al. (2019) to models with multiple types of agents. They study the mobility of workers with different but fixed skills.

2 Motivating Stylized Facts

This section examines the empirical relationship between cross-region human mobility and the development of the COVID-19 pandemic. We first assemble a biweekly panel dataset for US states containing interstate mobility flows, containment policies, and epidemiological data for COVID-19 from January 2020 to August 2021. The details of our data are in Appendix B.1. Next, we use reduced-form econometric tools to establish two stylized facts on interstate mobility and the COVID-19 pandemic in the US. The details of our empirical analyses are in Appendix B.2.

Fact 1. Higher mobility inflows from states with more cases are associated with an increase in local cases.

When studying the impact of mobility on pandemics, an intuitive question is whether having more people moving from states with more cases leads to an increase in local cases. We estimate the number of newly confirmed local COVID-19 cases and exposure to the COVID-19 pandemic in other states caused by inter-state human mobility. Exposure is measured by the number of cases in other states weighted by prepandemic mobility shares. Table 1 presents the results. Across all columns, we find that having incomers from states with a higher number of cases is associated with more local cases, conditional on local containment policies and state and time fixed effects.

Fact 2. People avoid moving to states with high infection rates.

We next study how interstate mobility responds to the COVID-19 pandemic. Our first analysis focuses on the effect of contemporary infection rates at the origin and destination states on the size of bilateral mobility flows, see Table 2. The dependent variable measures the share of individuals who move from one state to another during a two-week period. Columns (1) and (2) use the maximum daily mobility flows during a two-week period as the dependent variable, and columns (3) and (4) use the average daily mobility flows during the same period. Across all columns, higher contemporary infection rates at the destination are associated with lower mobility inflows.

However, contemporary infection rates may not reflect infection risk, and bilateral mobility flows do not capture overall population flows to a region. To solve these problems, we then examine the effect of cumulative infection rates on total net mobility inflows. We employ a difference-in-differences strategy with varying treatment intensity. We measure a state's treatment intensity using the cumulative infection rate at the end of the data sample, on August 21, 2021. We assume that the treatment started on March 13, 2020, when the US government declared a national emergency. Table 3 presents the results. The dependent variable is biweekly net mobility inflows for each state. Again, columns (1) and (2) use the maximum daily mobility flows during a two-week period, while columns (3) and (4) use the daily average. We find that states with higher cumulative infection rates reported larger declines in mobility inflows. This result holds when controlling for the stringency of containment policies with state and time fixed effects. We also estimate a specification with time-varying treatment effects, see Appendix Figure B.1. Again, we find that mobility inflows to states with higher cumulative infection rates started to decline after the national emergency was declared.

3 Model

We set up a multi-region SIRD model with endogenous cross-region mobility. Individuals consider the risk of infection and the probability of recovery and death in different regions when they make mobility decisions. The model assesses how mobility and disease dynamics are determined by mobility friction and other model parameters.

3.1 Mobility, Demographic, and Disease Dynamics

Consider an economy consisting of N regions. We adopt the SIRD model (Kermack and McKendrick, 1927) and divide the total population into four types: Susceptible (S), Infected (I), Recovered (R), and Deceased (D). We assume that time is discrete and use $\overline{S}_{i,t}, \overline{I}_{i,t}, \overline{R}_{i,t}, \overline{D}_{i,t}$ to denote the population of each type in region i at the beginning of period t, and $S_{i,t}, I_{i,t}, R_{i,t}, D_{i,t}$ the end of period t. The size of each type in each region changes due to both the pandemic and mobility across regions.

For mobility, we follow the recent quantitative economic geography literature (Caliendo et al., 2019; Redding and Rossi-Hansberg, 2017; Tombe and Zhu, 2019) and assume that individuals of each type have idiosyncratic preferences for locations drawn from extreme value distributions, an approach commonly adopted in discrete choice models (Anderson et al., 1992). Specifically, the problem faced by type S in region i in period t can be characterized by the following Bellman equation,

$$U^S_{i,t}(\boldsymbol{\varepsilon_{i,t}}) = u_i + \max_{j=1,...,N}\{\beta E_t[(1-\alpha_{j,t+1})U^S_{j,t+1}(\boldsymbol{\varepsilon_{j,t+1}}) + \alpha_{j,t+1}U^I_{j,t+1}(\boldsymbol{\varepsilon_{j,t+1}})] - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\}, \tag{1}$$

where u_i is the instantaneous utility of uninfected individuals in region i, and $\beta \in (0,1)$ is the discount rate. Conditional on moving to region j, with probability $\alpha_{j,t+1}$, the individual will become infected and join type I with a value function of $U^I_{j,t+1}$. With probability $1-\alpha_{j,t+1}$, the individual will remain uninfected and have a value function of $U^S_{j,t+1}$. Following Antràs et al. (2020) and Farboodi et al. (2021), we assume that agents form rational expectations about $\alpha_{j,t+1}$. Although an individual is uncertain about whether or not he/she will be infected in the next period, the aggregate risk of infection is determined by collective decisions and taken as given by each individual. The cost of moving from region i to region j is $\widetilde{\mu}_{ij}$. In addition, individuals currently in region i receive a vector of preference shocks $\boldsymbol{\varepsilon_{i,t}} \equiv \{\varepsilon_{i1,t}, ..., \varepsilon_{iN,t}\}$ at the end of each period with each element $\varepsilon_{ij,t}$ following an *i.i.d.* Gumbel distribution

$$\Pr\{\varepsilon_{ij,t} \leq \varepsilon\} = \exp\left\{-\exp\left\{-\varepsilon/\kappa - \gamma^{Euler}\right\}\right\},$$

κ is a parameter that controls the dispersion of a shock and γ^{Euler} is the Euler constant ensuring that the shock has zero mean. These shocks capture the reality that individuals have idiosyncratic reasons for traveling to each region. After observing the preference shock, an individual moves to the region that yields the highest utility net of the mobility cost. Similarly, for types I and R, their value functions are given by

$$U^I_{i,t}(\boldsymbol{\varepsilon_{i,t}}) = u^I_i + \max_{j=1,...,N}\{\beta E_t[(1-\gamma^R_j-\gamma^D_j)U^I_{j,t+1}(\boldsymbol{\varepsilon_{j,t+1}}) + \gamma^R_j U^R_{j,t+1}(\boldsymbol{\varepsilon_{j,t+1}}) + \gamma^D_j U^D] - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\}, \tag{2}$$

and

$$U_{i,t}^{R}(\boldsymbol{\varepsilon_{i,t}}) = u_i + \max_{j=1,\ldots,N}\{\beta E_t[U_{j,t+1}^{R}(\boldsymbol{\varepsilon_{j,t+1}})] - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\}, \tag{3}$$

where γ_j^R and γ_j^D are the recovery and death rates in region j, and u_i^I is the instantaneous utility of type I in region i.[4] We assume that infection reduces instantaneous utility $0 < u_i^I < u_i$. U^D is the utility of type D, which we normalize as a constant such that the value of death is less than the value of recovery.

Solving problems (1), (2), and (3), we have the following results.

Proposition 1. *The expected lifetime utility of types S, I, and R in region i at time t can be written recursively as*

$$\begin{aligned}
V_{i,t}^{S} &= \exp\left(\frac{u_i}{\kappa}\right)\sum_{j=1}^{N}\left(V_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})}\left(V_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}}\left(\mu_{ij}\right)^{-1}, \\
V_{i,t}^{I} &= \exp\left(\frac{u_i^I}{\kappa}\right)\sum_{j=1}^{N}\left(V_{j,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)}\left(V_{j,t+1}^{R}\right)^{\beta\gamma_j^R}\left(V^D\right)^{\gamma_j^D}\left(\mu_{ij}\right)^{-1}, \\
V_{i,t}^{R} &= \exp\left(\frac{u_i}{\kappa}\right)\sum\nolimits_{j=1}^{N}\left(V_{j,t+1}^{R}\right)^{\beta}\left(\mu_{ij}\right)^{-1},
\end{aligned} \tag{4}$$

where $V_{i,t}^{g} = \exp(E_{t-1}(U_{i,t}^{g})/\kappa), g \in \{S, I, R\}$ and $\mu_{ij} = \exp(\tilde{\mu}_{ij}/\kappa)$, and $V^D = exp(U^D/\kappa)$ is the value of type D.

Proof. See Appendix A.3.1. □

Equation (4) reflects that the expected lifetime utility of individuals depends on the contemporaneous utility and value of changing regions and types in the next period. Using the method to find choice probabilities in discrete choice models, we obtain the mobility rate of each type of individual. The following proposition summarizes the results.

Proposition 2. *The mobility rates of types S, I, and R are*

$$\begin{aligned}
m_{ij,t}^{S} &= \frac{\left(V_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})}\left(V_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}}\left(\mu_{ij}\right)^{-1}}{\sum_{k=1}^{N}\left(V_{k,t+1}^{S}\right)^{\beta(1-\alpha_{k,t+1})}\left(V_{k,t+1}^{I}\right)^{\beta\alpha_{k,t+1}}\left(\mu_{ik}\right)^{-1}}, \\
m_{ij,t}^{I} &= \frac{\left(V_{j,t+1}^{I}\right)^{\beta\left(1-\gamma_j^R-\gamma_j^D\right)}\left(V_{j,t+1}^{R}\right)^{\beta\gamma_j^R}\left(V^D\right)^{\beta\gamma_j^D}\left(\mu_{ij}\right)^{-1}}{\sum_{k=1}^{N}\left(V_{k,t+1}^{I}\right)^{\beta(1-\gamma_k^R-\gamma_k^D)}\left(V_{k,t+1}^{R}\right)^{\beta\gamma_k^R}\left(V^D\right)^{\beta\gamma_k^D}\left(\mu_{ik}\right)^{-1}}, \\
m_{ij,t}^{R} &= \frac{\left(V_{j,t+1}^{R}\right)^{\beta}\left(\mu_{ij}\right)^{-1}}{\sum_{k=1}^{N}\left(V_{k,t+1}^{R}\right)^{\beta}\left(\mu_{ik}\right)^{-1}},
\end{aligned} \tag{5}$$

where $m_{ij,t}^{g}$ is the probability that type g moves from region i to region j in period t, $g \in \{S, I, R\}$.

Proof. See Appendix A.3.2. □

[4] To keep the model non-trivial, we assume that $0 \leq \gamma_j^R < 1$, $0 \leq \gamma_j^D < 1$, and $\gamma_j^R + \gamma_j^D < 1$, $\forall j$. If $\gamma_j = 0$, $\forall j$, our model collapses to an SIR model.

It is worth mentioning two differences between these expressions and existing spatial models. First, our mobility rates can vary across types of individuals, whereas existing spatial models typically ignore mobility heterogeneity by assuming a single type of worker or commuter.[5] Second, in addition to mobility friction, disease parameters (infection, recovery, and death rates) directly affect the mobility rate. Individuals who can switch types (types S and I) take into account that they might switch to other types in the following period, whereas individuals of the terminated type R do not.

As is standard in SIRD models (Allen and den Driessche, 2008), the demographic dynamics within a period caused by the disease are given by

$$\begin{aligned} S_{i,t} &= \overline{S}_{i,t} - T_{i,t},\ I_{i,t} = T_{i,t} + \left(1 - \gamma_i^R - \gamma_i^D\right)\overline{I}_{i,t}, \\ R_{i,t} &= \overline{R}_{i,t} + \gamma_i^R \overline{I}_{i,t},\ D_{i,t} = \overline{D}_{i,t} + \gamma_i^D \overline{I}_{i,t}, \end{aligned} \tag{6}$$

where $T_{i,t}$, the number of newly infected people in region i in period t, satisfies

$$\frac{T_{i,t}}{\overline{L}_{i,t}} = \chi_i \frac{\overline{S}_{i,t}}{\overline{L}_{i,t}} \frac{\overline{I}_{i,t}}{\overline{L}_{i,t}}, \tag{7}$$

where $\overline{L}_{i,t} = \overline{S}_{i,t} + \overline{I}_{i,t} + \overline{R}_{i,t}$ is the stock of living individuals and χ_i is the transmission rate, the probability that the virus will transmit from type I to type S conditional on contact between them. Given the number of newly infected individuals, the infection rate of type S is

$$\alpha_{it} \equiv \frac{T_{i,t}}{\overline{S}_{i,t}}. \tag{8}$$

Last, the demographic dynamics across periods due to mobility are given by

$$\begin{aligned} \overline{S}_{i,t} &= \sum_{j=1}^{N} S_{j,t-1} m^S_{ji,t-1},\ \overline{I}_{i,t} = \sum_{j=1}^{N} I_{j,t-1} m^I_{ji,t-1}, \\ \overline{R}_{i,t} &= \sum_{j=1}^{N} R_{j,t-1} m^R_{ji,t-1},\ \overline{D}_{i,t} = D_{i,t-1}. \end{aligned} \tag{9}$$

The timing of events in the model is summarized in Figure 1.

3.2 Equilibrium

The fundamentals of our model are bilateral mobility costs $\mu \equiv \{\mu_{ij}\}_{i=1,j=1}^{N,N}$, the recovery rate $\gamma^R \equiv \left\{\gamma_i^R\right\}_{i=1}^{N}$, the death rate $\gamma^D \equiv \left\{\gamma_i^D\right\}_{i=1}^{N}$, the transmission rate $\chi \equiv \{\chi_i\}_{i=1}^{N}$, the value of type D V^D, the dispersion of the parameter of the Gumbel distribution κ, and the discount rate β. We denote the fundamentals by $\overline{\Omega} \equiv \left\{\mu, \gamma^R,\ \gamma^D,\ \chi,\ V^D,\ \kappa,\ \beta\right\}$.

