Trade, Migration, and Pandemics: Theory and Evidence from United States *

Xiao Chen† Hanwei Huang§ Jiandong Ju‡
Ruoyan Sun⋆ Jialiang Zhang¶

preliminary

Abstract

We study infectious diseases in a Susceptible-Infected-Recovered (SIR) model with interregional trade and endogenous migration. In our model, agents weigh the risk of infection and economic opportunities across regions when choosing the optimal location for work and life. They practice social distancing by avoiding high-risk areas. The model predicts that if migration cost is sufficiently low, the global basic reproductive number R_0 can be greater than one even if local R_0 is smaller than one. However, when global R_0 coincides with local R_0, further tightening of migration control becomes ineffective in reducing global R_0. We apply our theory to study the COVID-19 pandemic in the US. Reduced-form evidence based on smartphone user data indicates that the risk of infection at the destination reduces bilateral migration flows, and state-level containment policies effectively reduce the number of cases. A new method of "normalized hat algebra" is developed to solve and estimate the model. We estimate our model the US economy of 50 states. Counterfactual simulations based on the estimated model indicate that policies that altered cross-state mobility friction reduced the number of COVID-19 cases by 56.1 thousand while local containment policies reduced the number of cases by 31.6 million.

Key Words: SIR model, spatial economy, COVID-19 pandemic, endogenous migration, containment policy, normalized hat algebra

*For helpful discussions, we would like to thank Wenlan Luo. We are responsible for the remaining errors.

† School of International Trade and Economics, University of International Business and Economics. Email: chenxiao@uibe.edu.cn

§ Department of Economics and Finance, College of Business, City University of Hong Kong; Center for Economic Performance, London School of Economics. Email: huang.hanwei@cityu.edu.hk

‡ PBC School of Finance, Tsinghua University. Email: jujd@pbcsf.tsinghua.edu.cn

⋆ Department of Health Care Organization and Policy, School of Public Health, University of Alabama at Birmingham. Email: rysun@uab.edu

¶ PBC School of Finance, Tsinghua University. Email: zhangjl@pbcsf.tsinghua.edu.cn

1 Introduction

The 2019 novel coronavirus (COVID-19) pandemic has caused enormous loss of lives worldwide and brought the global economy into a deep recession. Depending on the vaccine's effectiveness, there is a real possibility that the pandemic could last until 2025 (Kissler et al., 2020). Containment policies have slowed down the spread of the virus (Hsiang et al., 2020; Maier and Brockmann, 2020) but also led to a collapse of global passenger and trade flow (Azevedo, 2020; ICAO, 2020). Reopening national borders for people and goods flows while safeguarding lives from the virus will be a paramount policy issue in the coming years. More broadly, due to rapid globalization and urbanization, infectious disease outbreaks have become more common in recent years, including SARS in 2003, H1N1 in 2009, MERS in 2012, Ebola in 2014, and Zika in 2015. How to cope with pandemics while reaping the benefits of globalization remains a long-term challenge for humanity.

In this paper, we study the dynamics of infectious diseases and the effect of containment policies using a multi-region Susceptible-Infected-Recovered (SIR) model with interregional trade and endogenous migration. In our model, agents of different types weigh the risk of infection and economic opportunities across regions when choosing the optimal location for work and life in each period. They automatically practice social distancing by avoiding high-risk areas. As migration is endogenous in our model, we can study the optimal mobility control. We provide the lower bound and upper bound for the range of migration restriction. When the migration cost is sufficiently low, the global basic reproductive number R_0 can be greater than one even if local R_0 is smaller than one, so the migration friction should be increased. On the other hand, when the migration friction is sufficiently high, such that global R_0 coincides with local R_0, further tightening of migration control becomes ineffective in reducing global R_0.

We apply our theory to study the real-life scenario of the US economy during the COVID-19 pandemic. First, we provide reduced-form evidence using smartphone user data, which suggests that the risk of infection at the destination reduces bilateral state-to-state migration flows. Later, when we simulate the model, we indeed find the Susceptible tend to move from regions with higher infection rates to those with lower infection rates. Second, we report that US state policies are effective at containing the pandemic and reducing the number of COVID-19 cases, while inter-state human mobility tends to worsen the pandemic and raise local COVID-19 cases. We confirm these findings in our simulated model by strengthening containment policies and show that stricter containment policies or higher inward migration costs reduce local infection. Interestingly, we find that unilateral imposition of migration restriction has "beggar-thy-neighbor" effect as it tends to increase the density of the infected population and the probability of getting infected in other regions. Without policy coordination, short-sighted governments in regions with outbreak tend to adopt loose containment policies while other regional governments would adopt policies as strict as possible.

We estimate our model to 50 US states and run counterfactual simulations to evaluate the effect of policies implemented by governments, including restrictions on cross-state movement and containment policies which reduced interaction of people within each state. We find both policies effectively reduced the number of accumulative infection during

March 2020 to July 2020. In aggregate, the observed changes in migration friction during this period reduced the total number of infection by about 56.1 thousand while the containment policies reduced infection by a total of 31.6 million. The stark difference in the effect of these two policies can be attributed to the fact that the US imposed minimum restriction of cross-state mobility during the migration. The effects were heterogeneous across states, in general, states with more stringent policies registered a bigger rise in the number of cases without these policies.

The first contribution of our paper to the literature of quantitative economic geography (Allen and Arkolakis, 2014; Redding and Rossi-Hansberg, 2017; Monte et al., 2018; Tombe and Zhu, 2019; Caliendo et al., 2019). The recent progress relies on models that deliver gravity equations for trade flows, migration, and commuting flows. These models are flexible and general enough to be brought to data and permit rich counterfactual analysis. We contribute to the literature by embedding an epidemiological model inside an economic geography model. In our model, agents' migration decisions are complicated by the fact that they can switch to other types in the next period while the probability of switching is endogenous, time-varying, and depends on the distribution of agents in each region.[1] This makes existing quantitative methods, including the "dynamic hat algebra" approach developed by Caliendo, Dvorkin, and Parro (2019), unsuitable to solve the model.[2] Instead, we make a methodological contribution by developing the approach "normalized hat algebra". We express the equilibrium conditions in terms of relative differences with respect to the steady state. It allows us to solve the dynamic model and conduct counterfactual simulations along the transition path without estimating all the fundamental parameters.[3]

The second contribution of our paper to the literature is in the field of epidemiological models (Kermack and McKendrick, 1927). We build a multi-region SIR model with endogenous migration across regions. Despite previous studies showing human mobility contributed to the spread of SARS-CoV-2 across Chinese cities (Fang et al., 2020) and large US cities (Glaeser et al., 2020) and our own evidence using smartphone user data capturing US inter-state mobility, migrations are typically taken as given or assumed away in existing SIR models with multiple regions (Bartlett, 1956; Muroya et al., 2013; Adda, 2016; Fajgelbaum et al. 2020; Antràs et al., 2020; Bisin and Moro, 2020). Moreover, such models do not capture the behavioural response in mobility to the virus which can reduce infection even without containment policies (Fang et al 2020).[4] Instead, agents in our model weigh the risk of infection and economic opportunities across regions and optimize the location for work and live in each period. Hence, they automatically follow social distancing measures and avoid going to high-risk regions. To verify this assumption, we

[1] We follow the literature (e.g., Acemoglu et al., 2020) to assume that the probability of infection depends on the density of the Susceptible and the Infected in the local population.

[2] Caliendo et al. (2017) assume that workers differ by skills. However, their skill type is time-invariant. Our method therefore can be applied to scenarios that agent can switch types, for example, from low-skill worker to high-skill worker.

[3] Since the differences are taken with respect to the steady state, compared to the "dynamic hat algebra" approach, the additional parameters that we need to estimate are the steady state variables.

[4] For example, Antràs et al. (2020) assume that buyers and sellers from different countries interact with each other only through international trade. Fajgelbaum et al. (2020) take the pre-pandemic commuting flows as given when studying the pandemic's optimal lockdown.

estimate how the state-to-state migration flow responded to disease conditions in each state during the pandemic. We find that people are less likely to move to a state with a higher infection rate of COVID-19, conditional on containment policies in each state and a number of controls.

We also contribute to the literature by investigating the effect of containment policies. By now, there is plenty of evidence that containment policies are effective in flattening the epidemic curve of COVID-19 (Fang et al., 2020; Hsiang et al., 2020; Maier and Brockmann, 2020). Our results confirm that US state policies were effective in reducing the number of COVID-19 cases. However, it is less clear whether these policies have been optimal or sub-optimal. There are normative studies on optimal containment policies along various dimensions, including quarantine and testing (Berger et al., 2020; Piguillem and Shi, 2020), lockdown (Alvarez et al., 2020; Acemoglu et al., 2020; Fajgelbaum et al., 2020), and general economic policies (Eichenbaum et al., 2020). However, to the best of our knowledge, no studies has provided optimal policy on human mobilities with endogenous migration. Our preliminary theoretical results show that if migration friction is sufficiently small, the global R_0 is greater than 1. Raising migration frictions tends to reduce R_0. However, once migration friction is high enough such that the global R_0 coincides with local R_0, further increases in migration friction would not reduce R_0. Therefore, a full lockdown would not be optimal, leaving scope for optimal mobility control.