If we denote the distribution of individuals across regions by $G_t \equiv \{S_{i,t}, I_{i,t}, R_{i,t}, D_{i,t}\}_{i=1}^{N}$, their mobility rates by $m_t = \{m^S_{ij,t},\ m^I_{ij,t},\ m^R_{ij,t}\}_{i=1,j=1}^{N,N}$ and expected lifetime utility by $V_t = \{V^S_{i,t},\ V^I_{i,t},\ V^R_{i,t}\}_{i=1}^{N}$, we can define the sequential competitive equilibrium.

[5] A notable exception is Caliendo et al. (2021), who study the migration of workers with different skill types. However, they continue to assume that workers do not change types.

Definition 1. *Given $\overline{\Omega}$ and G_0, the sequential competitive equilibrium is a sequence of $\{G_t,\ m_t,\ V_t\}_{t=0}^{\infty}$ satisfying equations (4), (5), (6), (7), (8), and (9).*

Next, we can define the *steady-state equilibrium*.

Definition 2. *A steady-state equilibrium is a sequential competitive equilibrium such that $\{G_t,\ m_t,\ V_t\}_{t=0}^{\infty}$ are constant over time.*

Finally, following the literature on the basic reproduction number R_0 (Van den Driessche, 2017), we define the *disease-free equilibrium* (DFE) as

Definition 3. *A DFE is defined as the point at which the whole population is susceptible, i.e., $S_{i,t}/L_{i,t} = 1,\ \forall i$.*

3.3 Mobility and Infection Risk

SIRD models have analytical solutions only under certain restrictions (Harko et al., 2014). Embedding them in an economic model with endogenous mobility makes it difficult to obtain analytical solutions. This subsection presents analytical results regarding how infection risk shapes individual mobility choices during the pandemic. First, given our assumptions that instantaneous utility and mobility friction are time-invariant, we have the following results.

Proposition 3. *The value functions of types I and R are time-invariant, that is, $V_{i,t}^R = V_i^R$ and $V_{i,t}^I = V_i^I$; so are their mobility rates, i.e., $m_{ij,t}^R = m_{ij}^R$ and $m_{ij,t}^I = m_{ij}^I$.*

Proof. See Appendix A.3.3. □

Types I and R have both been infected. They are not concerned about the risk of infection and do not respond to changes in infection rates. Only type S individuals, who have not been infected, care about the risk of infection when making mobility choices. Therefore, this proposition implies that the dynamics of human mobility during a pandemic are shaped mostly by type S individuals. Before discussing how susceptible individuals respond to the risk of infection, we must rank the value function of agents.

Lemma 1. *(a) $V_i^R > V_i^I$ and $V_i^R \geq V_{i,t}^S$; (b) Near the DFE, we have $V_i^I < V_{i,t+1}^S$.*

Proof. See Appendix A.3.4. □

Lemma 1 implies that as long as infection reduces instantaneous utility and death is strictly dominated by recovery, the value function of type R individuals is higher than that of types I and S. The intuition is that type I individuals might die while type S individuals might become infected. Near the DFE, the infection rate approaches 0. Then, the value of type S approaches the value of type R according to equation (4), which is greater than the value of type I. Using this lemma, we obtain the following result, which relates the risk of infection to the mobility of type S.

Proposition 4. *Near the DFE, if the probability of infection in region j increases, the mobility rate of type S to region j decreases weakly: $\frac{\partial \ln(m_{ij,t}^S)}{\partial \alpha_{j,t+1}} = (1 - m_{ij,t}^S) \ln(\frac{V_j^I}{V_{j,t+1}^S}) \leq 0$.*

Proof. See Appendix A.3.5. □

This result implies that type S individuals will avoid moving to regions with higher risks of infection. This is consistent with the stylized fact established in Section 2. We note that type S's response is proportional to the log difference of their value function from type I, $\ln(V_j^I) - \ln(V_{j,t+1}^S)$. The greater the welfare loss when becoming infected, the greater the response of type S individuals to infection risk. If the welfare of type I is the same as that of type S, i.e., $V_j^I = V_{j,t+1}^S$, we have $\partial \ln(m_{ij,t}^S)/\partial \alpha_{j,t+1} = 0$. In this case, there is no difference in welfare whether or not one becomes infected. Type S individuals simply do not respond to infection risk.

In addition, the magnitude of the response is proportional to the total mobility rate to other destinations $1 - m_{ij,t}^S$. When $1 - m_{ij,t}^S = 0$, again we have $\partial \ln(m_{ij,t}^S)/\partial \alpha_{j,t+1} = 0$, because there is nowhere else to go.[6] Naturally, the mobility of type S does not respond to changes in region j's infection risk. Otherwise, the higher the mobility rate to other destinations, the greater the response of type S to an increase in infection risk in region j. Our intuition is that if the mobility rate to other regions increases, other regions become more attractive as a destination than region j. Then, if the conditions in region j deteriorate, individuals respond more aggressively to these changes. Therefore, mobility also shapes individuals' response to infection risk.

3.4 Local and Global R_0

The basic reproduction number, R_0, is used to determine whether an emerging epidemic can be contained. R_0 measures the expected number of new infections generated by an infected individual near the DFE (Diekmann et al., 1990). In SIRD models without cross-region human mobility or other interactions, R_0 of region i is given by

$$R_{0,i}^L = \frac{\chi_i}{\gamma_i}, \tag{10}$$

which we define as *local* R_0 in region i. In this expression, χ_i is the expected number of new infections generated by an infected individual near the DFE in each period. $1/\gamma_i$ is the expected duration of infection. Therefore, $R_{0,i}^L = \chi_i/\gamma_i$ captures the expected number of new infections from an infected individual during their infection. If $R_{0,i}^L > 1$, the epidemic will continue to spread in the region and die out if $R_{0,i}^L < 1$.

This notion is applied to multi-group SIR models in which individuals may differ by age, gender, or other characteristics (Allen and Van den Driessche, 2008; Diekmann et al., 2010) or come from multiple regions (Antràs et al., 2020). We generalize the notion of R_0 to an environment with human mobility across regions.

If we combine equations (6) and (9) and use the result of Proposition 3 that $m_{ij,t}^I = m_{ij}^I$, the law of motion for type I can be written as:

$$I_{i,t} = \chi_i \frac{\overline{S}_{i,t}}{\overline{L}_{i,t}} \overline{I}_{i,t} + \left(1 - \gamma_i^R - \gamma_i^D\right) \overline{I}_{i,t} = \left(\chi_i \frac{\overline{S}_{i,t}}{\overline{L}_{i,t}} + 1 - \gamma_i^R - \gamma_i^D\right) \sum_{j=1}^{N} m_{ji}^I I_{j,t-1}. \tag{11}$$

[6] When the preference shocks follow an extreme value distribution, $m_{ij,t}^S = 1$ if mobility costs to destinations other than region j are infinite. If mobility costs are finite, the mobility rate to any destination is positive.

We follow Allen and Van den Driessche (2008) and vectorize the equation above near the DFE (which implies that $\frac{\overline{S}_{i,t}}{\overline{L}_{i,t}} = 1,\ \forall i$) and obtain

$$\mathbf{I}_{t+1} = \mathbf{F}\mathbf{I}_t + \mathbf{V}\mathbf{I}_t,$$

where $\mathbf{I_t} = [I_{1,t}, ..., I_{N,t}]'$, and

$$\mathbf{F} = \begin{bmatrix} \chi_1 & & & \\ & \chi_2 & & \\ & & \ddots & \\ & & & \chi_N \end{bmatrix} \begin{bmatrix} m^I_{11} & m^I_{21} & \cdots & m^I_{N1} \\ m^I_{12} & m^I_{22} & \cdots & m^I_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ m^I_{1N} & m^I_{2N} & \cdots & m^I_{NN} \end{bmatrix},$$

and

$$\mathbf{V} = \begin{bmatrix} 1-\gamma_1 & & & \\ & 1-\gamma_2 & & \\ & & \ddots & \\ & & & 1-\gamma_N \end{bmatrix} \begin{bmatrix} m^I_{11} & m^I_{21} & \cdots & m^I_{N1} \\ m^I_{12} & m^I_{22} & \cdots & m^I_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ m^I_{1N} & m^I_{2N} & \cdots & m^I_{NN} \end{bmatrix},$$

where $\gamma_i = \gamma^D_i + \gamma^R_i,\ (1 \leq i \leq N)$, is the removal rate of type I in region i. $\mathbf{F}$ is the transmission matrix, which captures new infections generated by transmission. $\mathbf{V}$ is the transition matrix, which captures the transition of type I to type R or D. In our model, individuals can move from one region to another, potentially generating new infections. This is captured by the mobility matrix $\mathbf{m} = \{m^I_{ji}\}_{N\times N}$, which appears in both $\mathbf{F}$ and $\mathbf{V}$. With $\mathbf{F}$ and $\mathbf{V}$, we follow Allen and Van den Driessche (2008) and define the *next-generation matrix* as

$$\mathbf{M} = \mathbf{F}\left(\mathbb{I} - \mathbf{V}\right)^{-1},$$

where $\mathbb{I}$ is an identity matrix. This is a concept borrowed from the literature on matrix population models that examines demographic dynamics and views infection as consecutive generations of infected people (Caswell, 2001). $\mathbf{M}$ connects the number of newly infected cases in different regions for consecutive periods. The basic reproduction number is given by the spectral radius, the largest norm of eigenvalues, of $\mathbf{M}$:

$$R^G_0 \equiv \rho(\mathbf{M}) \equiv \max_{1\leq i\leq N}\{|\Lambda_i|\}, \tag{12}$$

where Λ_i is the eigenvalue of $\mathbf{M}$. R^G_0 is the *global* R_0 of our multi-region economy.

This concept is similar to the interpretation of R^L_0 in the traditional SIRD model without cross-region mobility. First, the element on the i^{th} row and j^{th} column of $\mathbf{F}$, $\chi_i m^I_{ji}$, is the expected number of new cases generated by an infected individual from region j in region i. As for $\mathbf{V}$, the element on the i^{th} row and j^{th} column, $(1-\gamma_i)m^I_{ji}$, is the survival probability of type I from region j in region i. Then, the element on the i^{th} row and j^{th} column of $(\mathbb{I} - \boldsymbol{V})^{-1}$ is the expected duration of infection for an infected individual in region i who moved from region j.[7] Therefore, the element on the i^{th} row

[7]Let ξ_i be the expected duration of infection in region i and $\xi = [\xi_1, ..., \xi_N]$, then ξ satisfies a recursive relationship $\xi = \iota + \xi V$, and $\iota = [1, ..., 1]_{1\times N}$. Immediately, we have $\xi = \iota\left(\mathbb{I} - \boldsymbol{V}\right)^{-1}$.

and j^{th} column of $\mathbf{M}$ represents the expected number of new infections in region i as a result of an infected individual moving from region j.

Nevertheless, the multi-region environment requires further interpretation. We consider a small initial outbreak across regions near the DFE with $\delta_{\mathbf{0}} = \mathbf{I_0}$ and call it the 0-th generation of infection. Then, the expected number of new cases directly generated by the 0-th generation is given by $\delta_{\mathbf{1}} = \mathbf{M}\delta_{\mathbf{0}}$, which is the first generation of cases across regions. In general, we can calculate the k-th generation of infection as $\delta_{\mathbf{k}} = \mathbf{M}\delta_{\mathbf{k-1}} = \mathbf{M}^k\mathbf{I_0}$. Clearly, the growth rate of $\delta_{\mathbf{k}}$ is controlled by the largest norm of eigenvalues of $\mathbf{M}$, namely R_0^G.[8] We next discuss the properties of R_0^G.

Proposition 5. *R_0^G falls between the minimum and maximum of local reproduction numbers, i.e.,* $\min_i R_{0,i}^L \leq R_0^G \leq \max_i R_{0,i}^L$.

Proof. See Appendix A.3.6. □

Clearly, if everyone stays in the region with the lowest R_0^L, we have $R_0^G = \min_i R_{0,i}^L$. Similarly, if everyone stays in the region with the highest local R_0, we have $R_0^G = \max_i R_{0,i}^L$. In general, people can move across regions. R_0^G falls between the maximum and minimum of local R_0. Mobility becomes irrelevant for R_0^G when the disease parameters are identical across regions. R_0^G is the same as R_0^Ls as $\min_i R_{0,i}^L = \max_i R_{0,i}^L$.

Proposition 5 suggests that R_0^G is bounded from above by the region with the highest R_0^L. This has important policy implications. First, if $\max_i R_{0,i}^L < 1$, we have $R_0^G < 1$. Accordingly, if the pandemic can be controlled in the worst region, then allowing cross-region mobility will not cause an outbreak given that R_0^G is also less than 1. This result also provides the rationale for prioritizing disease control in the region with the highest R_0^L. If the pandemic can be controlled there, it can be contained globally.

In a world with cross-region mobility, all regions matter. The next proposition suggests that in a globalized world with cross-region mobility, improved infectious disease control in any region contributes to global disease control.

Proposition 6. *A decrease in the transmission rate χ_i or an increase in the removal rate γ_i in any region weakly decreases global R_0^G.*

Proof. See Appendix A.3.7. □

Combining these two propositions, we derive two important elements of infection control. First, all local regions should implement disease control policies. Second, disease control in the region with the highest R_0^L should be prioritized.

4 Normalized Hat Algebra

Hat Algebra is a set of methodologies for performing comparative static and counterfactual analysis. Jones (1965) was the first to propose the use of comparative statics in trade models by log linearization. Instead of log linearization, Dekle et al. (2007) transformed their model in terms of changes from the current equilibrium and called their method

[8] We note that $\lim_{k\to\infty} \frac{\boldsymbol{\delta_k}}{\zeta^k} = \lim_{k\to\infty}(\frac{\boldsymbol{M}}{\zeta})^k I_0$ is finite if $\zeta \geq \rho(\boldsymbol{M})$.