The rest of the paper is arranged as follows. Section 2 introduces the data and presents the stylized facts which motivate our model. Section 3 sets up the model and discusses an analytical result. Section 4 introduces the normalized hat algebra. Section 5 presents our strategy to estimate the model to the US economy of 50 states and demonstrates the model channels with by simulating a three-region economy. Section 6 runs counterfactual simulations for the US economy to evaluate the effect of containment policies and migration friction. Section 7 concludes and discusses two theoretical extensions to incorporate vaccination and testing to the model.

2 Data and Stylized Facts

This section presents three stylized facts about the effect of containment policy and the role of inter-state migration on the spread of COVID-19 across US states. These new stylized facts motivate our theoretical model in the next section.

2.1 Data

Epidemiological Data of COVID-19

Daily COVID-19 data for each state since January 2020 till July 2020 are collected from the COVID Tracking Project. It reports information on the total number of confirmed cases, hospitalized cases, and deaths. We aggregate the daily data to bi-weekly frequency and build a panel dataset with newly confirmed cases of US states.

Policy Data

The implementation dates of COVID-19 US state policies are documented by Raifman et al. (2020). This database tracks COVID-19 policices 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. The corresponding detailed classification of policies is listed in Table 1.

We follow Hale et al. (2020) 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 rescale each of the sub-policy under their indicator by their maximum value to create a score between 0 and 100, with a missing value contributing 0. 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. It also avoids the problem that some policies are implemented closely, and the interpretation of individual policy responses could be over or mis-estimated.

Figure 1 plots the mean of the composite index for 50 states and DC over time. We find that US state containment policies peaked around April 2020 and were then gradually relaxed. Figure 2 shows the composite index's regional variation and trends every four weeks from March to July, 2020. As we can see, there are large variations across states in the strictness of containment policies.

Migration Data

To estimate the impact of the pandemic on migration, we use *anonymous* smartphone data from *PlaceIQ*, a location analytics firm, which track mobile phone users' location over time. Such data have been used to measure mobility during the pandemic (Glaeser, et al. 2020). We aggregate the Location Exposure Index (LEX) constructed by Couture et al. (2021) to a bi-weekly level to obtain real-time state-to-state migration flows during the pandemic.

Other Socio-Economic Statistics

US state population, wage, and income data in 2019 are from the US Census Bureau. Our source of state-level price data is the Regional Price Parities (RPPs) from the US Bureau of Economic Analysis. RPPs measures the differences in price levels across states and metropolitan areas for a given year. State-to-state trade flows are constructed from the 2017 Commodity Flow Survey (CFS), provided by the Bureau of Transportation Statistics and the US Census Bureau. State-to-state migration flows at annual frequency from 2014 to 2018 are from the IPMUS-USA database of Minnesota Population Center[5].

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

2.2 Stylized Facts

Exploiting the data collected, we use reduced-form econometric tools to establish three stylized facts about inter-state migration, containment policies, and the development of COVID-19 in the US. We are interested in how the risk of infection affected agents' mobility during the pandemic. We would also like to know whether containment policies

[5]Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0

implemented by state governments were effective in slowing down the pandemic and inter-state migration flows exacerbated infection.

Fact 1: People avoid high risk areas when migrating during the pandemic

Compared with existing studies on COVID-19, a key innovation of our model is that migration is endogenous and agents perceive the risk of infection in different regions when moving across regions. To support this assumption, we use migration data from *PlanceIQ* during the pandemic and study how migration responded to infection risk in different regions. To investigate whether people responded to the risk of infection, we estimate the following equation.

$$\begin{aligned} ln(m_{ij,t}) &= a_0 + a_1 Infection_{i,t-1} + a_2 Infection_{j,t-1} + a_3 Quarantine_{j,t-1} + ... \quad (1) \\ &+ \sum_k b_k X_{ij,k} + d_t + O_i + D_j + \xi_{ij,t} \end{aligned}$$

while $m_{ij,t}$ is the flow of people from state i to j at time t, $Infection_{i,t-1}$ captures the lagged infection rate at the origin state and $Infection_{j,t-1}$ the destination state. $Quarantine_{j,t-1}$ is the lagged quarantine policy imposed by the destination state. $X_{ij,k}$ captures the time variant gravity variables including distance, shared common border between the origin and destination. Finally, d_t captures the time fixed effect, O_i and D_j the origin and destination state fixed effects, respectively, and $\xi_{ij,t}$ the error term. Estimated results are shown in Table 2. We estimate the model using the method of Poisson Pseudo Maximum Likelihood (PPML), which deals with zeros in the migration data and potential heteroskedasticity (Silva and Tenreyro, 2006). Table 2 presents the estimated result. The gravity variables are consistent with the prior literature (Poncet, 2006; Tombe and Zhu, 2019): distances discourage inter-state migration while a shared border encourages it. We find that higher infection rate at the destination strongly and significantly reduced the state-to-state migration flow, while quarantine policy at the destination or infection rate at the origin did not seem to matter. It suggests that migration behavior during the pandemic changes with people spontaneously avoiding areas with high infection risks, after controlling for the effect of quarantine policies across states.

We next study the impact of containment policies and cross-state human mobility on the spread of the pandemic by estimating the following equation:

$$y_{i,t} = b_0 + b_1 ClosureIndex_{i,t} + b_3 \sum_{j \neq i} m^0_{ji} \ln(cases_{j,t-1}) + d_t + S_i + \nu_{ij,t}, \quad (2)$$

while $y_{i,t}$ is our measure of newly confirmed COVID-19 cases in state i at period t, and $ClosureIndex_{i,t}$ is an index capturing the stringency of closure measure as described in section 2.1, which ranges from 0 to 100. We measure the exposure of state i to the pandemic in other states due to inter-state mobility using $\sum_{j \neq i} m^0_{ji} \ln(cases_{j,t-1})$. m^0_{ji} is the share of people moving from state j to state i in the initial period of our sample, which is before the outbreak of the pandemic. $\ln(cases_{j,t-1})$ is the logarithm of the lagged number of cases in state j. We also control for time fixed effect d_t and state fixed effect S_i.

Table 3 presents the estimation results. In columns (1) - (4), we estimate the model using the method of OLS and using the logarithm of the number of new cases as the

outcome. To deal with zeros in the number of cases and heteroskedasticity (Silva and Tenreyro, 2006), columns (5) - (8) estimate the model using the method of PPML. As we can see, across all columns, the closure index which measures the stringency of containment policies had a significant and negative impact on the number of new cases. We can therefore summarize the finding as

Fact 2: Stringent containment policies are effective on reducing new infection.

The original PlaceIQ data measure cross state mobility in terms of share at daily frequency. Our analysis is at the bi-weekly frequency. We measure the exposure to pandemic in other states using either the average daily migration shares within initial two weeks or the maximum. As for the pandemic, we measure either in terms of the number of new cases or the cumulative number of cases. Across these measures, we find the positive and significant effect of exposure to the COVID-19 pandemic in other regions via cross-state migration on the local pandemic, which we summarize as the follows.

Fact 3: More migration from states of higher number new cases is associated with higher number of local cases.

3 Model

Motivated by the stylized facts, we set up a multi-region SIR model with interregional trade and endogenous migration. Agents make dynamic migration decisions taking into account the risk of infection and the probability of recovery in different regions. The model predicts how the dynamics of the pandemic and demography, trade, and the welfare depend on migration frictions and other model parameters.

3.1 Consumption, Production, and Trade

Consider an economy consisting of N regions. Agents derive utility from consuming a non-tradable composite final good given by

$$u(c_{i,t}) = \left[\int_0^1 q_{i,t}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}},$$

while $q_{i,t}(\omega)$ is the quantity of intermediate good variety ω consumed by agents in region i at period t, and σ is the elasticity of substitution between the differentiated varieties.

Intermediate goods are tradable. Following Eaton and Kortum (2002), we assume that agents purchase from the region offering the lowest price for each variety. Intermediate goods are produced with a linear production technology in labor input while productivities are random and drawn from the Fréchet distributions given by

$$z \sim F_i(z) = e^{-A_i z^{-\theta}},$$

while A_i captures the average productivity of region i and θ the dispersion of productivity. The cost of purchasing one unit of goods from region i in region j is given by

$$p_{ij}(\omega) = \frac{\tau_{ij} w_{i,t}}{z_i(\omega)},$$

where $w_{i,t}$ is the wage at region i at period t and τ_{ij} is the iceberg trade cost from region i to j. According to Eaton and Kortum (2002), the expenditure share of region j on goods from region i is given by

$$\pi_{ij,t} = \frac{A_i \left(\tau_{ij} w_{i,t}\right)^{-\theta}}{\Phi_{j,t}}, \tag{3}$$

while $\Phi_{j,t} = \sum_{k=1}^{N} A_k \left(\tau_{kj} w_{k,t}\right)^{-\theta}$. The aggregate price index in region j is given by

$$P_{j,t} = \gamma_1 \Phi_{j,t}^{-1/\theta}, \tag{4}$$

while $\gamma_1 \equiv \Gamma(\frac{\theta+1-\sigma}{\theta})$ is a constant.