"Exact Hat Algebra"(EHA), as it is an approach without approximation. However, EHA can only be applied to static or steady models. Caliendo et al. (2019) extended EHA and developed the method of "Dynamic Hat Algebra" (DHA) to solve models with intrinsic migration dynamics. Their idea is to transform the model in terms of growth rates. They show that this approach allows them to derive the transitional path and conduct a counterfactual analysis along the transitional path in a class of models.

Our model does not belong to the class of models that can be solved by DHA. To check that, we follow Caliendo et al. (2019) and rewrite the value function of type S in terms of growth rates as follows:

$$\dot{V}^S_{i,t+1} = \sum_{k=1}^{N} m^S_{ik,t} (\dot{V}^S_{k,t+2})^{\beta(1-\alpha_{k,t+2})} (\dot{V}^I_{k,t+2})^{\beta\alpha_{k,t+2}} \left(\frac{V^S_{k,t+1}}{V^I_{k,t+1}}\right)^{\beta(\alpha_{k,t+1}-\alpha_{k,t+2})},$$

where $\dot{V}^g_{i,t+1} \equiv \frac{V^g_{i,t+1}}{V^g_{i,t}}$ is the growth rate of the value function of type g, $g \in \{S, I\}$. However, the last term in the equation, $(V^S_{k,t+1}/V^I_{k,t+1})^{\beta(\alpha_{k,t+1}-\alpha_{k,t+2})}$ is still expressed in terms of the value function *level* and appears whenever the probability of infection fluctuates over time: $\alpha_{k,t+1} \neq \alpha_{k,t+2}$. Because $\alpha_{k,t+1}$ captures the probability of infection, which is endogenous and can vary over time, we cannot rewrite our model purely in terms of growth rates of value functions, and we solve it using DHA.

To solve the problem, we develop a Normalized Hat Algebra approach by normalizing all value functions to their corresponding steady-state value. We denote the normalized value by $\widehat{x}_t = x_t/x_{ss}$, while x_{ss} is the steady-state of x and prove the following result.

Proposition 7. *Given the initial distribution of agents, G_0, and the steady state of the economy, the solution to the sequential competitive equilibrium solves equations (6), (7), (8), and (9), together with the following equations for all regions in all periods*

$$\widehat{V}^S_{i,t} = \sum_{j=1}^{N} m^S_{ij,ss} \left(\widehat{V}^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V^I_{j,ss}}{V^S_{j,ss}}\right)^{\beta\alpha_{j,t+1}},$$

$$m^S_{ij,t} = \frac{m^S_{ij,ss} \left(\widehat{V}^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V^I_{j,ss}}{V^S_{j,ss}}\right)^{\beta\alpha_{j,t+1}}}{\sum_{k=1}^{N} m^S_{ik,ss} \left(\widehat{V}^S_{k,t+1}\right)^{\beta\left(1-\alpha_{k,t+1}\right)} \left(\widehat{V}^I_{k,t+1}\right)^{\beta\alpha_{k,t+1}} \left(\frac{V^I_{k,ss}}{V^S_{k,ss}}\right)^{\beta\alpha_{j,t+1}}},$$

where $m^S_{ij,ss}$ is the steady-state mobility rates of type S.

Proof. See Appendix A.3.8 □

We note that the value functions and mobility rates of types I and R are the same as their steady-state values given Proposition 3, i.e., $\widehat{V}^I_{i,t} = \widehat{V}^R_{i,t} = 1$, $m^I_{ij,t} = m^I_{ij,ss}$, and $m^R_{ij,t} = m^R_{ij,ss}$. As for the corresponding values for type S, they contain the term $(V^I_{j,ss}/V^S_{j,ss})^{\beta\alpha_{j,t+1}}$. However, the problem here is much simpler because $\frac{V^I_{j,ss}}{V^S_{j,ss}}$ is a constant.

This approach requires us to solve the steady-state values. The steady-state values for types S, I, and R can be solved using equation (4) by replacing $\alpha_{j,ss} = 0$ $(\forall j)$ for given

instantaneous utility and model fundamentals $\overline{\Omega}$. Then, the steady-state mobility rates, $m^g_{ij,ss}$, $g \in \{S, I, R\}$, can be computed from equation (5). The detailed algorithm of the approach is presented in Appendix B.5.

5 Mobility, R_0, and Cumulative Infection Rate

Restrictions of human mobility have been one of the most prominent anti-contagion policies implemented during the COVID-19 pandemic (Hale et al., 2021). In this section, we investigate how mobility friction affects R_0^G and the cumulative infection rate. We calibrate our model to a three-region economy. The model parameters are specified in Appendix Table B.2. We assume that these three regions have the same removal rate, but Region 1 has the lowest disease transmission rate, Region 2 the medium rate, and Region 3 the highest rate. Therefore, R_0^L values are ranked as follows: $R^L_{0,1} < R^L_{0,2} < R^L_{0,3}$.

Finding 1. Type S individuals avoid moving to regions with higher infection risk and drive the overall mobility dynamics during the pandemic.

We use the Normalized Hat Algebra approach developed in the previous section to simulate the model and compute the cumulative infection rate for each region. Figure 2 plots the transitional dynamics when the pandemic originated from Region 1.[9] Panel (a) shows the number of type S individuals in each region, and Panel (f) shows their mobility rate to Region 1 relative to the steady state. Consistent with Stylized Fact 1 and Proposition 4, type S individuals avoid moving to Region 1, the region with outbreaks and the highest infection risk. However, as type I individuals in Region 1 recover or die (Panel b), although the cumulative number of cases increases (Panel e), infection risk gradually declines (Panel d), fewer type S individuals choose to move to another region, and more type S individuals from other regions move to Region 1. Therefore, the share of type S individuals rebounds in Region 1. Overall, Region 1's share of type S individuals exhibits a "V" shape, and the total population size in Region 1 shows a similar trend (Panel c).

Finding 2. A universal increase in cross-region mobility friction increases R_0^G.

We now consider how a universal rise in mobility friction affects R_0^G and the cumulative infection rate. Figure 3 (a) plots the results.[10] We find that increasing cross-region mobility friction *increases* R_0^G. The intuition for this result is as follows. When the mobility matrix is an identity matrix, i.e., $\mathbf{m} = \mathbb{I}$, there is no cross-region mobility. It is easy to verify that R_0^G reaches the maximum $R_0^G = \rho(\mathbf{M}) = \max_i\{\frac{\chi_i}{\gamma_i}\}$ using equation (12) because type S individuals tend to move from regions with higher infection risk to regions with lower infection risk. When mobility restrictions are relaxed, type S individuals can move to regions with lower infection risk, which tends to reduce R_0^G. However, a lower R_0^G is not equivalent to fewer infected people, which is our next finding.

[9] We assume that 1% of the population got infected in the initial period.

[10] To capture mobility friction, we decompose the mobility matrix as $\mathbf{m}(\sigma) = (1-\sigma)\mathbf{m^I_{ss}} + \sigma\mathbb{I}$, $\sigma \in [0, 1]$, and $\mathbf{m^I_{ss}}$ is the steady-state mobility matrix. σ controls mobility friction in a parsimonious manner. A higher σ implies higher mobility friction. $\mathbf{m}(1) = \mathbb{I}$ is the case of full isolation, as the probability of moving to nonlocal regions is 0.

Finding 3. Increasing R_0^G does not always increase the cumulative infection rate.

R_0^G is difficult to calculate and cannot be directly observed. Governments may care more about the cumulative infection rate, which is a straightforward measure of the severity of the pandemic. Figure 3 (a) also shows that the effect of mobility friction on the cumulative infection rate depends on the origin of the pandemic. When it originates from the region with the highest R_0^L, Region 3, the cumulative infection rate increases with mobility friction. The main reason is that fewer individuals move from Region 3 to other regions with lower transmission rates when mobility costs increase. This tends to increase overall infections. In contrast, if the pandemic originates from Region 1, the region with the lowest R_0^L, the cumulative infection rate decreases when mobility friction increases because people are less likely to move to regions with higher transmission rates.

Finding 4. Increasing a region's mobility friction leads to fewer local infections. However, increasing mobility friction in a region with low R_0^L increases the global cumulative infection rate.

So far, we have only considered universal changes in cross-region mobility friction. Now, we consider changes in mobility friction with respect to moving in and out of a region. Figure 3 (b) plots the results when mobility friction changes at the regional level and outbreaks happen in all three regions.[11] In all cases, reducing mobility friction with other regions leads to more local infections. However, the global cumulative infection rate declines if the region with the lowest R_0^L reduces its mobility friction with other regions. This is intuitive as agents move to the region with the lowest transmission rate, which reduces the global cumulative infection rate. In contrast, reducing mobility friction for the region with the highest R_0^L raises overall infections. This result suggests that if the objective of governments is to reduce the cumulative infection rate, the preferred mobility friction policy can differ between local and central governments. Whereas local governments always prefer to limit mobility, the central government prefers regions with low R_0^L to relax mobility control while regions with high R_0^L to tighten mobility control.

6 Conclusion

Rapid globalization and urbanization have generated enormous gains from human mobility, but may also have led to more frequent outbreaks of infectious diseases. We develop a theoretical framework to jointly study infectious diseases and human mobility. The model allows us to analyze how human mobility responds to infection risk and study the properties of global and local R_0s. We enrich the existing toolboxes of spatial economies by developing a new Normalized Hat Algebra approach, which solves the dynamic mobility model with endogenous type switching. We use the calibrated model to investigate the effects of mobility friction on R_0 and the cumulative infection rate. We find some

[11] 1% of the local population is infected in each region initially. Instead of scaling the entire mobility matrix, we scale the specific row and column associated with the inflows and outflows of a region.

limitations of R_0^G and a potential conflict in mobility policy preferences between local and central governments.

Our model is general and versatile. It can be extended to study relevant disease containment policies implemented by governments. For example, vaccination programs are effective in reducing the rate of infection (Chen et al., 2022). In Appendix B.6, we discuss an extension of our model to incorporate vaccinations. Other components could also be added. For example, following Redding and Rossi-Hansberg (2017), production and interregional trade could be introduced to the model. Incorporating them could introduce economic forces of social distancing (Antràs et al., 2020) but would not change our model predictions about R_0 and how people respond to infection risk. Future research could build on our framework to further study epidemics and economics.

References

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, M.D., 2021. Optimal targeted lockdowns in a multigroup SIR model. American Economic Review: Insights, 3(4), pp.487-502.

Adda, J., 2016. Economic activity and the spread of viral diseases: Evidence from high frequency data. The Quarterly Journal of Economics, 131(2), pp.891-941.

Allen, L.J., and Van den Driessche, P., 2008. The basic reproduction number in some discrete-time epidemic models. Journal of Difference Equations and Applications, 14(10-11), pp.1127-1147.

Allen, T., and Arkolakis, C., 2014. Trade and the topography of the spatial economy. The Quarterly Journal of Economics, 129(3), pp.1085-1140.

Alvarez, F.E., Argente, D., and Lippi, F., 2021. A simple planning problem for COVID-19 lock-down, testing, and tracing. American Economic Review: Insights, 3(3), pp.367-382.

Anderson, S.P., De Palma, A., and Thisse, J.F., 1992. Discrete choice theory of product differentiation. MIT Press.

Antràs, P., Redding, S.J., and Rossi-Hansberg, E., 2020. Globalization and pandemics (No. w27840). National Bureau of Economic Research.

Artuç, E., Chaudhuri, S., and McLaren, J., 2010. Trade shocks and labor adjustment: A structural empirical approach. American Economic Review, 100(3), pp.1008-45.

Atkeson, A., 2020. What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios (No. w26867). National Bureau of Economic Research.

Bartlett, M.S., 1956. Deterministic and stochastic models for recurrent epidemics. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 4(81), p.109.

Berger, D.W., Herkenhoff, K.F., and Mongey, S., 2020. An SEIR infectious disease model with testing and conditional quarantine (No. w26901). National Bureau of Economic Research.

Bethune, Z.A., and Korinek, A., 2020. Covid-19 infection externalities: Trading off lives vs. livelihoods (No. w27009). National Bureau of Economic Research.

Bisin, A., and Moro, A., 2022. JUE insight: Learning epidemiology by doing: The empirical implications of a Spatial-SIR model with behavioral responses. Journal of Urban Economics, 127, p.103368.

Caliendo, L., Dvorkin, M., and Parro, F., 2019. Trade and labor market dynamics: General equilibrium analysis of the China trade shock. Econometrica, 87(3), pp.741-835.

Caliendo, L., Opromolla, L.D., Parro, F., and Sforza, A., 2021. Goods and factor market integration: A quantitative assessment of the EU enlargement. Journal of Political Economy, 129(12), pp.3491-3545.

Caswell, H., 2000. Matrix population models (Vol. 1). Sunderland, MA: Sinauer.

Chen, X., Huang, H., Ju, J., Sun, R., and Zhang, J., 2022. Impact of vaccination on the COVID-19 pandemic in US states. Scientific Reports, 12(1), pp.1-10.

Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., y Piontti, A.P., Mu, K., Rossi, L., Sun, K., and Viboud, C., 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science, 368(6489), pp.395-400.

Costinot, A., and Rodríguez-Clare, A., 2014. Trade theory with numbers: Quantifying the consequences of globalization. In Handbook of International Economics, 4, pp. 197-261. Elsevier.

Couture, V., Dingel, J.I., Green, A., Handbury, J., and Williams, K.R., 2021. JUE insight: Measuring movement and social contact with smartphone data: A real-time application to COVID-19. Journal of Urban Economics, p.103328.

Dekle, R., Eaton, J., and Kortum, S, 2007. Unbalanced trade. American Economic Review, 97(2), pp.351-355.

Diekmann, O., Heesterbeek, J.A.P., and Metz, J.A., 1990. On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations. Journal of Mathematical Biology, 28(4), pp.365-382.

Diekmann, O., Heesterbeek, J.A.P., and Roberts, M.G., 2010. The construction of next-generation matrices for compartmental epidemic models. Journal of the Royal Society Interface, 7(47), pp.873-885.

Eichenbaum, M.S., Rebelo, S., and Trabandt, M., 2021. The macroeconomics of epidemics. The Review of Financial Studies, 34(11), pp.5149-5187.

Fajgelbaum, P.D., Khandelwal, A., Kim, W., Mantovani, C., and Schaal, E., 2021. Optimal lockdown in a commuting network. American Economic Review: Insights, 3(4), pp.503-22.

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, p.104272.

Farboodi, M., Jarosch, G., and Shimer, R., 2021. Internal and external effects of social distancing in a pandemic. Journal of Economic Theory, 196, p.105293.

Glaeser, E.L., Gorback, C., and Redding, S.J., 2022. JUE insight: How much does COVID-19 increase with mobility? Evidence from New York and four other US cities. Journal of Urban Economics, p.103292.

Goolsbee, A., and Syverson, C., 2021. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. Journal of Public Economics, 193, p.104311.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., and Tatlow, H., 2021. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nature Human Behaviour, 5(4), pp.529-538.

Harko, T., Lobo, F.S., and Mak, M.K., 2014. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. Applied Mathematics and Computation, 236, pp.184-194.

Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L.Y., Hultgren, A., Krasovich, E., and Lau, P., 2020. The effect of large-scale anti-contagion policies on the COVID-19 Pandemic. Nature, 584(7820), pp.262-267.

ICAO, Air Transport Bureau, 2020. Effects of novel coronavirus (COVID-19) on civil aviation: Economic impact analysis. International Civil Aviation Organization (ICAO), Montréal, Canada.

Jones, R.W., 1965. The structure of simple general equilibrium models. Journal of Political Economy, 73(6), pp.557-572.

Kermack, W.O., and McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character, 115(772), pp.700-721.

Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H., and Lipsitch, M., 2020. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science, 368(6493), pp.860-868.

Li, C.K., and Schneider, H., 2002. Applications of Perron-Frobenius theory to population dynamics. Journal of Mathematical Biology, 44(5), pp.450-462.

Maier, B.F., and Brockmann, D., 2020. Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. Science, 368(6492), pp.742-746.

Monte, F., Redding, S.J., and Rossi-Hansberg, E., 2018. Commuting, migration, and local employment elasticities. American Economic Review, 108(12), pp.3855-90.

Muroya, Y., Enatsu, Y., and Kuniya, T., 2013. Global stability of extended multi-group SIR epidemic models with patches through migration and cross patch infection. Acta Mathematica Scientia, 33(2), pp.341-361.

Piguillem, F., and Shi, L., 2020. Optimal COVID-19 quarantine and testing policies. EIEF Working Papers Series 2004.

Raifman, J., Nocka, K., Jones, D., Bor, J., Lipson, S., Jay, J., and Galea, S., 2020. COVID-19 US state policy database. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2020-07-16. https://doi.org/10.3886/E119446V17

Redding, S.J., and Rossi-Hansberg, E., 2017. Quantitative spatial economics. Annual Review of Economics, 9, pp.21-58.

Silva, J.S., and Tenreyro, S., 2006. The log of gravity. The Review of Economics and Statistics, 88(4), pp.641-658.

Tombe, T., and Zhu, X., 2019. Trade, migration, and productivity: A quantitative analysis of China. American Economic Review, 109(5), pp.1843-72.

Valsecchi, M., and Durante, R., 2021. Internal migration networks and mortality in home communities: Evidence from Italy during the Covid-19 pandemic. European Economic Review, 140, p.103890.

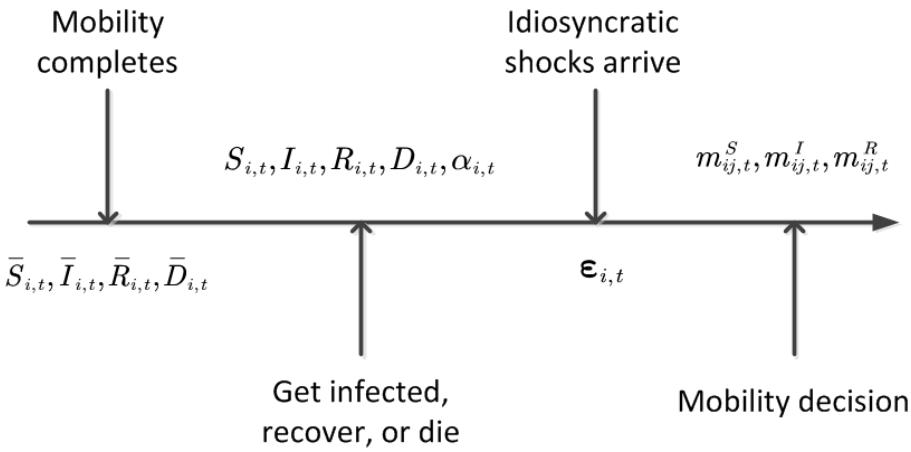
Van den Driessche, P., 2017. Reproduction numbers of infectious disease models. Infectious Disease Modelling, 2(3), pp.288-303.

Xiong, C., Hu, S., Yang, M., Luo, W., and Zhang, L., 2020. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. Proceedings of the National Academy of Sciences, 117(44), pp.27087-27089.

Appendix

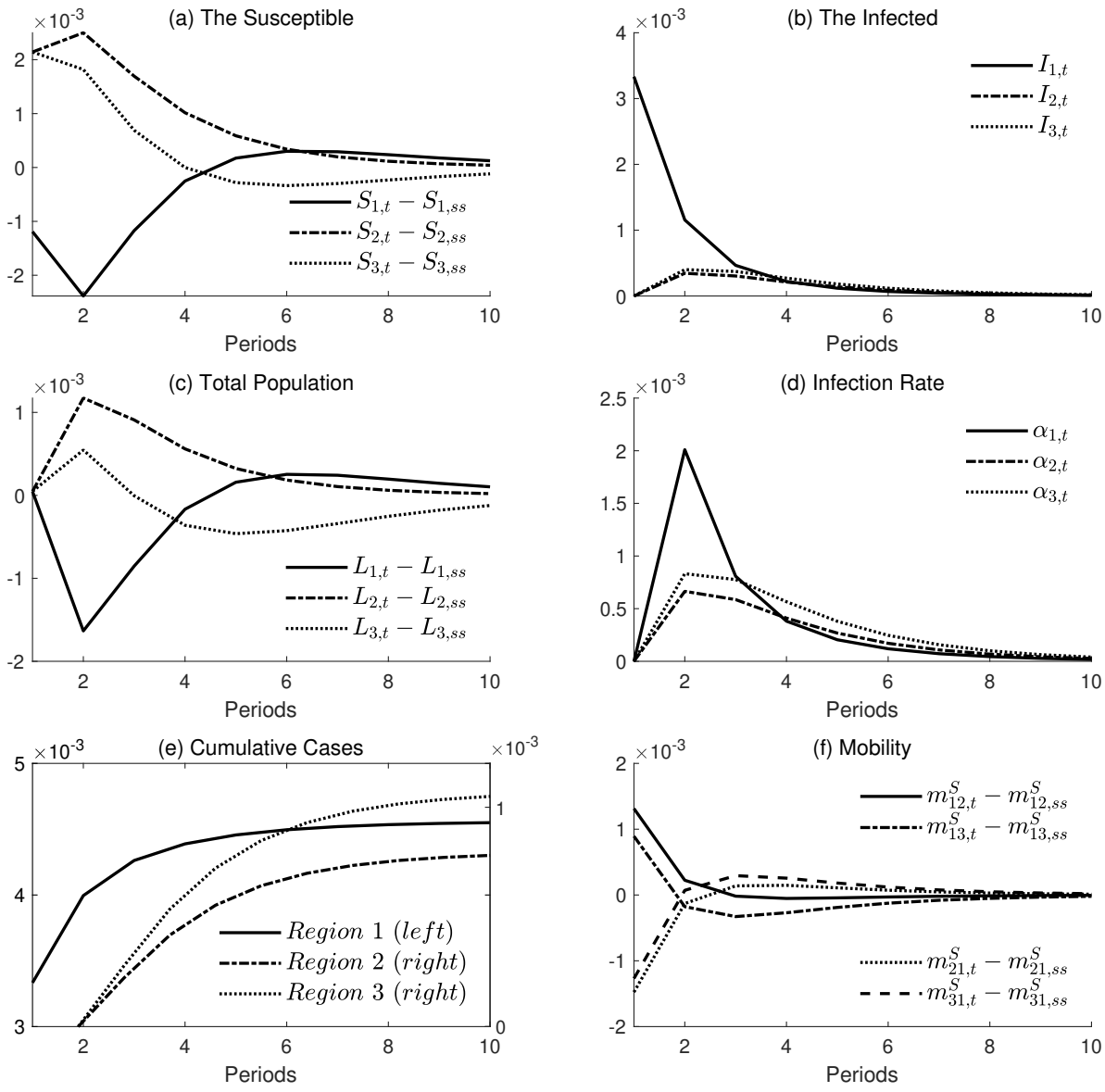
A.1 Figures

Figure 1: Model Timeline



Notes. This figure plots the timeline of events in each model in period t. At the beginning of the period (after mobility in the previous period is complete), the stock of type g individuals in region i is $\overline{g}_{i,t}, g \in \{S, I, R, D\}$. Next, possible infection (with probability $\alpha_{i,t}$), recovery, and death take place. The stock of type g individuals becomes $g_{i,t}$. At the end of the period, individuals derive their idiosyncratic preference shock $\varepsilon_{i,t}$ and make a mobility decision. The mobility rate of type g is $m^g_{ij,t}$.

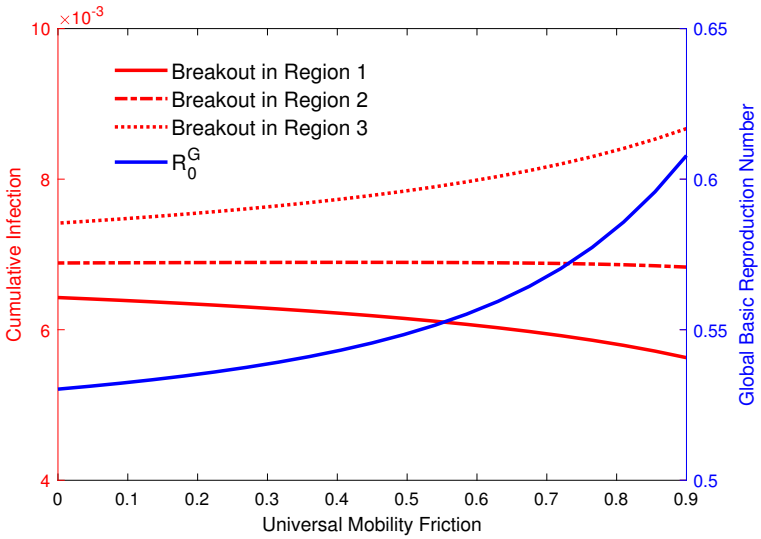
Figure 2: Epidemic and Mobility Dynamics with an Outbreak in Region 1



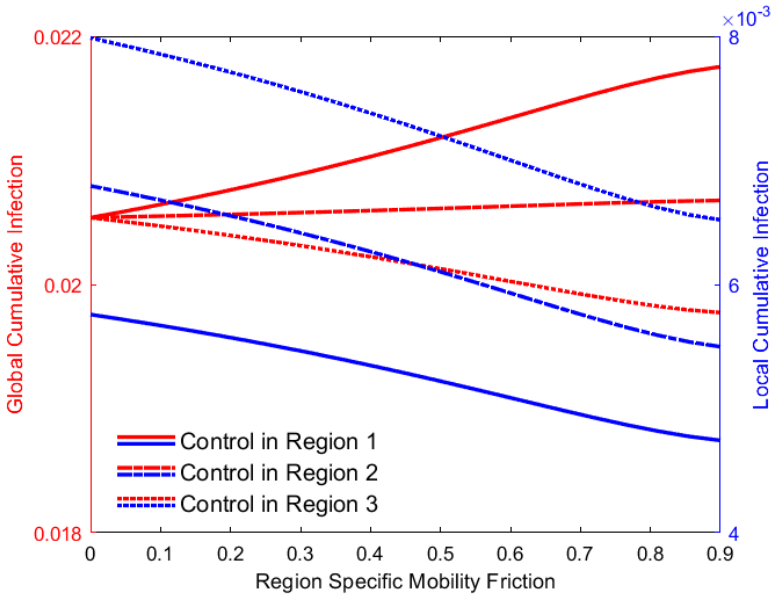
Notes. This figure plots the simulation results from a calibrated three-region economy with an initial outbreak in Region 1 (1% of the local population is infected). The model parameters are specified in Table B.2. Panel (a) plots the stock of type S individuals relative to the steady state in each region; Panel (b) plots the number of type I individuals; Panel (c) plots the size of the local population relative to the steady state; and Panel (d) plots the infection rate. Panel (e) plots the cumulative number of cases for Regions 1 (left axis), 2, and 3 (right axis) by summing the number of new cases across all periods. Panel (f) plots the mobility rate of type S individuals moving in and out of Region 1 relative to the steady state.

Figure 3: Mobility Friction, Global R_0, and Cumulative Infection Rate

(a) Universal Mobility Control



(b) Regional Mobility Control



Notes. This figure plots the simulation results from a calibrated three-region economy. The model parameters are specified in Table B.2. Panel (a) plots how universal mobility control, which increases all cross-region mobility costs (horizontal axis), affects the global cumulative infection rate (left axis) and R_0^G (right axis). We simulate three scenarios, with outbreaks starting from Regions 1, 2, and 3, respectively. Panel (b) plots how global (left axis) and local (right axis) cumulative infection rates vary with regional mobility friction (horizontal axis) when outbreaks occur in all regions. We simulate three scenarios, with increasing mobility friction in and out of Regions 1, 2, and 3, respectively.