3.2 Migration, Demographic and Disease Dynamics

We adopt the SIRD model (Kermack and McKendrick, 1927) and divide the total population into four groups, the Susceptible (S), Infected (I), the Recovered (R), and the Deceased (D).[6] For the purpose of empirical analysis, 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 group at the beginning of the period t in region i, and $S_{i,t}, I_{i,t}, R_{i,t}, D_{i,t}$ at the end of period t. The size of each group evolves over time due to both the pandemic and migration across regions.

For migration, we follow recent quantitative economic geography literature (Redding and Rossi-Hansberg, 2017; Caliendo et al. 2019; Tombe and Zhu, 2019) and assume that agents make their migration decision by maximizing the expected utility. To be specific, for the Susceptible group S in region i at period t, their problem can be characterized by the following Bellman equation,

$$U_{i,t}^{S}(\boldsymbol{\varepsilon_{i,t}}) = u_{i,t} + \max_{j}\{\beta E_t[(1-\alpha_{j,t+1})U_{j,t+1}^{S}(\boldsymbol{\varepsilon_{j,t+1}}) + \alpha_{j,t+1}U_{j,t+1}^{I}(\boldsymbol{\varepsilon_{j,t+1}})] - \widetilde{\mu}_{ij,t} + \varepsilon_{ij,t}\}, \tag{5}$$

while $u_{i,t}$ is the instantaneous utility given by $u_{i,t} = \frac{w_{i,t}}{P_{i,t}}$, and β is the discount rate. With a probability of $\alpha_{j,t+1}$, the agent will get infected in region j at period $t+1$ and join group I with a value function of $U_{j,t+1}^{I}$. With a probability of $1-\alpha_{j,t+1}$, the agent will remain uninfected and join group S with a value function of $U_{j,t+1}^{S}$. The migration cost from i to j for is $\widetilde{\mu}_{ij,t}$ in period t. On the top of that, agents currently in region i receive a vector of preference shock $\boldsymbol{\varepsilon_{i,t}}$ at the end of each period with each element $\varepsilon_{ij,t}$ following an *i.i.d.* Gumbel distribution

$$\Pr\{\varepsilon_{ij,t} \leq \varepsilon\} = \exp\left\{-\exp\left\{-\varepsilon/\kappa - \gamma^{Euler}\right\}\right\}, \tag{6}$$

while κ is a parameter that controls the dispersion of the shock while γ^{Euler} is the Euler constant ensuring that the shock has zero mean. Similarly, for groups I and R, their value functions are given by

$$U_{i,t}^{I}(\boldsymbol{\varepsilon_{i,t}}) = u_{i,t}^{I} + \max_{j}\{\beta E_t[(1-\gamma_j^R-\gamma_j^D)U_{j,t+1}^{I}(\boldsymbol{\varepsilon_{j,t+1}}) + \gamma_j^R U_{j,t+1}^{R}(\boldsymbol{\varepsilon_{j,t+1}}) + \gamma_j^D U^D] - \widetilde{\mu}_{ij,t} + \varepsilon_{ij,t}\}, \tag{7}$$

[6] If the death rate is zero, the setup is the same as the standard SIR model.

and

$$U_{i,t}^{R}(\boldsymbol{\varepsilon_{i,t}}) = u_{i,t} + \max_{j}\{\beta E_t[U_{j,t+1}^{R}(\boldsymbol{\varepsilon_{j,t+1}})] - \widetilde{\mu}_{ij,t} + \varepsilon_{ij,t}\}, \tag{8}$$

while γ_j^R and γ_j^D, are the recovered and death rates in region j, respectively. For group I, their instantaneous utility $u_{i,t}^I = \frac{\eta_{i,t} w_{i,t}}{P_{i,t}}$ depends on the effective labor supply rate $\eta_{i,t} \in [0,1]$ which depends on the containment policy. Following Fajgelbaum et al. (2020), we assume that a stricter containment policy will urge more infected people to telecommute, thus reduce the effective wage for those infected. For group D, the Deceased, their value function is given by U^D which is a constant that we normalize as 0.[7] Solving problems (5), (7), and (8), we find that values of residents can be recursively defined as the present option values of future living choices portfolio:

$$\begin{aligned}
V_{i,t}^{S} &= \exp\left(\frac{u_{i,t}}{\kappa}\right) \sum_{k=1}^{N} \left(V_{k,t+1}^{S}\right)^{\beta(1-\alpha_{k,t+1})} \left(V_{k,t+1}^{I}\right)^{\beta\alpha_{k,t+1}} \left(\mu_{ik,t}\right)^{-1}, \\
V_{i,t}^{I} &= \exp\left(\frac{u_{i,t}^{I}}{\kappa}\right) \sum_{k=1}^{N} \left(V_{k,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(V_{k,t+1}^{R}\right)^{\beta\gamma_j^R} \left(V^D\right)^{\gamma_k^D} \left(\mu_{ik,t}\right)^{-1}, \\
V_{i,t}^{R} &= \exp\left(\frac{u_{i,t}}{\kappa}\right) \sum\nolimits_{k=1}^{N} \left(V_{k,t+1}^{R}\right)^{\beta} \left(\mu_{ik,t}\right)^{-1},
\end{aligned} \tag{9}$$

and the migration rates are given by:

$$\begin{aligned}
m_{ij,t}^{S} &= \frac{\left(V_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})} \left(V_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}} \left(\mu_{ij,t}\right)^{-1}}{\sum_{k=1}^{N} \left(V_{k,t+1}^{S}\right)^{\beta(1-\alpha_{k,t+1})} \left(V_{k,t+1}^{I}\right)^{\beta\alpha_{k,t+1}} \left(\mu_{ik,t}\right)^{-1}}, \\
m_{ij,t}^{I} &= \frac{\left(V_{j,t+1}^{I}\right)^{\beta\left(1-\gamma_j^R-\gamma_j^D\right)} \left(V_{j,t+1}^{R}\right)^{\beta\gamma_j^R} \left(V^D\right)^{\gamma_j^D} \left(\mu_{ij,t}\right)^{-1}}{\sum_{k=1}^{N} \left(V_{k,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(V_{k,t+1}^{R}\right)^{\beta\gamma_j^R} \left(V^D\right)^{\gamma_k^D} \left(\mu_{ik,t}\right)^{-1}}, \\
m_{ij,t}^{R} &= \frac{\left(V_{j,t+1}^{R}\right)^{\beta} \left(\mu_{ij,t}\right)^{-1}}{\sum_{k=1}^{N} \left(V_{k,t+1}^{R}\right)^{\beta} \left(\mu_{ik,t}\right)^{-1}},
\end{aligned} \tag{10}$$

while $V_{i,t}^{g} = \exp(E_t(U_{i,t}^{g})/\kappa), g \in \{S, I, R\}$, and $\mu_{ij,t} = \exp(\tilde{\mu}_{ij,t}/\kappa)$. Compared with existing models on migration (Redding and Rossi-Hansberg, 2017; Caliendo et al. 2019; Tombe and Zhu, 2019), there are a few major differences. First, the migration rates are heterogeneous across different of groups of agents. Second, other than migration frictions, disease parameters (infection, recovery, and death rates), also affect the migration probability directly. Lastly, the agents take into account the fact that they might switch to other types in following periods. It is worth mentioning that though agents are uncertain about whether they will get infected next period, the aggregate risks of infection $\alpha_{j,t+1}$ are endogenously determined by collective migration decisions, thus will be considered as given from the perspective of each individual when making location choices.

[7] We adopt the linear utility function of $u(x) = x$ following Artuç et al (2010). If we follow Caliendo et al (2019) to assume $u(x) = \ln x$, U^D by definition should be the utility of zero consumption stream, which would be $-\infty$ and collapse the value functions to zero.