A.2 Tables

Table 1: Interstate Mobility and the COVID-19 Pandemic in the US

	(1)	(2)	(3)	(4)
VARIABLES	ln (number of new cases)			
$\sum_{j\neq i} m_{ji}^{max} ln(NewCases_{j,t-1})$	0.374***			
	(22.695)			
$\sum_{j\neq i} m_{ji}^{mean} ln(NewCases_{j,t-1})$		0.394***		
		(23.017)		
$\sum_{j\neq i} m_{ji}^{max} ln(CumulativeCases_{j,t-1})$			0.159***	
			(13.737)	
$\sum_{j\neq i} m_{ji}^{mean} ln(CumulativeCases_{j,t-1})$				0.171***
				(13.973)
$stringency\ index_{t-1}$	-0.004*	-0.004*	-0.002	-0.002
	(-1.766)	(-1.742)	(-0.629)	(-0.623)
Observations	2,040	2,040	2,040	2,040
R-squared	0.971	0.972	0.943	0.943
BiWeek FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes

Notes. This table estimates the effect of interstate mobility on state-level COVID-19 cases in the US. The dependent variable is the logarithm of the number of new cases in state i during a two-week period. For the independent variables, we use four measures of local exposure to outside cases caused by interstate mobility (columns 1-4). They combine two measures of initial interstate mobility and two measures of outside cases. The two measures of initial bilateral mobility are as follows: m_{ji}^{max}, the maximum daily mobility rate from state j to state i during the initial two-week period; and m_{ji}^{mean}, the average daily mobility rate during the initial two-week period. Two measures of outside cases are used: $ln(NewCase_{j,t-1})$ is the logarithm of the lagged number of new cases in origin state j, and $ln(CumulativeCase_{j,t-1})$ is the logarithm of the lagged number of cumulative positive cases in origin state j. The $stringency\ index_{t-1}$ measures the stringency of closure policies, ranging from 0 to 100 at the state and biweekly levels. A higher stringency index indicates stricter containment policies. A one-period lag is used to deal with the potential endogeneity of containment policies. The sample includes all 50 states and DC from January 2020 to August 2021 with 40 periods of two weeks. The numbers in parentheses are robust t-statistics with standard errors two-way clustered at the state and biweekly levels. Significance is indicated by *, **, and *** at the 0.1, 0.05 and 0.01 levels, respectively.

Table 2: The COVID-19 Pandemic and US State-to-State Human Mobility

	(1)	(2)	(3)	(4)
VARIABLES	Max mobility		Average mobility	
new infection rate of origin	-0.940	-0.491	-0.601	-0.146
	(-0.436)	(-0.642)	(-0.272)	(-0.201)
new infection rate of destination	-7.633***	-7.844***	-7.456***	-7.698***
	(-3.432)	(-10.862)	(-3.245)	(-11.085)
mobility restrictions of origin	-0.048***	-0.048***	-0.048***	-0.048***
	(-3.531)	(-9.873)	(-3.408)	(-9.602)
mobility restrictions of destination	-0.007	-0.008	-0.003	-0.005
	(-0.448)	(-1.498)	(-0.201)	(-0.827)
log(distance)	-1.364***		-1.384***	
	(-271.503)		(–214.016)	
shared state border	1.069***		1.100***	
	(95.547)		(96.924)	
Observations	102,000	102,000	102,000	102,000
Time FE	Yes	Yes	Yes	Yes
Origin FE	Yes	No	Yes	No
Destination FE	Yes	No	Yes	No
Origin-Destination Pair FE	No	Yes	No	Yes

Notes. This table estimates how interstate bilateral mobility flows respond to the risk of COVID-19 infection. The main dependent variables are as follows: $Max\ mobility_{ij,t}$, which is the maximum daily bilateral mobility flow between a state pair within a two-week period; and $Average\ mobility_{ij,t}$, which is the average daily bilateral mobility flow between a state pair within a two-week period. *new infection rate of origin* is measured by the number of new cases divided by the population in the state of origin, and *new infection rate of destination* is the same measure for the destination state. To deal with potential endogeneity, both variables are lagged by one period. *mobility restrictions of origin* measures the stringency of the restrictions on mobility imposed by the state of origin, and *mobility restrictions* is the measure imposed by the destination state. Time-invariant gravity variables include distance and a common border between the state of origin and the destination. The sample includes all 2,550 state pairs in the US from January 2020 to August 2021, with 40 periods of two weeks. Columns (2) and (4) control for Origin-Destination state pair fixed effects, *log*(*distance*) and *shared state border* are absorbed from the pair fixed effects and are not reported in the table. The equations are estimated using Poisson Pseudo Maximum Likelihood (PPML). The numbers in parentheses are robust t-statistics with standard errors two-way clustered at the state pair and biweekly levels. Significance is indicated by *, **, and *** at the 0.1, 0.05, and 0.01 levels, respectively.

Table 3: Effect of the COVID-19 Pandemic on Net Mobility inflows to US states

VARIABLES	(1)	(2)	(3)	(4)
	Share of net inflow (by max mobility)		Share of net inflow (by average mobility)	
cumulative infection rate × post pandemic	-0.056**	-0.069**	-0.032	-0.045**
	(-2.039)	(-2.494)	(-1.524)	(-2.094)
stringency index		-0.010***		-0.009***
		(-2.893)		(-2.966)
Observations	2,040	2,040	2,040	2,040
R-squared	0.956	0.956	0.956	0.956
Biweek FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes

Notes. This table estimates how state-level net mobility inflows respond to the risk of COVID-19 infection. The main dependent variables are as follows: *Share of net inflow by max mobility*, which is net inflows divided by the population at the state level, and net inflow is calculated by aggregating the inflow (measured by the maximum daily flow within a biweekly period) from other states, netting out outflows to other states. *Share of net inflow by average mobility* is constructed similarly, except that the average daily bilateral mobility flow is used. *cumulative infection rate* is computed as cumulative cases over the total state population at the end of the sample period, the last two-week period in August 2021. *post pandemic* is a dummy variable taking a value of 1 after the US declared a national emergency on March 13, 2020, or the fourth biweekly period in our time window. *stringency index* illustrates the stringency of closure policies, which ranges from 0 to 100 and varies at the state and biweekly levels. We lag this variable by one period to deal with the potential endogeneity of containment policies. A higher value of *Stringency Index* indicates stricter containment policies. The sample includes all 50 states and DC from January 2020 to August 2021 with 40 periods of two weeks. The numbers in parentheses are robust t-statistics with standard errors two-way clustered at the state and biweekly levels. Significance is indicated by *, **, and *** at the 0.1, 0.05, and 0.01 levels, respectively.

A.3 Proofs

A.3.1 Proof of Proposition 1

We start from the value function and mobility matrix of the recovered individuals. Taking expectation on both sides of equation (3) and denote $\widetilde{V}^R_{i,t} \equiv E_{t-1}U^R_{i,t}(\boldsymbol{\varepsilon_{i,t}})$, we have

$$\widetilde{V}^R_{i,t} = u_i + E_{\boldsymbol{\varepsilon_{i,t}}} \max_j \{\beta \widetilde{V}^R_{j,t+1} - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\}. \tag{A.1}$$

Let's denote $V \equiv \max_j \{\beta \widetilde{V}^R_{j,t+1} - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\}$, we can calculate its expectation under our assumption that $\varepsilon_{ij,t}$ follows the Gumbel distribution:

$$\begin{aligned}
E_{\boldsymbol{\varepsilon_{i,t}}} V &= \int V \mathrm{d} Pr\{\forall j, \varepsilon_{ij,t} \leq V + \widetilde{\mu}_{ij} - \beta \widetilde{V}^R_{j,t+1}\} \\
&= \int V \mathrm{d} \prod_j \exp\left\{-\exp\left\{-[V + \widetilde{\mu}_{ij} - \beta \widetilde{V}^R_{j,t+1}]/\kappa - \gamma^{Euler}\right\}\right\} \\
&= \int V \mathrm{d} \exp\left\{-\sum_j (V^R_{j,t+1})^\beta \mu_{ij}^{-1} \exp\left\{-V/\kappa - \gamma^{Euler}\right\}\right\}
\end{aligned}$$

.

Let $Z \equiv \kappa \ln \sum_j (V^R_{j,t+1})^\beta \mu_{ij}^{-1}$, then the integral above can be transformed as:

$$\begin{aligned}
E_{\boldsymbol{\varepsilon_{i,t}}} V &= \int V \mathrm{d} \exp\left\{-\exp\left\{-(V - Z)/\kappa - \gamma^{Euler}\right\}\right\} \\
&= \int (V' + Z) \mathrm{d} \exp\left\{-\exp\left\{-V'/\kappa - \gamma^{Euler}\right\}\right\} = Z,
\end{aligned}$$

where the last equality is due to our normalization of the mean of the Gumbel distribution as 0. Substituting the result above to equation (A.1), we get

$$\widetilde{V}^R_{i,t} = u_i + E_{\boldsymbol{\varepsilon_{i,t}}} V \Rightarrow V^R_{i,t} = \exp\left(\frac{u_i}{\kappa}\right) \sum\nolimits_{j=1}^{N} \left(V^R_{j,t+1}\right)^\beta (\mu_{ij})^{-1}.$$

The value functions of type S and type I can be derived in similar ways.

A.3.2 Proof of Proposition 2

To simply notations, we denote $Z_{ij} = \beta \widetilde{V}^R_{j,t+1} - \widetilde{\mu}_{ij}$, the mobility rate from i to j for type R satisfies:

$$\begin{aligned}
m^R_{ij,t} &= Prob\{\forall k \neq j, \beta \widetilde{V}^R_{k,t+1} - \widetilde{\mu}_{ik} + \varepsilon_{ik,t} \leq \beta \widetilde{V}^R_{j,t+1} - \widetilde{\mu}_{ij} + \varepsilon_{ij,t}\} \\
&= Prob\{\forall k \neq j, \varepsilon_{ik,t} \leq \varepsilon_{ij,t} + Z_{ij} - Z_{ik}\} \\
&= \int \prod_{k \neq j} \exp\left\{-\exp\left\{-(\varepsilon_{ij,t} + Z_{ij} - Z_{ik})/\kappa - \gamma^{Euler}\right\}\right\} \mathrm{d} \exp\left\{-\exp\left\{-\varepsilon_{ij,t}/\kappa - \gamma^{Euler}\right\}\right\}.
\end{aligned}$$

Denote $Y = \exp\left\{-\exp\left\{-\varepsilon_{ij,t}/\kappa - \gamma^{Euler}\right\}\right\}$, we have:

$$m^R_{ij,t} = \int Y^{\sum_{k\neq j}\exp\{(Z_{ik}-Z_{ij})/\kappa\}}\mathrm{d}Y = \frac{1}{\sum_{k\neq j}\exp\{(Z_{ik}-Z_{ij})/\kappa\}+1}$$
$$= \frac{\exp\{Z_{ij}/\kappa\}}{\sum_k \exp\{Z_{ik}/\kappa\}} = \frac{\left(V^R_{j,t+1}\right)^{\beta}(\mu_{ij})^{-1}}{\sum_{k=1}^N \left(V^R_{k,t+1}\right)^{\beta}(\mu_{ik})^{-1}}.$$

The mobility rates for type S and I can be derived in similar ways.

A.3.3 Proof of Proposition 3

Denote $\widehat{x} = x/x_{ss}$, while x_{ss} is the steady-state value of x, we have $\hat{V}^R_{i,t}$ and $\hat{V}^I_{i,t}$ converging to 1 at the steady state. It implies that $\forall\delta > 0$ small enough, $\exists\ T$ such that $\forall t > T$, $|\hat{V}^R_{i,t} - 1| < \delta$. Using the Taylor expansion, we know that $|\left(\hat{V}^R_{i,T+1}\right)^{\beta} - 1| < \beta\delta + o(\delta)$. Suppose for any period s, $\exists\ \epsilon$ such that $|\hat{V}^R_{i,s} - 1| = \epsilon > 0$, from equations (4) and (5), we know

$$\hat{V}^R_{i,T} - 1 = \sum_{j=1}^N m^R_{ij}\left[\left(\hat{V}^R_{j,T+1}\right)^{\beta} - 1\right]$$
$$\Leftrightarrow |\hat{V}^R_{i,T} - 1| < \sum_{j=1}^N m^R_{ij}(\beta\delta + o(\delta)) = \beta\delta + o(\delta)$$
$$\Leftrightarrow |\hat{V}^R_{i,s} - 1| < \beta^{T+1-s}\delta + o(\delta)$$

Then for the given δ, since $\beta < 1$ we can always find some T large enough such that

$$|\hat{V}^R_{i,s} - 1| < \beta^{T+1-s}\delta + o(\delta) < \epsilon, \tag{A.2}$$

contradiction, we have $\hat{V}^R_{i,t} = 1$, $\forall t$. Similarly, we can show that $\hat{V}^I_{i,t} = 1$. When the value functions of type R and I are constant, it is easy to see that $m^R_{ij,t} = m^R_{ij}$ and $m^I_{ij,t} = m^I_{ij}$ from equation (5).

A.3.4 Proof of Lemma 1

Proposition 3 tells that the lifetime utility type R and I are time-invariant, then equation (4) becomes

$$V^S_{i,t} = e^{u_i/\kappa}\sum_{j=1}^N (V^S_{j,t+1})^{\beta(1-\alpha_{j,t+1})}(V^R_I)^{\beta\alpha_{j,t+1}}\mu^{-1}_{ij}$$
$$V^I_i = e^{u^I_i/\kappa}\sum_{j=1}^N (V^I_j)^{\beta(1-\gamma^R_j-\gamma^D_j)}(V^R_j)^{\beta\gamma^R_j}(V^D_j)^{\beta\gamma^D_j}\mu^{-1}_{ij}$$
$$V^R_i = e^{u_i/\kappa}\sum_{j=1}^N (V^R_j)^{\beta}\mu^{-1}_{ij}$$

Divide the equation of type I by the one of type R, and use equation (5), we get:

$$\frac{V_i^I}{V_i^R} = e^{\frac{u_i^I - u_i}{\kappa}} \sum_{j=1}^{N} m_{ij}^R \left(\frac{V_j^I}{V_j^R}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(\frac{V_j^D}{V_j^R}\right)^{\beta\gamma_j^D} < \sum_{j=1}^{N} m_{ij}^R \left(\frac{V_j^I}{V_j^R}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)}$$

The inequality comes from the fact that $0 < u_i^I < u_i$ and $V_i^D < V_i^R$. Denote $\overline{\lambda} \equiv \max_i\{(V_i^I/V_i^R)^{\beta(1-\gamma_i^R-\gamma_i^D)}\}$ and $k \equiv \arg\max_i\{(V_i^I/V_i^R)^{\beta(1-\gamma_i^R-\gamma_i^D)}\}$ and note that $\beta(1-\gamma_k^D-\gamma_k^R) < 1$, the inequality above implies that

$$\overline{\lambda}^{\frac{1}{\beta(1-\gamma_k^R-\gamma_k^D)}} < \overline{\lambda} \Rightarrow \overline{\lambda} < 1 \Rightarrow V_i^I < V_i^R, \ \ \forall i.$$

Similarly, we divide the equation of type S group by the one of type R and get:

$$\frac{V_{i,t}^S}{V_i^R} = \sum_{j=1}^{N} m_{ij}^R \left(\frac{V_{j,t+1}^S}{V_j^R}\right)^{\beta(1-\alpha_{j,t+1})} \left(\frac{V_j^I}{V_j^R}\right)^{\beta\alpha_{j,t+1}} \leq \sum_{j=1}^{N} m_{ij}^R \left(\frac{V_{j,t+1}^S}{V_j^R}\right)^{\beta(1-\alpha_{j,t+1})}$$

Say, if there exists some $V_{i,t}^S/V_i^R > 1$, from the inequality above we know that there must exist some $V_{j,t+1}^S/V_j^R > 1$. Suppose $\overline{\xi}_t \equiv \max_i V_{i,t}^S/V_i^R > 1$, then the equation above implies that

$$\overline{\xi}_t \leq \sum_{j=1}^{N} m_{ij}^R \overline{\xi}_{t+1}^{\beta(1-\alpha_{j,t+1})} \leq \sum_{j=1}^{N} m_{ij}^R \overline{\xi}_{t+1}^{\beta} = \overline{\xi}_{t+1}^{\beta}$$

Therefore, $\overline{\xi}_{t+1} \geq \overline{\xi}_t^{1/\beta} > \overline{\xi}_t$ and $\{\overline{\xi}_{t+1}\}_{t=0}^{\infty}$ is a growing series greater than 1. However, at the steady state, as the infected individuals are all removed from the population and infection rates fall to zero, we have $V_{i,\infty}^S = V_{i,\infty}^R$ according to equation (4). We should have $\overline{\xi}_\infty = 1$. Contradiction, we have $V_{i,t}^S \leq V_{i,t}^R$, $\forall i$ and t. Finally, near DFE, the measure of type I is close to 0 and the whole population is susceptible, we have $V_{i,t}^S = V_i^R$ according to equation (4). Since $V_i^I < V_i^R$, we immediately have $V_i^I < V_{i,t}^S$ near the DFE.

A.3.5 Proof of Proposition 4

Taking logs on both sides of the expression for $m_{ij,t}^S$ in equation (5), we have

$$\ln(m_{ij,t}^S) = \beta(1-\alpha_{j,t+1})\ln(V_{j,t+1}^S) + \beta\alpha_{j,t+1}\ln(V_{j,t}^I) - \ln(\mu_{ij,t}) - ln(OV_{i,t}^S), \tag{A.3}$$

while $OV_{i,t}^S \equiv \sum_{k=1}^{N}\left(V_{k,t+1}^S\right)^{\beta(1-\alpha_{k,t+1})}\left(V_{k,t+1}^I\right)^{\beta\alpha_{k,t+1}}(\mu_{ik,t})^{-1}$ is the option value of region i for type S. Taking partial derivative of equation (A.3) with respect to $\alpha_{j,t+1}$, we have

$$\frac{\partial \ln(m_{ij,t}^S)}{\partial \alpha_{j,t+1}} = \beta\ln(\frac{V_{j,t}^I}{V_{j,t}^S}) - \frac{\partial \ln(OV_{i,t}^S)}{\partial \alpha_{j,t+1}}. \tag{A.4}$$

Denote $v_{k,t} = \left(V_{k,t+1}^S\right)^{\beta(1-\alpha_{k,t+1})}\left(V_{k,t+1}^I\right)^{\beta\alpha_{k,t+1}}(\mu_{ik,t})^{-1}$, then $OV_{i,t}^S = \sum_{k=1}^{N} v_{k,t}$ and

$$\begin{aligned}\frac{\partial \ln(OV_{i,t}^S)}{\partial \alpha_{j,t+1}} &= \frac{\partial \ln(OV_{i,t}^S)}{\partial OV_{i,t}^S}\frac{\partial OV_{i,t}^S}{\partial v_{j,t}}\frac{\partial v_{j,t}}{\partial ln(v_{j,t})}\frac{\partial \ln(v_{j,t})}{\partial \alpha_{j,t+1}} = \frac{v_{j,t}}{OV_{i,t}^S}\beta\ln(\frac{V_{j,t+1}^I}{V_{j,t+1}^S}) \\ &= \beta m_{ij,t}^S \ln(\frac{V_{j,t+1}^I}{V_{j,t+1}^S}).\end{aligned}$$

The last equality uses the expression for $m^S_{ij,t}$ in equation (5). Therefore, we have

$$\frac{\partial \ln(m^S_{ij,t})}{\partial \alpha_{j,t+1}} = \beta(1 - m^S_{ij,t}) \ln(\frac{V^I_{j,t+1}}{V^S_{j,t+1}}), \tag{A.5}$$

Since $V^I_j < V^S_j$ according to Lemma 1 and $m^S_{ij,t} \leq 1$ we have $\frac{\partial \ln(m^S_{ij,t})}{\partial \alpha_{j,t+1}} \leq 0$.

A.3.6 Proof of Proposition 5

We know that the law of motion for type I is given by $\boldsymbol{I_{t+1}} = \boldsymbol{F}\boldsymbol{I_t} + \boldsymbol{V}\boldsymbol{I_t}$, and the next-generation matrix $\boldsymbol{M}$ is $\boldsymbol{M} = \boldsymbol{F}(\mathbb{I} - \boldsymbol{V})^{-1}$. Given that $\chi_i \geq 0$ and $m^I_{ij} \geq 0$, $\forall i$ and $\forall j$, $\mathbf{F}$ is obviously non-negative. As $\mathbf{V}$ has a spectral radius of matrix less than 1, we have $(\mathbb{I} - \boldsymbol{V})^{-1} = \sum_{t=1}^{\infty} \mathbf{V}^t$. Since $\mathbf{V}$ is non-negative, as $\gamma_i < 1$ $\forall i$, $(\mathbb{I} - \boldsymbol{V})^{-1}$ must be non-negative as well. Therefore, $\mathbf{M}$ is non-negative. Then using the *Collatz - Wielandt* formula, the spectral radius of $\boldsymbol{M}$ satisfies:

$$\rho(\boldsymbol{M}) = \max_{\boldsymbol{x} \geq 0} \min_{i, x_i \neq 0} \frac{(\boldsymbol{x}\boldsymbol{M})_i}{x_i} = \min_{\boldsymbol{x} \geq 0} \max_{i, x_i \neq 0} \frac{(\boldsymbol{x}\boldsymbol{M})_i}{x_i}. \tag{A.6}$$

Denote $\boldsymbol{e} = [1, ..., 1]_{1 \times N}$, a row vector that all elements are one. According to the *Perron-Frobenius* theorem, $\boldsymbol{e}$ is the left eigenvector of the mobility matrix $\boldsymbol{m}$ with an eigenvalue of 1, given that $\boldsymbol{m}$ is a Markov matrix. So we have $\boldsymbol{e}\boldsymbol{m} = \boldsymbol{e}$. We can construct a row vector $\boldsymbol{\varepsilon}$ whose i^{th} element is $\varepsilon_i = \gamma_i / \chi_i$, i.e., $\boldsymbol{\varepsilon} = [\gamma_1/\chi_1, ..., \gamma_N/\chi_N]_{1 \times N}$. We next show

$$\boldsymbol{\varepsilon}\boldsymbol{M} = \boldsymbol{\varepsilon}\boldsymbol{F}(\mathbb{I} - \boldsymbol{V})^{-1} = \boldsymbol{e}.$$

First, let $Diag(\boldsymbol{x})$ be a square matrix with elements of vector $\boldsymbol{x}$ on the main diagonal, and $\boldsymbol{\chi} = \{\chi_1, ..., \chi_N\}$ and $\boldsymbol{\gamma} = \{\gamma_1, ..., \gamma_N\}$. Then $\boldsymbol{\varepsilon}\boldsymbol{F} = \boldsymbol{\varepsilon} Diag(\boldsymbol{\chi})\boldsymbol{m} = \boldsymbol{\gamma}\boldsymbol{m}$. Second, $\boldsymbol{e}(\mathbb{I} - \boldsymbol{V}) = \boldsymbol{e}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m}) = \boldsymbol{e} - \boldsymbol{e}Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m} = \boldsymbol{e}\boldsymbol{m} - (\boldsymbol{e}\boldsymbol{m} - \boldsymbol{e}Diag(\boldsymbol{\gamma})\boldsymbol{m}) = \boldsymbol{\gamma}\boldsymbol{m}$. Therefore, $\boldsymbol{\varepsilon}\boldsymbol{F} = \boldsymbol{e}(\mathbb{I} - \boldsymbol{V})$, or equivalently, $\boldsymbol{\varepsilon}\boldsymbol{F}(\mathbb{I} - \boldsymbol{V})^{-1} = \boldsymbol{e}$. Then if we replace $\boldsymbol{x}$ in equation (A.6) by the constructed vector $\boldsymbol{\varepsilon}$, we have

$$\min_i \frac{\chi_i}{\gamma_i} = \min_i \frac{1}{\varepsilon_i} \leq \rho(\boldsymbol{M}) \leq \max_i \frac{1}{\varepsilon_i} = \max_i \frac{\chi_i}{\gamma_i}, \tag{A.7}$$

so the global R_0 is bounded by the lowest and highest local R_0.

A.3.7 Proof of Proposition 6

Following the previous proof, we denote the next-generation matrix as $\boldsymbol{M} = \boldsymbol{F}(\mathbb{I} - \boldsymbol{V})^{-1} = Diag(\boldsymbol{\chi})\boldsymbol{m}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1}$, therefore,

$$\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}\chi_i} = \frac{\mathrm{d}Diag(\boldsymbol{\chi})}{\mathrm{d}\chi_i}\boldsymbol{m}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1} \geq \mathbf{0},$$

and

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{M}}{\mathrm{d}\gamma_i} &= -Diag(\boldsymbol{\chi})\boldsymbol{m}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1} \frac{(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})}{\mathrm{d}\gamma_i} (\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1} \\
&= Diag(\boldsymbol{\chi})\boldsymbol{m}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1} \frac{\mathrm{d}Diag(\boldsymbol{e} - \boldsymbol{\gamma})}{\mathrm{d}\gamma_i} \boldsymbol{m}(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1} \leq \mathbf{0},
\end{aligned}$$

using the fact that matrices $\boldsymbol{m}$ and $(\mathbb{I} - Diag(\boldsymbol{e} - \boldsymbol{\gamma})\boldsymbol{m})^{-1}$ are both nonnegative (see the previous proof). According to Theorem 2.1 (c) in Li and Schneider (2002), the spectral radius of an irreducible nonnegative matrix increases if any entry of it increases, vice versa if any entry of it decreases. Therefore, the global R_0 increases with the transmission rate χ_i and decreases with the removal rate $\gamma_i, \forall\, i$.

A.3.8 Proof of Proposition 7

We divide the value function for type S in equation (4) by its steady state and have

$$\begin{aligned}
\hat{V}^S_{i,t} = \frac{V^S_{i,t}}{V^S_{i,ss}} &= \frac{\sum_{j=1}^N \left(V^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(V^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} (\mu_{ij})^{-1}}{\sum_{j=1}^N \left(V^S_{j,ss}\right)^{\beta} (\mu_{ij})^{-1}} \\
&= \sum_{j=1}^N \frac{\left(V^S_{j,ss}\right)^{\beta} (\mu_{ij})^{-1}}{\sum_{k=1}^N \left(V^S_{k,ss}\right)^{\beta} (\mu_{ik})^{-1}} \frac{\left(V^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(V^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}}}{\left(V^S_{j,ss}\right)^{\beta}} \\
&= \sum_{j=1}^N m^S_{ij,ss} \left(\hat{V}^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(\hat{V}^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V^I_{j,ss}}{V^S_{j,ss}}\right)^{\beta\alpha_{j,t+1}}.
\end{aligned}$$

As for the mobility matrix, we do the same and get

$$\begin{aligned}
\frac{m^S_{ij,s}}{m^S_{ij,ss}} &= \frac{\left(V^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(V^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} (\mu_{ij})^{-1}}{\left(V^S_{j,ss}\right)^{\beta} (\mu_{ij})^{-1}} \frac{V^S_{i,ss}}{V^S_{i,t+1}} \\
&= \frac{\left(V^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(V^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}}}{\left(V^S_{j,ss}\right)^{\beta}} \frac{1}{\hat{V}^S_{i,t+1}} \\
&= \frac{\left(\hat{V}^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(\hat{V}^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V^I_{j,ss}}{V^S_{j,ss}}\right)^{\beta\alpha_{j,t+1}}}{\sum_{j=1}^N m^S_{ij,ss} \left(\hat{V}^S_{j,t+1}\right)^{\beta(1-\alpha_{j,t+1})} \left(\hat{V}^I_{j,t+1}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V^I_{j,ss}}{V^S_{j,ss}}\right)^{\beta\alpha_{j,t+1}}}.
\end{aligned}$$

Appendix B: Supplementary Materials

B.1 Data

Epidemiological Data of COVID-19 Daily COVID-19 data for each state from January 2020 till August 2021 are collected from the COVID Tracking Project. It reports information on the total number of confirmed cases, hospitalized cases, and deaths for each state. We aggregate the daily COVID-19 data to a biweekly level and build a panel dataset of US states.