As standard in SIR models, the demographic dynamics within period due to the disease are given by

$$\begin{aligned} S_{i,t} &= \overline{S}_{i,t} - T_{i,t},\ I_{i,t} = T_{i,t} + \left(1 - \gamma_i^R - \gamma_i^D\right) \overline{I}_{i,t}, \\ R_{i,t} &= \overline{R}_{i,t} + \gamma_i^R \overline{I}_{i,t},\ D_{i,t} = \overline{D}_{i,t} + \gamma_i^D \overline{I}_{i,t}, \end{aligned} \tag{11}$$

where

$$T_{i,t} = M_i \left(\frac{\overline{S}_{i,t}}{L_{i,t}}, \frac{\eta_{i,t} \overline{I}_{i,t}}{L_{i,t}} \right) L_{i,t} \tag{12}$$

is the number of newly infected people in region i at time t and $M_i(.,.)$ is the matching function between groups S and I (Acemoglu et al., 2020). Then the infection rate is given by $\alpha_{it} \equiv \frac{T_{i,t}}{\overline{S}_{i,t}}$. To simplify matters, we adopt the functional form assuming that $M_i \left(\frac{\overline{S}_{i,t}}{L_{i,t}}, \frac{\eta_{i,t}\overline{I}_{i,t}}{L_{i,t}} \right) = \min \left\{ \frac{\overline{S}_{i,t}}{L_{i,t}}, \chi_i \frac{\overline{S}_{i,t}}{L_{i,t}} \left(\frac{\eta_{i,t}\overline{I}_{i,t}}{L_{i,t}} \right)^{\phi} \right\}$, while χ_i is a parameter that can be shaped by regional features influencing spread of virus like area and climate, $\eta_{i,t}$ which is region and time specific and depends on and crucially local containment policies, and ϕ is the matching elasticity regarding infected population.[8] Similar to other dynamic models of migration, the demographic dynamics across periods due to migration are given by

$$\begin{aligned} \overline{S}_{i,t} &= \sum_{j=1}^{N} S_{j,t-1} m^S_{ji,t-1},\ \overline{I}_{i,t} = \sum_{j=1}^{N} I_{j,t-1} m^I_{ji,t-1}, \\ \overline{R}_{i,t} &= \sum_{j=1}^{N} R_{j,t-1} m^R_{ji,t-1},\ \overline{D}_{i,t} = \overline{D}_{i,t-1}. \end{aligned} \tag{13}$$

To sum up, the timing of the event in the model is given by Figure 3.

Finally, to close the model, the goods market condition for each market is give by

$$w_{i,t} \left(S_{i,t} + R_{i,t} + \eta_{i,t} I_{i,t} \right) = \sum_{j=1}^{N} \pi_{ij,t} w_{j,t} \left(S_{j,t} + R_{j,t} + \eta_{j,t} I_{j,t} \right). \tag{14}$$

3.3 Equilibrium

Given our assumptions, the time-varying fundamentals of the economy are the bilateral migration costs $\mu_t = \{\mu_{ij,t}\}_{i=1,j=1}^{N,N}$ and containment policies $\eta_t = \{\eta_{i,t}\}_{i=1}^{N}$, while the constant fundamentals include productivities of each region $A = \{A_i\}_{i=1}^{N}$, bilateral trade costs $\tau = \{\tau_{ij}\}_{i=1,j=1}^{N,N}$, recovery rate $\gamma^R = \{\gamma_i^R\}_{i=1}^{N}$, death rate $\gamma^D = \{\gamma_i^D\}_{i=1}^{N}$, and the scale parameter of the matching function $\chi = \{\chi_i\}_{i=1}^{N}$. We denote the time-varying fundamentals as $\Omega_t \equiv \{\mu_t,\ \eta_t\}$ and the constant fundamentals as $\overline{\Omega} \equiv \{A,\ \tau,\ \gamma^R,\ \gamma^D,\ \chi\}$.

[8] Here we add a min operator in the definition of the matching function to allow for the scenario that the pandemic explodes and everyone susceptible is infected. Then all susceptible people will get infected at some period, which cannot be simulated if traditional economic matching function of $M(x,y) = x^\theta y^{1-\theta}$ or epidemiological matching function of $M(x,y) = axy$ is used.

The other parameters which are assumed to be constant across regions and time are: the trade elasticity (θ), the migration elasticity (κ), the discount rate (β), and the elasticity of substitution between differentiated varieties in final consumption (σ).

Our first goal is to find the equilibrium wage $w_t = \{w_{i,t}\}_{i=1}^{N}$ and allocation of expenditure $\{\pi_{ij,t}\}_{i=1,j=1}^{N,N}$ for period t, given the fundamentals of the economy Ω_t and $\overline{\Omega}$, and the distribution of each of type of agents across regions $G_t \equiv \{S_{i,t}, I_{i,t}, R_{i,t}, D_{i,t}\}_{i=1}^{N}$, which will be referred as the *temporary equilibrium*.

Definition 1 *Given (Ω_t, $\overline{\Omega}$, and G_t), a temporary equilibrium is a vector of wages $w(\Omega_t, \overline{\Omega}, G_t)$ satisfying equations (3), (4) and (14).*

The temporary equilibrium corresponds to the equilibrium of a static multi-regional trade model.[9] We then define the *sequential competitive equilibrium* for a given path of $\{\Omega_t\}_{t=0}^{\infty}$ and Ω. If we denote $m_t = \{m_{ij,t}^S,\ m_{ij,t}^I,\ m_{ij,t}^R\}_{i=1,j=1}^{N,N}$ and $V_t = \{V_{i,t}^S,\ V_{i,t}^I,\ V_{i,t}^R\}_{i=1}^{N}$ as the migration share and lifetime utility of each group, the definition of the sequential competitive equilibrium is

Definition 2 *Given ($\{\Omega_t\}_{t=0}^{\infty}$, $\overline{\Omega}$, and G_0), the sequential competitive equilibrium is a sequence of $\{G_t,\ m_t,\ V_t,\ w(\Omega_t, \overline{\Omega}, G_t)\}_{t=0}^{\infty}$ satisfying equations (10), (11), (12), (12), and (14).*

Then we are ready to define the *stationary equilibrium*.

Definition 3 *A stationary equilibrium is a sequential competitive equilibrium such that $\{G_t,\ m_t,\ V_t,\ w(\Omega_t, \overline{\Omega}, G_t)\}_{t=0}^{\infty}$ are constant over time.*

(sth is wrong with the numbering the equations. Should be 9-13)

3.4 Migration and R_0

Only under certain restrictions, SIR models have analytical solutions (Harko et al., 2014). Embedding such models in an economic geography model with interregional trade and endogenous migration by multiple groups of agents makes it even more complicated. Nevertheless, we can provide analytical results about the R_0 to demonstrate the interaction between migration and the pandemic. If we rewrite the law of motion for group I given in Eq. (12) and (14) [10]:

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

Following Diekmann et al. (1990), we vectorize the equation above as

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

[9] Redding (2016) provides a proof of existence and uniqueness to a similar problem.

[10] Now we assume that $\phi = 1$ and the pandemics will not explode, as usually people do when calculating R_0

while

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

$$\mathbf{V} = \begin{bmatrix} \gamma^D_1+\gamma^R_1 & & & \\ & \gamma^D_2+\gamma^R_2 & & \\ & & \ddots & \\ & & & \gamma^D_N+\gamma^R_N \end{bmatrix} \begin{bmatrix} m^I_{11,ss} & m^I_{21,ss} & \cdots & m^I_{N1,ss} \\ m^I_{12,ss} & m^I_{22,ss} & \cdots & m^I_{N2,ss} \\ \vdots & \vdots & \ddots & \vdots \\ m^I_{1N,ss} & m^I_{2N,ss} & \cdots & m^I_{NN,ss} \end{bmatrix}.$$

and $m^I_{ji,ss}$ is the steady state value of $m^I_{ji,t-1}$. The first term refers to all the historical infected people, including the incumbent and the newly infected. And the second refers to those who either die or recover, not infected any longer. Then R_0 can be calculated as the spectral radius of the following matrix

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

Intuitionally, $\mathbf{F} - \mathbf{I}$ indicates the newly infected while $\mathbf{V}^{-1}$ is for the expected duration of infection. So by definition, matrix M represents the number of people an infected person is expected to contaminate during his illness. In the symmetric case, $\chi_i = \chi$, $\gamma^R_i + \gamma^D_i = \gamma$, the migration matrix for group I is

$$m^I = \begin{bmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{bmatrix},$$

while a is the probability of staying in local region and b the probability of migrating to other regions. Matrix m^I has two eigenvalues, one that equals $a - b$ and another that equals 1. Then matrix $\mathbf{M}$ can be diagonalized as

$$\mathbf{M} = P^{-1} \begin{bmatrix} \frac{\chi+1-\frac{1}{a-b}}{\gamma} & & & \\ & \frac{\chi+1-\frac{1}{a-b}}{\gamma} & & \\ & & \ddots & \\ & & & \frac{\chi}{\gamma} \end{bmatrix} P, \tag{15}$$

which has two eigenvalues given by χ/γ and $(\chi + 1 - 1/(a-b))/\gamma$. We find that even if the local $R_0 = \chi/\gamma < 1$, the global R_0 is greater than 1 as long as $b \geq \frac{\chi+\gamma}{1+\chi+\gamma}\frac{1}{N}$. If $\frac{\chi+\gamma}{1+\chi+\gamma}\frac{1}{N} > b > \frac{2\chi}{1+2\chi}\frac{1}{N}$, then the global $R_0 = (1/(a-b) - 1 - \chi)/\gamma$ which decreases with b. Finally, as long as $b \leq \frac{2\chi}{1+2\chi}\frac{1}{N}$, the global R_0 coincides with the local R_0, which does not depend on migration costs. We summarize the results in the following proposition.

Proposition 4 *In a symmetric equilibrium,*
(a) if migration cost is sufficiently small such that $b \geq \frac{\chi+\gamma}{1+\chi+\gamma}\frac{1}{N}$, i.e., the probability of

migrating to other region is sufficiently large, then the global basic reproductive number $R_0 > 1$*, even if the local* $R_0 < 1$*;*
(b) if migration cost is sufficiently large such that $b \leq \frac{2\chi}{1+2\chi}\frac{1}{N}$*, i.e, the probability of migrating to other region is sufficiently low, then the global basic reproductive number* $R_0 > 1$ *coincides with the local* R_0*.*

According to result (a), the pandemic may be waning in a region but waxing globally. One way to put down the pandemic globally is to reduce migration across regions. However, according to result (b), beyond a certain threshold, further increasing the migration cost becomes ineffective as the global R_0 is determined by the local R_0.

4 Normalized Hat Algebra

The "Hat Algebra" approach has been widely used in quantitative trade and economic geography models (Dekle et al, 2007; Costinot and Rodriguez-Clare, 2014; Redding and Rossi-Hansberg, 2017). This approach allows researchers to fit models to observed data without estimating model fundamentals such trade and migration costs but still permits rich counterfactual analysis. However, it is only useful in static or steady models. Caliendo et al. (2019) extended this approach and developed the method of "Dynamic Hat Algebra" to solve models with intrinsic migration dynamics.[11] They show how to derive the transition path and conduct counterfactual analysis along the transition path in a class of models.

Unfortunately, the dynamic hat algebra cannot be applied to our setting. Firstly, the dynamic hat algebra relies on the logarithm preference while we follow Artuç et al. (2010) and assume an linear preference.[12] Secondly, due to the pandemic, group S transits to group I over time and the transition probability is time-variant. This makes the dynamic hat algebra impractical even if we adopt the logarithm preference. To see that, we can rewrite the growth rate of group S's value function in Equation (10) following Caliendo et al. (2019):

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

while $\dot{V}^g_{i,t+1} = \frac{V^g_{i,t+2}}{V^g_{i,t+1}}$, $g \in \{S, I\}$. As we can see, the term $(V^S_{k,t+1}/V^I_{k,t+1})^{\beta(\alpha_{k,t+1}-\alpha_{k,t+2})}$ emerges any time infection rates fluctuates: $\alpha_{k,t+1} \neq \alpha_{k,t+2}$. Therefore, we cannot rewrite the equation purely in terms of the growth rate of the value function and proceed to solve the model in terms of changes as Caliendo et al. (2019).

To solve this problem, we instead develop a "Normalized Hat Algebra" approach by normalizing all value function to their corresponding steady-state values. If we denote $\widehat{x}_t = x_t/x_{ss}$, while x_{ss} is the steady state value of variable x. Then if we define

[11] In new economic geography models such as Fujita et al (1999), migration typically follows an exogenous process which depends on the wage difference between regions.

[12] This is because the Deceased group have streams of zero consumption in our model. Under the logarithm preference, the utility of the Deceased will be minus infinity, which collapses all value function to zero as long as the probability of death is positive.

Definition 5 *A converging sequence is such that* $lim_{t\to\infty}\Phi_t = \Phi_{ss}$,

we have the following result

Proposition 6 *Given the initial allocation of the economy* (G_0, π_0) *and the steady state of economy* $(\Phi_{SS}, \overline{\Phi})$*, and an anticipated sequence of fundamentals* $\{\Phi_t\}_{t=0}^{\infty}$*, the solution to the sequential competitive equilibrium solves equations (14), (12) , (12), together with the following nonlinear equations*

$$\begin{aligned}
\widehat{V}_{i,t}^{S} &= \exp(\frac{u_{i,t}(w_t) - u_{i,ss}(w_{ss})}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^{S} \left(\widehat{V}_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}} \left(\frac{V_{j,ss}^{I}}{V_{j,ss}^{S}}\right)^{\beta\alpha_{j,t+1}} \\
\widehat{V}_{i,t}^{I} &= \exp(\frac{u_{i,t}^{I}(w_t) - u_{i,ss}^{I}(w_{ss})}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^{I} \left(\widehat{V}_{j,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(\widehat{V}_{j,t+1}^{R}\right)^{\beta\gamma_j^R} \\
\widehat{V}_{i,t}^{R} &= \exp(\frac{u_{i,t}(w_t) - u_{i,ss}(w_{ss})}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^{R} \left(\widehat{V}_{j,t+1}^{R}\right)^{\beta}
\end{aligned}$$

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

for all regions i *and* j *at each period* t*, while* w_t *is the solution to the temporary equilibrium given* $\{\Phi_t, G_t\}$ *and* $w_{ss} = lim_{t\to\infty} w_t$.

Therefore, given the initial conditions, steady state values and anticipated sequence of converging time varying fundamentals, we are able to solve the transitional dynamics. The key to get there results is that at the steady state, the infection rate in each region should reach zero. In other words, economic activities achieve stability either all the susceptible get infected or pandemic just gradually fades away as all infection are eliminated. In both cases, steady state values for the Recovered and the Susceptible $V_{i,ss}^{R}$, $V_{i,ss}^{S}$ can be solved first, and then the infected $V_{i,ss}^{I}$ from Equation (10).

Compared with dynamic hat algebra, our approach requires the value of steady states. But counterfactual analysis can be conducted directly by inserting counterfactual parameters into , rather than develop new equation systems as in dynamic hat algebra. The detailed algorithm to use the normalized hat algebra is in the Appendix A.1.

5 Quantifying the Model

In this section, we first estimate our model to the spatial economy of 50 US states. We use data from the pre-pandemic US economy to estimate the steady-state trade and migration parameters (trade elasticity σ, migration elasticity κ, trade and migration costs). For the matching function between groups S and I, $M(.,.)$, we estimate it using epidemiological data of COVID-19 and containment policies. Then we apply our normalized hat algebra approach to estimate time-varying migration friction during the pandemic by fitting the model to observed the dynamics during the pandemic.

We then simulate an artificial three-region economy to demonstrate the mechanism of our model, before running full counterfactuals in the section for the US economy.

5.1 Estimation

We categorize the model parameters into four groups: Preference, Trade, Migration, and Pandemics, which are calibrated using separate data sets.

Preference

To ease our quantitative analysis and calibration, we set the following parameters by adopting common values from the literature. We set the value of $\beta = 0.96$, the discount factor, to match the 4% yearly interest rate. σ captures the demand elasticity of consumption goods. We set $\sigma = 2$.[13]

Trade

We estimate trade elasticity and trade costs using the 2017 Commodity Flow Survey by aggregating the data to state level. According to Equation (3), if we normalize the bilateral trade share with the home trade share in the destination, we have

$$\ln\left(\frac{\pi_{ij,t}}{\pi_{jj,t}}\right) = -\theta \ln \tau_{ij} + S_i^T - S_j^T - \theta \ln \frac{w_{i,t}}{w_{j,t}}, \tag{16}$$

while $S_i^T \equiv \ln A_i$ is the state fixed effect indicating technology and the coefficient before the income ratio of origin and destination states $\frac{w_{i,t}}{w_{j,t}}$ is the trade elasticity. Following Waugh (2010), we decompose trade friction into distance, contiguity and importer fixed effect to allow for asymmetric trade costs between states

$$\ln \tau_{ij} = \beta_0^T + \beta_1^T \ln dist_{ij} + \beta_2^T cont_{ij} + \delta_j^T + \varepsilon_{ij}^T,$$

while $ln(dist)_{ij}$ is the log of distance between two states, $cont_{ij}$ is a dummy indicating contiguous or not, δ_i^T is the importer fixed effect , and ε_{ij}^T the error term. Without loss of generality, we can normalize that $\sum_{i=1}^N S_i^T = \sum_{i=1}^N \delta_i^T = 0$.

Table A.1 presents our estimates of θ, which varies from about 2 to 3. We will set $\theta = 2.961$ in our benchmark simulation according to the estimation result in column (2).[14]

[13] To ensure that the price index in equation (4) is well defined, we need $\theta > \sigma - 1$. Later, we will do robustness checks on the value of σ.

[14] We later explore the sensitivity of our results using alternative values of θ.

After obtaining the value of trade elasticity, we infer the set of trade costs between any two states using the following equation

$$\widehat{\tau}_{ij} \quad = \quad \exp\left(\widehat{\beta}_0^T + \widehat{\beta}_1^T \ln dist_{ij} + \widehat{\beta}_2^T cont_{ij} + \widehat{\delta}_j^T\right),$$

and plot the distribution of these estimated trade costs in Appendix Figure A1.