Policy Data The implementation dates of various state-level COVID-19 policies in the US are documented by Raifman et al. (2020). This database tracks COVID-19 policies implemented by each state over time and covers a wide range of policies, including public place closure, physical distance closures, stay home policy, face-mask-wearing, and quarantine policy. We follow Hale et al. (2021) to construct a US state-level policy index. The index is a composite measure that combines different indicators of containment policies into a general index. We first re-scale each of the sub-policy by their maximum value to create a score between 0 and 100, with a missing value contributing 0. Next, these scores of different indicators are averaged to get the final composite measure for each state. The composite measure allows us to compare government responses to COVID-19 across states.

Mobility Data We use anonymous smartphone data from *PlaceIQ*, a location analytics firm, to estimate the interaction between the pandemic and human movement. *PlaceIQ* data tracks mobile phone users' location over time and describes smartphone devices "pinging" in a given geographic unit on a given day. Couture et al. (2021) used these data to compute a location exposure index "LEX", which measures the share of devices that have been pinged in a state other than their present location at least once in the past 14 days. The "LEX" measure is a state-to-state matrix at a daily frequency. The definition of "LEX", as stated in Couture et al. (2021), is "What fraction of active phones in geographic unit g_0 on day d have been active in geographic unit g at any point in the last 14 days?" We further aggregate the "LEX" from daily to a biweekly frequency to obtain real-time state-to-state mobility flows during the pandemic from January 2020 to August 2021. We aggregated the "LEX" data in two different ways: 1), we take the max of daily bilateral mobility flows between states during bi-week t; 2), we take the average of daily mobility flows between states during bi-week t.

The measurement we use to assess inter-state mobility flow is based on *PlaceIQ* and "LEX" data, and it includes all types of cross-state human activities, such as migration, commuting, and business travel. The biweekly state-to-state mobility flow is suitable to provide evidence of empirical relevance, as it is compatible with our theoretical framework, where agents make multi-period dynamic mobility decisions.

Table B.1 shows the summary statistics of the above-mentioned variables. Panel A is at the state-pair biweekly level, while Panel B is at the biweekly state level.

B.2 Reduced Form Analysis

Stylized Fact 1

We first investigate how inter-state human mobility affects the transmission of the virus using the following empirical specification:

$$ln(y_{i,t}) = a_0 + a_1 \sum_{j \neq i} m^0_{ji} \ln(cases_{j,t-1}) + a_2 StringencyIndex_{i,t-1} + d_t + S_i + \nu_{ij,t},$$

where $y_{i,t}$ is the number of newly confirmed COVID-19 cases in state i at a bi-week t. We use $\sum_{j \neq i} m^0_{ji} \ln(cases_{j,t-1})$ to measure the exposure of state i to the pandemic in other states due to inter-state human mobility. m^0_{ji} is the share of people moving from state j to state i in the initial period of our sample, which is the second bi-week of January 2020, before the outbreak of the pandemic. We use two measures to capture the initial mobility flow: the maximum daily mobility flows during a bi-week period, and the average daily mobility flows. $cases_{j,t-1}$ captures the severity of the pandemic at the origin state j, of which we use two measures: the number of new cases in period $t-1$, and the number of cumulative cases by period of $t-1$. Combined with the two measures of initial mobility m^0_{ji}, we have four approaches to measure a state i's exposure to the pandemic in other states.

Since containment policies are critical for containing the virus, we control for time-variant containment policies $Stringency\ Index_{i,t-1}$. A higher value in $StringencyIndex$ indicates stricter containment policies. The stringency of containment policies could be affected by the severity of the disease in a region, which creates potential endogeneity concerns. To deal with such concerns, we employ lagged variables. In addition, we control for time fixed effect d_t to capture national shocks and state fixed effect S_i for state-specific factors. Our sample includes all 50 US states at the cross-section across 40 bi-weeks, i.e., 2,040 bi-weekly state observations for January 2020 to August 2021.

Stylized Fact 2

We then adopt two empirical strategies to examine how infection risk shape inter-state human mobility. The first is a two-way fixed effect specification given by:

$$\begin{aligned} m_{ij,t} &= b_0 + b_1 Infection_{i,t-1} + b_2 Infection_{j,t-1} + b_3 MobilityRestriction_{i,t-1} \\ &+ b_4 MobilityRestriction_{j,t-1} + \sum_k \beta_k X_{ij,k} + d_t + O_i + D_j + \xi_{ij,t}, \end{aligned}$$

where $m_{ij,t}$ is defined as the share of people who move from state i to state j in a bi-week t relative to the people that are currently in state j. For independent variables, $Infection_{i,t-1}$ captures the lagged infection rate at the origin state, whereas $Infection_{j,t-1}$ represents the lagged infection rate at the destination state. $MobilityRestriction_{i,t-1}$ is a policy indicator reflecting restrictions from the origin state (i) on cross-region mobility. It equals 0 when the local government implements no restrictions, 1 when the government recommends people not to travel across regions, and 2 when the government imposes cross-region movement restrictions. $MobilityRestriction_{j,t-1}$ refers to the same set of policies

on movement restrictions but imposed by the destination state (j). $X_{ij,k}$ captures the time-invariant gravity variables, including distance, and shared common border between the origin and destination. The primary coefficients of interest are b_1 and b_2, which assess how bilateral state-to-state mobility responds to infection rates at origin and destination state, respectively. In addition, we include d_t to capture the time fixed effect. We control for origin state and destination state fixed effects (O_i, D_j). We also control for bilateral state pair fixed effects OD_{ij} for alternative specifications to capture all time-invariant unobservables at the state-pair level, such as pre-existing pairwise social and economic linkages. Since our sample includes biweekly state-pair observations that experience no human mobility flows during the bi-week. We estimate the model using Poisson Pseudo Maximum Likelihood (PPML), which handles zeros in the mobility data and potential heteroskedasticity (Silva and Tenreyro, 2006).

In the second specification, we investigate how the heterogeneity of COVID-19 across states affects state-level net mobility inflows. We expect states that suffered more from the pandemic to have smaller net inflows and estimate the following Difference-in-Differences (DiD) specification:

$$\begin{aligned} NetInflowShare_{j,t} &= c_0 + c_1 CumulativeInfectionRate_j \times AfterPandemic_t \\ &+ c_2 StringencyIndex_{j,t-1} + d_t + S_j + \nu_{j,t}, \end{aligned}$$

where $NetInflowShare_{j,t}$ is the net total mobility inflows from other states to a state j at period t divided by the population of j, $CumulativeInfectionRate_j$ is the cumulative infection rate at state j measured by the total number of confirmed cases divided by the population of j, and $AfterPandemic_t$ is dummy that equals 1 if t is after the COVID-19 pandemic started in the US. Again, we control for the stringency of closure policies as measured by $Stringency\ Index_{i,t-1}$, state fixed effect S_j and time fixed effect d_t.

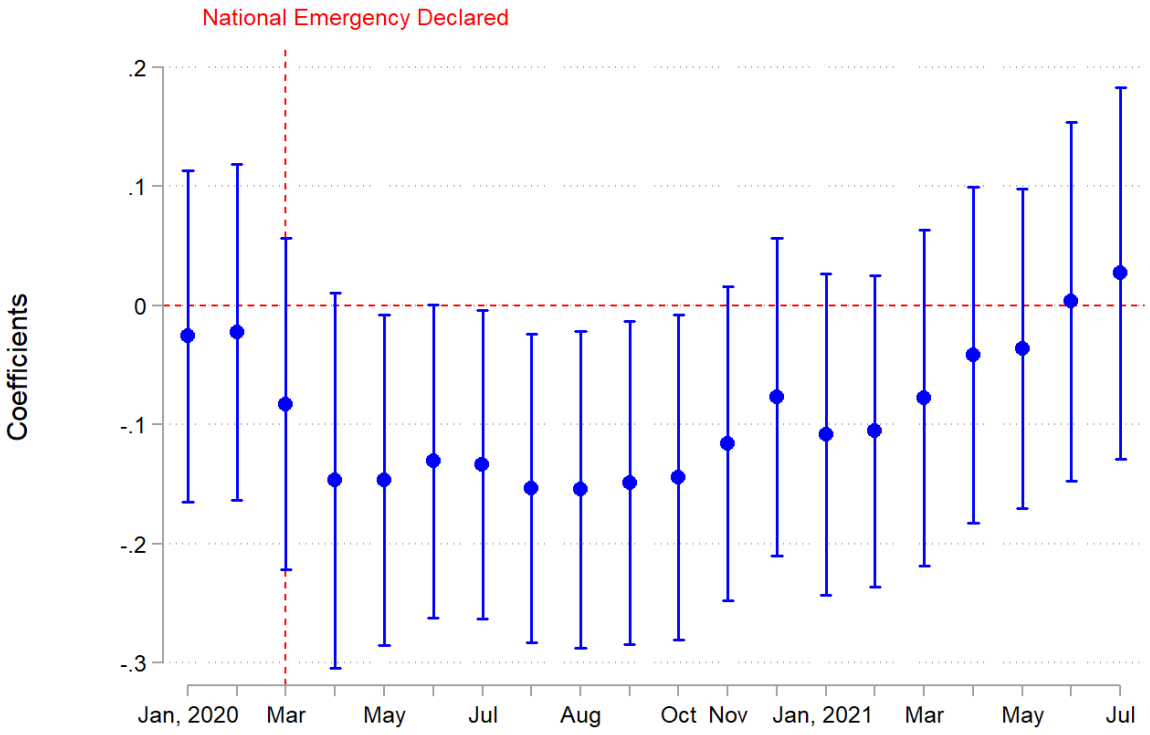
In addition, we augment the DiD specification above and allow for time-varying treatment effects.

$$\begin{aligned} NetInflowShare_{j,t} &= \alpha_0 + \sum_k \alpha_k CumulativeInfectionRate_j \times T_{k,t} \\ &+ c_2 StringencyIndex_{j,t-1} + d_t + S_j + \nu_{j,t}, \end{aligned}$$

while T_k is time dummy that equal 1 if period t falls within a specific year-month k. We note the state fixed effect would absorb $CumulativeInfectionRate_j$ while the time fixed effect would absorb T_k. The key coefficient of interest is α_k, which we plot in Figure B.1.

B.3 Additional Figures

Figure B.1: Effect of Cumulative Infection Rates on Net Mobility Inflows



Notes. This figure depicts the effect of cumulative infection rate on net mobility inflows of US states before and after the national emergency declaration in March 2020. The dependent variable is the net biweekly mobility inflows of a state, measured by the maximum daily net inflows within a bi-week. The markers represent the estimated coefficients of interaction terms between year-month dummies and state-level cumulative infection rate in August 2021. The capped spikes represent the associated 95% confidence intervals. The reference group is the final month of the data sample, August 2021. The dashed vertical line represents March 2020, when a national emergency was declared. The regression considers state and time fixed effects. Standard errors are clustered at the state and biweekly levels. We also consider measuring mobility inflows using the average daily net inflows within a bi-week and get a qualitatively similar figure.

B.4 Additional Tables

Table B.1: Summary Statistics

	Definition	Obs	Mean	Std.Dev	Min	Max
Panel A. Variables at the bilateral state pair - biweek level						
Max mobility	Max of bilateral mobility flow between states at bi-weekly frequency	102,000	0.012	0.036	0	0.766
Mean mobility	Average of bilateral mobility flow between states at bi-weekly frequency	102,000	0.011	0.034	0	0.742
Distance	bilateral distance between states	102,000	1,981.038	1,461.291	37.356	8.176.055
Border	Dummy variable of whether two states share a border	102,000	0.084	0.277	0	1
Panel B. Variables at the state - biweek level						
New infection rate	New infection rate (new cases /population) measured at state-biweek level	2,040	0.003	0.003	0	0.024
Mobility Restrictions	Restrictions on internal travel across regions/cities implemented by state government	2,040	1.116	0.537	0	2
Stringency index	US state-level containment policy index, ranged [0, 100]	2,040	52.380	18.300	0	93.52
Number of new cases	Number of new cases at state bi-weekly level	2,040	17,930	37,035	0	575,710
$\sum_{j\neq i} m_{ji}^{max} ln(newcases_{j,t})$	sum of ln (new cases) across other states, weighted by max mobility in the initial period	2,040	13.328	5.713	0	26.959
$\sum_{j\neq i} m_{ji}^{mean} ln(newcases_{j,t})$	sum of ln (new cases) across other states, weighted by mean mobility in the initial period	2,040	12.912	5.497	0	26.037
$\sum_{j\neq i} m_{ji}^{max} ln(cumulativecases_{j,t})$	sum of ln (cumulative cases) across other states, weighted by max mobility in the initial period	2,040	17.387	7.493	0.001	33.006
$\sum_{j\neq i} m_{ji}^{mean} ln(cumulativecases_{j,t})$	sum of ln (cumulative cases) across other states, weighted by mean mobility in the initial period	2,040	16.834	7.199	0.001	31.700
Share of net inflow (by max mobility)	percentage of people net inflow over population, aggregated by maximum of bilateral mobility	2,040	0.136	3.431	-10.204	11.802
Share of net inflow (by average mobility)	percentage of people net inflow over population, aggregated by average of bilateral mobility	2,040	0.121	3.156	-9.667	11.021
Cumulative infection rate	cumulative number of cases over population at the end of data period	2,040	10.976	2.583	3.545	14.905

Notes: Panel A shows the summary statistics of variables at the bilateral state pair - biweek level. The sample includes all 2,550 state pairs in the U.S., from January 2020 to August 2021, with 40 periods of bi-weeks. Panel B shows the summary statistics of variables at the state - biweek level. The sample includes all 51 states in the U.S., from January 2020 to August 2021, with 40 periods of bi-weeks.