Migration

We next estimate the migration elasticity within the U.S. We also estimate the steady-state migration costs by fitting the pre-pandemic migration data to a model without pandemic, and estimate changes in migration costs during the pandemic by fitting our dynamic model with pandemic to observed migration and epidemiological data.

First, in the model without pandemic, it is easy to show that the migration share from region i to j is given by[15]

$$m_{ij,t} = \frac{V_{j,t+1}^{\beta}\mu_{ij}^{-1}}{\sum_{k=1}^{N} V_{k,t+1}^{\beta}\mu_{ik}^{-1}}.$$

Normalize $m_{ij,t}$ by the share of non-migrants who stay at the home state $m_{ii,t}$, we get

$$\ln\left(\frac{m_{ij,t}}{m_{ii,t}}\right) = \beta \ln V_{j,t+1} - \beta \ln V_{i,t+1} - \ln \mu_{ij}. \tag{17}$$

The model implies the share of migrants from state i to state j at time t, relative to the share of non-migrants who stay at the home of state i, is a function of value function differences and migration costs. Taking the first difference of equation (17), we estimate the following dynamic panel regression model to estimate migration elasticity κ.

$$\ln\frac{m_{ij,t}}{m_{ii,t}} - \beta \ln\frac{m_{ij,t+1}}{m_{jj,t+1}} = \frac{\beta}{\kappa}\left(\frac{w_{j,t+1}}{P_{j,t+1}} - \frac{w_{i,t+1}}{P_{i,t+1}}\right) - (1-\beta)\ln\mu_{ij} + \varepsilon_{ij,t}^{\kappa}, \tag{18}$$

which is similar to the estimation equation in Artuç et al (2010). But the differences of real wage, instead of relative ratio, now enter the equation, so scale of real wage matters when estimation. Fortunately, when estimating trade cost, technology across states is normalized. $\sum_{i=1}^{N} S_i^T = 0$ indicates $\prod_{i=1}^{N} A_i = 1$, thus we can scale the real wage by:

$$\prod_{i=1}^{N}\frac{w_i}{P_i} = \left(\frac{1}{\gamma}\right)^N \left(\prod_{i=1}^{N}\pi_{ii}\right)^{-1/\theta} \tag{19}$$

We further assume that migration costs depend on distance, contiguity, and destination effect:

$$\ln\widehat{\mu}_{ij} = \widehat{\beta}_0^M + \widehat{\beta}_1^M \ln dist_{ij} + \widehat{\beta}_2^M cont_{ij} + \widehat{\delta}_j^M. \tag{20}$$

For a given β, κ can be obtained by taking ratio of β and the coefficient in front of $\frac{w_{j,t+1}}{P_{j,t+1}} - \frac{w_{i,t+1}}{P_{i,t+1}}$. We provide the resulting estimates for κ in Table A.2. According to our

[15] Check Caliendo et al (2019) or Tombe and Zhu (2019).

preferred specification of column (4), the implied migration elasticity is $\kappa = 2.97$.[16] We infer the set of migration costs between any two states using the equation (20) and plot the distribution of these estimated migration costs in Figure A2.

Next, we estimate changes in migration costs during pandemics by matching migration and epidemiological data during the pandemic with our dynamic model with pandemic. Our estimated changes in migration costs minimize the difference between cases predicted by the model and the observed data across states for the period of March 7^{th} 2020 to July 12^{th} 2020 at bi-weekly frequency. We set the distribution of agents across states G_0 as the observed data in first period. Then we assume that migration costs to a state j change in a constant rate $\delta_{j,t}^{\dot{M}}$ until the last period of our data. The model would then generate series of changes in the number of newly infected cases for each state $\{\widehat{I}_{i,t}\}$. We then estimate the changes in migration costs by solving the following problem

$$min_{\delta_{j,t}^{\dot{M}}} \sum_i (\widehat{I}_{i,18}/I_{i,18} - 1)^2.$$

As shown in Figure 4, during the early pandemics period, migration inflow barriers in most states stayed relatively stable with a weekly change rate less than 5%, but Vermont, New Hampshire, Utah and Maine witnessed an increasing migration inflow barrier with rate 20.5%, 8.8%, 5.6% and 5.1%. On the contrary, Washington, Indiana, Minnesota and Georgia muted their inflow barriers by rates of 13.2%, 9.3%, 7.6% and 5.8%.

Pandemics

In this subsection, we report our estimates of epidemiological parameters. We set the mortality rate γ_i^D equal to 0.5 percent, which is a weighted average of U.S. population for different age groups (Eichenbaum et al, 2020). It takes on average 18 days to either recover or die for infected group (Atkeson, 2020), and our model is set at bi-weekly level, there is $\gamma_i^R + \gamma_i^D = 14/18$. Thus we obtain recovery rate γ_i^R of 0.77.

We adopt the following empirical model to estimate the parameters in the matching function χ_i, ϕ,and $\eta_{i,t}$:

$$\ln \frac{T_{i,t}}{\overline{S}_{i,t}} = \ln \chi_i + \phi \ln \eta_{i,t} + \phi \ln \frac{\overline{I}_{i,t}}{L_{i,t}}.$$

By assuming that $\ln \eta_{i,t} = -b \cdot Policy_{i,t}$[17], while $Policy_{i,t}$ is the containment policy adopted by state i at period t, we get the following specification for estimation.

$$\ln \frac{T_{i,t}}{\overline{S}_{i,t}} = \beta_0^P + \beta_1^P Policy_{i,t} + \beta_2^P \ln \frac{\overline{I}_{i,t}}{L_{i,t}} + \iota_i + \varepsilon_{i,t}^P.$$

Only the accumulated number of positive, recovery and death are available in the data. We therefore calculate the stock of infected $\overline{I}_{i,t}$ as the difference of accumulated

[16] Fajgelbaum et al.(2019) estimate the migration elasticity using a static migration model applied to the United States and explore the variation in state-level taxes. Their estimates vary across different model specifications and range from 1.36 to 2.25.

[17] Actually, we assume that $\ln \eta_{i,t} = a - b \cdot Policy_{i,t}$ as linear form. But before the pandemic, there was no containment policy and $\eta_{i,t} = 1$, which means that $a = 0$.

number of positive and recovery. Once empirical model is estimated, we infer coefficients $\phi = \beta_2^P$, $b = -\beta_1^P/\beta_2^P$, and $\chi_i = exp\{\beta_0^P + \iota_i\}$.

We find that the estimated that $\phi = 0.519$ and $b = 0.0154$. Figure 5 plots the region specific matching scale parameter χ_i. We find that χ_i tends to be higher in states with higher population density. For example, χ_i is greater than 0.2 in populous states like Massachusetts, Maryland, New Jersey, Rhode Island, DC, Illinois and Delaware but below 0.02 in states with small population like Montana and Hawaii.

5.2 Simulating a Three-region Economy

Before running the counterfactual simulation, we first simulate an artificial three-region economy, which helps us understand how endogenous migration affects the propagation of shocks across regions in our model. Other than the basic parameters mentioned above, we assume the remaining parameters take the value in Table 4. Therefore, the three regions are symmetric in fundamentals. In dynamic simulation with disease outbreaks, we assume 10% people in region 1 get infected by the virus at the initial period.

Comparative Statics at Steady States

At the steady state, there should be no infected people. If the pandemic gets under control, a fraction of the susceptible remain uninfected while those who got infected die or recover. In this case, the Recovered and the Susceptible coexist at the steady state. However, if the pandemic exploded, infection does not stop as long as there are susceptible people. In this case, we will only observe the Recovered at the steady state.

Figure 6 plots the welfare level of each type of agent at the steady state and how it varies with model parameters. The blue lines are for the welfare of the Infected, and red ones for the Susceptible and Recovered who share the same welfare at the steady-state. It is obvious that the Infected are much worse off due to a positive probability of death.[18] Lowering trade friction and migration cost for the Susceptible and Recovered will increase their welfare. Besides, a higher recovery rate, lower death rate, higher effective labor supply, and lower migration cost for the Infected will raise the welfare of the Infected but does not affect the welfare of the Susceptible and Recovered.

Comparative Dynamics of Containment Policies

We next examine the effect of local containment policy on the transition path by varying the effective labor supply, η. It determines the frequency of interaction between the Infected and other groups and measures the strictness of the within-region containment policy. Figure 7 plots the simulation results. The solid lines are for the case that $\eta = 0.5$, while dashed ones $\eta = 0.3$, a scenario with lower effective labor supply, therefore stricter containment policy.

With stricter local containment policy, pandemics fade out more rapidly, cause fewer death, and preserve a higher labor force at the steady state. Infection declines in region 1, but forms a hump shape in region 2 and 3. This is because migration spreads pandemics from region 1 to region 2 and 3 within a few periods. However, a small reproduction

[18] Note the left axis is for the Infected while the axis is for the Susceptible and Recovered.

number guarantees a converging steady state with the Susceptible surviving, leading to the hump shape for regions 2 and 3.

As for the total labor supply, we find a U-shape for region 1 but an inverted U-shape for regions 2 and 3. To better illustrate how different groups react, we compare dynamic migration rates with steady state, as shown in the last row of Figure 7. We find that migration rates of the Infected and the Recovered barely deviate (blue lines basically coincide with black lines) but the migration rate of the Susceptible significantly deviates from the steady state. To avoid infection in region 1, the outflow of the Susceptible people from region 1 increases, while the inflow decreases. As a substitution, migration of the Susceptible between regions 2 and 3 rises. The endogenous migration of the Susceptible explains the first half of total labor dynamics. A few periods later, infection rates across regions converge to a nearly equal level as the virus spreads to regions 2 and 3. Then migration rates return to steady-state, and so does the total labor force, which accounts for the second half of total labor dynamics. Figure 8 plots the effect on the real economy. Again, we find that the dynamics are mainly driven by the Susceptible. The changes of their welfare level are much bigger than other groups, at a different order of magnitude.

Comparative Dynamics of Migration Cost

Suppose region 2 puts forward a policy aiming at reducing immigrants from region 1, the origin of the pandemic, by raising the migration cost moving from region 1 μ_{12}. Figure 9 plots the simulation results. From the figure, we find that the policy levels down the infection rate in region 2 but increase those in region 1 and 3. Correspondingly, the number of Susceptible in region 2 rises, but those in region 1 and 3 falls, resulting in a seemingly "beggar-thy-neighbor" effect. However, when we consider the total labor, region 2 experiences a sharp decline while regions 1 and 3 attract more labor. The reason is relatively straightforward: people in region 1 find it hard to move to region 2, while region 3 absorbs some immigrants from region 1 who would have moved to region 2. Moreover, migration of the Susceptible is the main driving force of demographic dynamics.

Policy Competition and Coordination

In this subsection, we explore the scope of regional competition and coordination in containment policies (measured by η) during the pandemic. Before we begin, we first define the welfare criteria adopted by policy makers.

Definition 1: regional welfare is the sum of welfare across all local residents at the steady-state;

Definition 2: regional welfare is the sum of values across all residents in the first period.

We consider the following three policy scenarios:

Policy scenario 1: Taking η_3 as given, regions 1 and 2 make policy simultaneously to maximize its welfare

Policy scenario 2: Taking η_3 as given, a social planner chooses η_1 and η_2 to maximize the total welfare of regions 1 and 2

Policy scenario 3: Keeping η_3 constant, a social planner chooses η_1 and η_2 to maximize the total welfare of regions 1, 2 and 3

Welfare definition 1 focuses on long-term well-being, while welfare definition 2 puts more weight on short-term well-being. Policy scenario 1 is the case with local policy competition. Policy scenario 2 can be thought of as a federal government coordinating among states, given foreign policies. And policy scenario 3 examines the global welfare.

As shown in Figure 10, if the pandemic starts in region 3, only the outcome using welfare definition 2 in policy scenario 3 indicates a policy combination, as $\eta_1 = \eta_2 = 1$. Otherwise, $\eta_1 = \eta_2 = 0$ would be the optimal policy, i.e., regions 1 and 2 make containment policy as strict as possible.

Next, we consider the case that the pandemic start in both regions 1 and 2. Figure 11 plots how welfares varies with policies. It is interesting to note that if policy makers are short-sighted, all three policy scenarios call for a loose policy combination such that $\eta_1 = \eta_2 = 1$. But if long-term welfare matters, strict policies are adopted in all scenarios.

Finally, if the pandemic starts in region 1, far-sighted policy makers always adopt strict containment policies. Short-sighted global or federal policy markers support a loose policy combination. However, short-sighted local policy makers would prefer a rather interesting policy combination of $\eta_1 = 1$ and $\eta_2 = 0$ in competition. According to Figure 12, if short-sighted local governments make their policy choices without coordination, states with more severe pandemics will not tighten containment policies while other states with tend to adopt strict containment policies.

6 Counterfactual Analysis

In this section, we use our calibrated US economy with 50 states to simulate the effect of containment policies and migration policies on the pandemic and real economy. We first show the baseline calibrated economy. We then conduct two counterfactual simulations. The first one removes the estimated changes in migration friction. The second one assumes away containment policies implemented by state governments.

6.1 The Baseline Economy

In this subsection, we first take the calibrated US states parameters into the model and run a dynamic series of infection cases from March 7^{th} to July 12^{th}. As shown in Figure 13, the red dashed lines are model-predicted accumulative infection while black solid lines are reported cases in reality, all normalized with the states population. Our results captures the general rising infection across all states. There are some states with relatively good fit, such as Colorado (CO), Illinois (IL), Massachusetts (MA), New Jersey (NJ) and Pennsylvania (PA). A few states display significant discrepancy. States like California (CA) and Florida (FL) have missing data on recovered cases, which lead to exaggeration in the number of continuing infected cases. Some deviations can be attributed to the pooling of recovery and mortality rate when we calibrate the model. This tends to overestimate the infection curves in states with relative poor medical conditions and underestimate infection in states with more abundant medical resources.

6.2 No Change in Migration Costs

Migration barriers play roles in preventing the spread of pandemics. During pandemics, tests and quarantine measures taken by local governments increase the cost of moving across states. To the extreme, governments can shut down the borders and bring all movements to a full stop. Our model would help us know the extent that such changes in migration friction slow down the pandemic in the US.

We do that by eliminating the calibrated changes in effective migration costs shown in Figure 4. Figure 14 presents the corresponding counterfactual result. From the top panel of the figure, we find that states with high calibrated migration barriers increments suffer more from removing the migration friction. For example, infected cases in Vermont and New Hampshire at July 12^{th} will grow for an extra 17.2% and 7.1%, respectively; but those in Washington, Indiana and Minnesota will reduce by 9.6%, 6.9% and 5.4%, respectively.

However, when we dig deeper, the bottom panel in Figure 14 shows that infection probabilities in almost all the states rise. Though bigger states do not impose significant change in migration cost, like DC, New Jersey, Delaware and Maryland, they experience greater increase in infection risks due to geographic centrality. People will on average get infected with a probability 0.07 basic points higher, and 56.1 thousands more will be involved into infection by July 12^{th} 2020. So actually the reduction of infection cases in some states should be attributed to the incentive to move out, and the migration changes during early pandemics do work in controlling epidemic situation. As for economic impacts, it is not surprising that states with lower migration inflow barriers grow relatively faster. Indeed migration cost change will make a trade-off between economic growth and pandemics control.

6.3 No Containment Policies

There is evidence that containment policies reduce infection effectively (Hsiang et al, 2020). In our endogenous migration context, isolation of infected individuals within regions not only reduces the state-specific infection probability of the susceptible population, but benefits all the other states via migration linkages.

Our conterfactual simulation removes the estimated changes in containment policies, shown in Figure 5. The results in Figure 15 indicate that containment policies contribute significantly to slow down the pandemic. Without the policies, infected cases in all states at July 12^{th} would have grown more than 10%. In states such as New York (NY) and California (CA), the number of cases would near double, a growth of 83.9% for NY and 71.5% for CA. Again, infection probabilities in all the states raise, but in fairly greater magnitudes. Residents in Massachusetts even find themselves more likely to get infected by 1.5 percent points. Overall, containment policy reduce infection cases up to 31.6 million across US by July 12^{th} 2020. And intuitively, containment policies do more economic harm to populous states. But the counterfactual GDP shares fluctuate within 1%.

7 Conclusion

Although the rapid globalization and urbanization generate enormous gains from international trade and immigration, they also lead to more frequent outbreaks of infectious diseases. In this paper, we develop a framework to analyze the role of human migration in both economic and disease dynamics under the context of the COVID-19 pandemic. We combine a multi-region spatial general equilibrium model of international trade with endogenous migration across regions, adding the SIR model from epidemiology. We address several important features by building on two branches of models. First, as we allow endogenous migration during the pandemic, agents weigh the risk of infection and economic opportunities across regions and optimize the location chosen for work and life in each period. Hence, agents automatically practice social distancing by avoiding high-risk areas. This could help to clearly identify the real effect of containment policies through counterfactual analysis. Second, although embedding an SIR model in an economic geography model with interregional trade and dynamic migration makes it very complicated to solve, we can still provide some analytical results about R_0 to demonstrate the interaction between migration and the pandemic.

Despite adding many moving parts to the SIR model, our model is quite versatile and can be extended to study other relevant policies taken by governments. In this section, we consider two important policies, vaccination and testing, which are in the Appendix A.2.

References

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, MD, 2020. A multi-risk SIR model with optimally targeted lockdown (No. w27102). National Bureau of Economic Research.