Table B.2: Model Parameters for the Three-Region Economy

Parameters	Definition	Target/Source	Value
β	Discount rate	4% annual interest rate	0.9985
κ	Dispersion of preference shock	estimates from biweekly Baidu mobility data	8
γ^D	Mortality rate	US case fatality rate in year 2020	0.0174
γ^R	Recovery rate	average removal period of 18 days	0.76
χ_1	Transmission rate of Region 1	authors' choice	0.3
χ_2	Transmission rate of Region 2	authors' choice	0.4
χ_3	Transmission rate of Region 3	authors' choice	0.5
η	utility discount for group I ($u_i^I = \eta u_i$)	authors' choice	0.5
μ_{ij}	mobility costs from i to j, $i \neq j$	authors' choice	4
μ_{ii}	local mobility costs	authors' choice	1
u_i	instantaneous utility of type I and R	authors' choice	1
u^D	instantaneous utility of type D	authors' choice	0

B.5 The Normalized Hat Algebra Algorithm

First, we note that at the steady state, there is no type I agent, therefore, the infection rate $\alpha = 0$. Normalizing the value of the deceased as one, i.e., $V^D = 1$, the steady state of the model is characterized by:

$$\begin{aligned}
V_i^S &= \exp(\frac{u_i}{\kappa}) \sum_{j=1}^{N} \left(V_j^S\right)^{\beta} (\mu_{ij})^{-1}, \\
V_i^I &= \exp(\frac{u_i^I}{\kappa}) \sum_{j=1}^{N} \left(V_j^I\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(V_j^R\right)^{\beta\gamma_j^R} (\mu_{ij})^{-1}, \\
V_i^R &= \exp(\frac{u_i}{\kappa}) \sum_{j=1}^{N} \left(V_j^R\right)^{\beta} (\mu_{ij})^{-1},
\end{aligned}$$

$$\begin{aligned}
m_{ij}^S &= \frac{\left(V_j^S\right)^{\beta} (\mu_{ij})^{-1}}{\sum_{k=1}^{N} \left(V_k^S\right)^{\beta} (\mu_{ik})^{-1}}, \\
m_{ij}^I &= \frac{\left(V_j^I\right)^{\beta\left(1-\gamma_j^R-\gamma_j^D\right)} \left(V_j^R\right)^{\beta\gamma_j^R} (\mu_{ij})^{-1}}{\sum_{k=1}^{N} \left(V_k^I\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(V_k^R\right)^{\beta\gamma_j^R} (\mu_{ik})^{-1}}, \\
m_{ij}^R &= \frac{\left(V_j^R\right)^{\beta} (\mu_{ij})^{-1}}{\sum_{k=1}^{N} \left(V_k^R\right)^{\beta} (\mu_{ik})^{-1}}.
\end{aligned}$$

The size of type S and R in each region satisfies

$$S_i = \sum_{j=1}^{N} S_j m_{ji}^S, \; R_i = \sum_{j=1}^{N} R_j m_{ji}^R.$$

Then for given κ, β, $\{u_i, u_i^I, \gamma_i^R, \gamma_i^D\}_{i=1}^N$, $\{\mu_{ij}\}_{N\times N}$, we can solve for the steady-state value function and mobility rate of S, I, and R using the equations above.

Denote $\widehat{x}_t = x_t/x_{ss}$ where x_{ss} is the steady-state value of x, following Proposition 7, the model can be rewritten as:

$$\widehat{V}_{i,t}^S = \sum_{j=1}^N m_{ij}^S \left(\widehat{V}_{j,t+1}^S\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}_{j,t+1}^I\right)^{\beta\alpha_{j,t+1}} \left(\frac{V_j^I}{V_j^S}\right)^{\beta\alpha_{j,t+1}} \tag{E.1}$$

$$\widehat{V}_{i,t}^I = \sum_{j=1}^N m_{ij}^I \left(\widehat{V}_{j,t+1}^I\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(\widehat{V}_{j,t+1}^R\right)^{\beta\gamma_j^R} \tag{E.2}$$

$$\widehat{V}_{i,t}^R = \sum_{j=1}^N m_{ij}^R \left(\widehat{V}_{j,t+1}^R\right)^{\beta} \tag{E.3}$$

$$m_{ij,t}^S = \frac{m_{ij}^S \left(\widehat{V}_{j,t+1}^S\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}_{j,t+1}^I\right)^{\beta\alpha_{j,t+1}} \left(\frac{V_j^I}{V_j^S}\right)^{\beta\alpha_{j,t+1}}}{\sum_{k=1}^N m_{ik}^S \left(\widehat{V}_{k,t+1}^S\right)^{\beta\left(1-\alpha_{k,t+1}\right)} \left(\widehat{V}_{k,t+1}^I\right)^{\beta\alpha_{k,t+1}} \left(\frac{V_k^I}{V_k^S}\right)^{\beta\alpha_{k,t+1}}} \tag{E.4}$$

$$m_{ij,t}^I = \frac{m_{ij}^I \left(\widehat{V}_{j,t+1}^I\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(\widehat{V}_{j,t+1}^R\right)^{\beta\gamma_j^R}}{\sum_{k=1}^N m_{ik}^I \left(\widehat{V}_{k,t+1}^I\right)^{\beta(1-\gamma_k^R-\gamma_k^D)} \left(\widehat{V}_{k,t+1}^R\right)^{\beta\gamma_k^R}} \tag{E.5}$$

$$m_{ij,t}^R = \frac{m_{ij}^R \left(\widehat{V}_{j,t+1}^R\right)^{\beta}}{\sum_{k=1}^N m_{ik}^R \left(\widehat{V}_{k,t+1}^R\right)^{\beta}} \tag{E.6}$$

Algorithm

1. Solve the steady state V_i^S, V_i^I, V_i^R, m_{ij}^S, m_{ij}^I and m_{ij}^R.
2. For an exogenously given $\{L_0\}$, take a large enough T.
3. Take initial guess $\left\{\widehat{V}_t^{S(0)}, \widehat{V}_t^{I(0)}, \widehat{V}_t^{R(0)}\right\}_{t=1}^T$.
4. For each iteration of $\left\{\widehat{V}_t^{S(k)}, \widehat{V}_t^{I(k)}, \widehat{V}_t^{R(k)}\right\}_{t=1}^T$:
   (a) Given $\{S_{t-1}, I_{t-1}, R_{t-1}, D_{t-1}\}$, solve forward simultaneously for $\{S_t, I_t, R_t, D_t, \alpha_t\}$ from equations (6), (7), (8), (9), (E.4), (E.5), and (E.6).
   (b) Use equations (E.1), (E.2), (E.3) to solve backward to get $\left\{\widehat{V}^{S(k)\prime}, \widehat{V}_t^{I(k)\prime}, \widehat{V}_t^{R(k)\prime}\right\}_{t=1}^T$
5. If $\left\{\widehat{V}_t^{S(k)\prime}, \widehat{V}_t^{I(k)\prime}, \widehat{V}_t^{R(k)\prime}\right\}$ is close to $\left\{\widehat{V}_t^{S(k)}, \widehat{V}_t^{I(k)}, \widehat{V}_t^{R(k)}\right\}$, finish. Otherwise, set next guess $\left\{\widehat{V}_t^{S(k+1)}, \widehat{V}_t^{I(k+1)}, \widehat{V}_t^{R(k+1)}\right\} = \left\{\widehat{V}_t^{S(k)\prime}, \widehat{V}_t^{I(k)\prime}, \widehat{V}_t^{R(k)\prime}\right\}$.

B.6 Extension: a Model with Vaccinations

We now generalize the baseline model to incorporate vaccinations. We assume that people get vaccinated will be permanently immune, thus join group R. We further assume that region i randomly vaccinate a fraction δ_i of the susceptible population in each period, so a susceptible person from region i moving to region j will have a chance of δ_j directly transferring to type R. Now the Bellman equation for group S people becomes:

$$\begin{aligned} U_{i,t}^{S}(\boldsymbol{\varepsilon_{i,t}}) =& u_i + \max_j \left\{ \beta E_t \left[(1-\delta_j)(1-\alpha_{j,t+1}) U_{j,t+1}^{S}(\boldsymbol{\varepsilon_{j,t+1}}) + (1-\delta_j)\alpha_{j,t+1} U_{j,t+1}^{I}(\boldsymbol{\varepsilon_{j,t+1}}) \right.\right. \\ & \left.\left. + \delta_j U_{j,t+1}^{R}(\boldsymbol{\varepsilon_{j,t+1}}) \right] - \widetilde{\mu}_{ij} + \varepsilon_{ij,t} \right\}. \end{aligned}$$

As for the demographic dynamic, a fraction δ_i of type S will join type R:

$$\overline{S}_{i,t} = (1-\delta_i) \sum_{j=1}^{N} S_{j,t-1} m_{ji,t-1}^{S}, \ \overline{R}_{i,t} = \sum_{j=1}^{N} R_{j,t-1} m_{ji,t-1}^{R} + \delta_i \sum_{j=1}^{N} S_{j,t-1} m_{ji,t-1}^{S},$$

therefore, vaccination can slow down the spread of pandemics directly. If $\delta_i = 0$, the model collapses to our benchmark. We can show that the value function of type S satisfies

$$V_{i,t}^{S} = \exp\left(\frac{u_i}{\kappa}\right) \sum_{j=1}^{N} \left(V_{j,t+1}^{S}\right)^{\beta(1-\delta_j)(1-\alpha_{j,t+1})} \left(V_{j,t+1}^{I}\right)^{\beta(1-\delta_j)\alpha_{j,t+1}} \left(V_{j,t+1}^{R}\right)^{\beta\delta_j} (\mu_{ij})^{-1},$$

while the value functions of types I and R remain the same as Proposition 1. The mobility rate of type S is given by

$$m_{ij,t}^{S} = \frac{\left(V_{j,t+1}^{S}\right)^{\beta(1-\delta_j)(1-\alpha_{j,t+1})} \left(V_{j,t+1}^{I}\right)^{\beta(1-\delta_j)\alpha_{j,t+1}} \left(V_{j,t+1}^{R}\right)^{\beta\delta_j} (\mu_{ij})^{-1}}{\sum_{k=1}^{N} \left(V_{k,t+1}^{S}\right)^{\beta(1-\delta_k)(1-\alpha_{k,t+1})} \left(V_{k,t+1}^{I}\right)^{\beta(1-\delta_k)\alpha_{k,t+1}} \left(V_{k,t+1}^{R}\right)^{\beta\delta_k} (\mu_{ik})^{-1}},$$

while the mobility rates of types I and R remain the same as Proposition 2.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1859	Xavier Jaravel Danial Lashkari	Nonparametric measurement of long-run growth in consumer welfare
1858	Leonardo Bursztyn Jonathan Kolstad Aakaash Rao Pietro Tebaldi Noam Yuchtman	Political adverse selection
1857	Attila Lindner Balázs Muraközy Balázs Reizer Ragnhild Schreiner	Firm-level technological change and skill demand
1856	Jeremiah Dittmar Ralf R. Meisenzahl	The university, invention and industry: evidence from German history
1855	Donna Brown Jonathan Wadsworth	Accidents will happen: (de)regulation of health and safety legislation, workplace accidents and self employment
1854	Fabrizio Leone	Foreign ownership and robot adoption
1853	Andrew E. Clark Conchita D'Ambrosio Niccoló Gentile Alexandre Tkatchenko	What makes a satisfying life? Prediction and interpretation with machine-learning algorithms
1852	Xiang Ding Teresa C. Fort Stephen J. Redding Peter K. Schott	Structural change within versus across firms: evidence from the United States
1851	Christos Genakos Eleni Kyrkopoulou	Social policy gone bad educationally: unintended peer effects from transferred students
1850	Antonin Bergeaud Cyril Verluise	A new dataset to study a century of innovation in Europe and in the US

1849	Jo Blanden Mattias Doepke Jan Stuhler	Education inequality
1848	Martina Manara Tanner Regan	Ask a local: improving the public pricing of land titles in urban Tanzania
1847	Rebecca Freeman Kalina Manova Thomas Prayer Thomas Sampson	Unravelling deep integration: UK trade in the wake of Brexit
1846	Nicholas Bloom Scott W. Ohlmacher Cristina J. Tello-Trillo Melanie Wallskog	Pay, productivity and management
1845	Martin Gaynor Adam Sacarny Raffaella Sadun Chad Syverson Shruthi Venkatesh	The anatomy of a hospital system merger: the patient did not respond well to treatment
1844	Tomaz Teodorovicz Raffaella Sadun Andrew L. Kun Orit Shaer	How does working from home during Covid-19 affect what managers do? Evidence from time-use studies
1843	Giuseppe Berlingieri Frank Pisch	Managing export complexity: the role of service outsourcing
1842	Hites Ahir Nicholas Bloom Davide Furceri	The world uncertainty index
1841	Tomaz Teodorovicz Andrew L. Kun Raffaella Sadun Orit Shaer	Multitasking while driving: a time use study of commuting knowledge workers to access current and future uses