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, 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., 2020. A simple planning problem for covid-19 lockdown (No. w26981). National Bureau of Economic Research.

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.

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.

Bartlett, M.S., 1956. Deterministic and stochastic models for recurrent epidemics. In Proceedings of the third Berkeley symposium on mathematical statistics and probability (Vol. 4, No. 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. 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.

Bisin, A. and Moro, A., 2020. Learning epidemiology by doing: The empirical implications of a spatial-sir model with behavioral responses (No. w27590). National Bureau of Economic Research.

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., 2017. Goods and factor market integration: a quantitative assessment of the EU enlargement (No. w23695). National Bureau of Economic Research.

Chen, X., Huang, H., Ju, J., Sun, R. and Zhang, J., 2021. Impact of vaccination on the COVID-19 pandemic: Evidence from US states. Available at SSRN 3845163.

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.

Costinot, A. and Rodrguez-Clare, A., 2014. Trade theory with numbers: Quantifying the consequences of globalization. In Handbook of international economics (Vol. 4, pp. 197-261). Elsevier.

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.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. "Unbalanced trade." American Economic Review 97, no. 2 (2007): 351-355.

Eaton, J. and Kortum, S., 2002. Technology, geography, and trade. Econometrica, 70(5), pp.1741-1779.

Eichenbaum, M.S., Rebelo, S. and Trabandt, M., 2020. The macroeconomics of epidemics (No. w26882). National Bureau of Economic Research.

Fajgelbaum, P., Khandelwal, A., Kim, W., Mantovani, C. and Schaal, E., 2020. Optimal lockdown in a commuting network (No. w27441). National Bureau of Economic Research.

Fajgelbaum, P.D., Morales, E., Surez Serrato, J.C. and Zidar, O., 2019. State taxes and spatial misallocation. The Review of Economic Studies, 86(1), pp.333-376.

Fang, Hanming, and Long Wang, and Yang Yang. Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in China. Forthcoming. Journal of Public Economics, 2020.

Glaeser, E.L., Gorback, C. and Redding, S.J., 2020. 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.

Hale, Thomas, Noam Angrist, Emily Cameron-Blake, Laura Hallas, Beatriz Kira, Saptarshi Majumdar, Anna Petherick, Toby Phillips, Helen Tatlow, Samuel Webster (2020). Oxford COVID-19 Government Response Tracker, Blavatnik School of Government.

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.

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.

Fujita, M., Krugman, P.R. and Venables, A., 1999. The spatial economy: Cities, regions, and international trade. MIT press.

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. Working paper.

Poncet, S., 2006. Provincial migration dynamics in China: Borders, costs and economic motivations. Regional Science and Urban Economics, 36(3), pp.385-398.

Raifman, Julia, Nocka, Kristen, Jones, David, Bor, Jacob, Lipson, Sarah, Jay, Jonathan, Galea, Sandro. 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.

Waugh, Michael E. (2010). International Trade and Income Differences, The American Economic Review, 100(5), pp.2093-2124

Redding, S.J., 2016. Goods trade, factor mobility and welfare. Journal of International Economics, 101, pp.148-167.

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

Figures

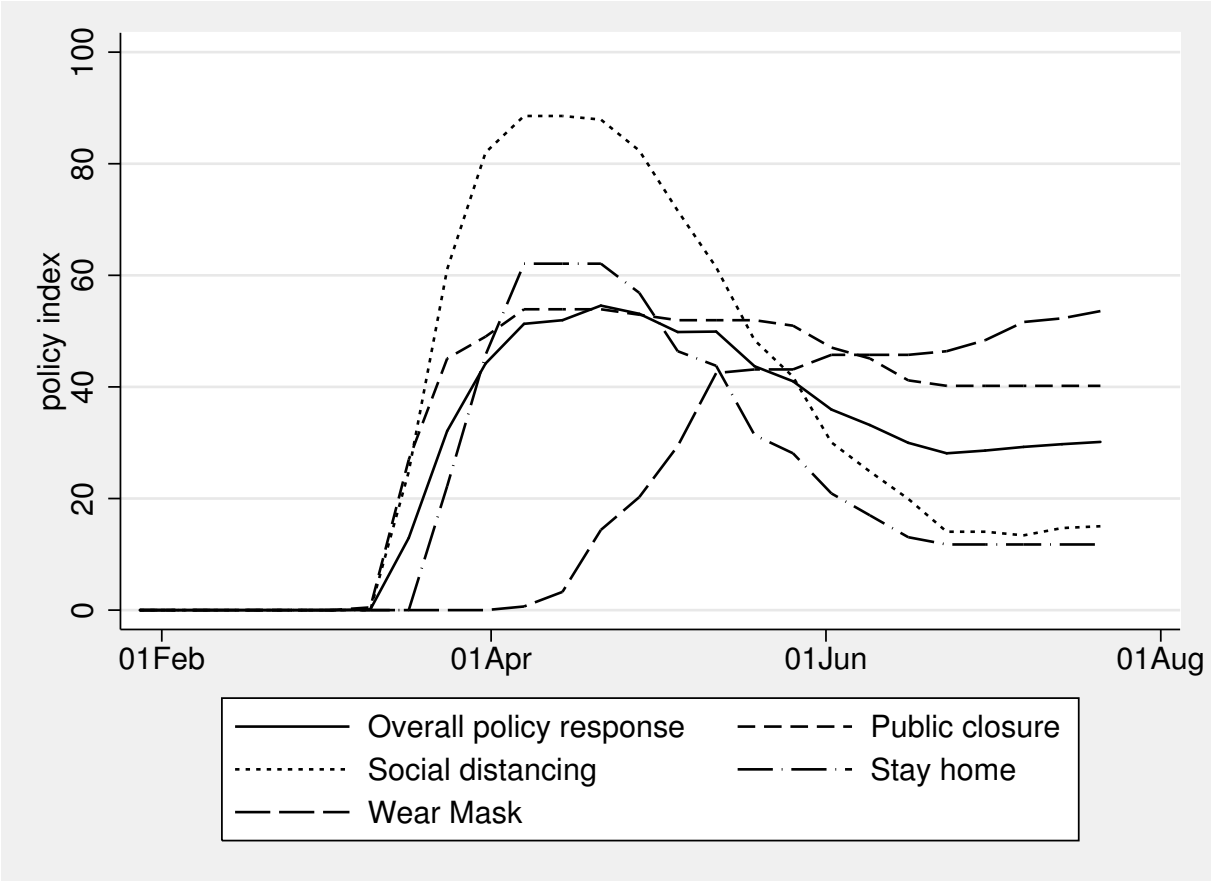

Figure 1: Evolution of COVID-19 policies averaged across US states

Notes: The figure plots the average of policy indexes across states over time.

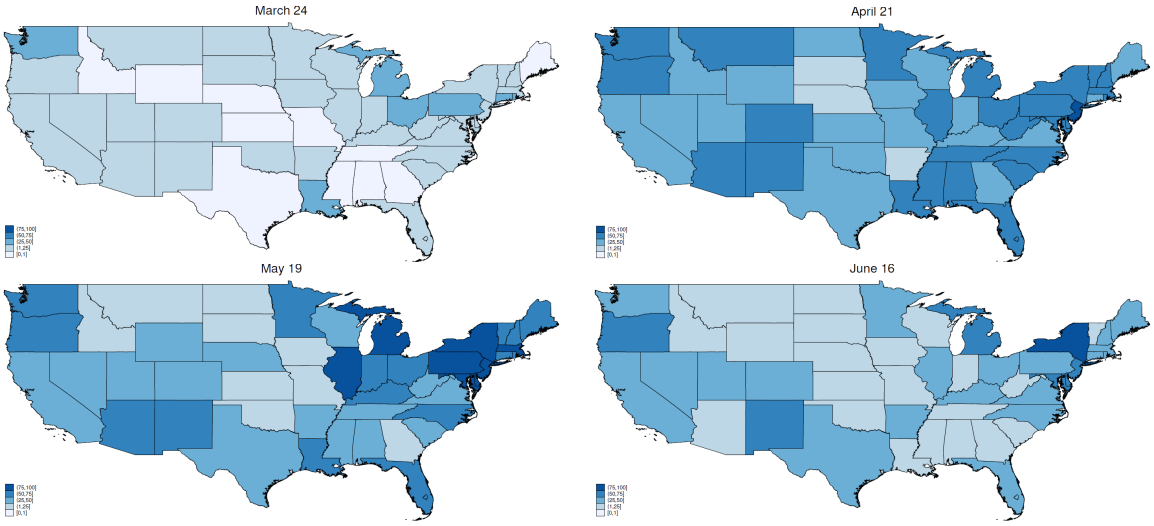

Figure 2: Overall policy response across states and time

Notes: The figure plots the overall containment policy across states at different points in time. Darker colors denote more stringent lockdowns.

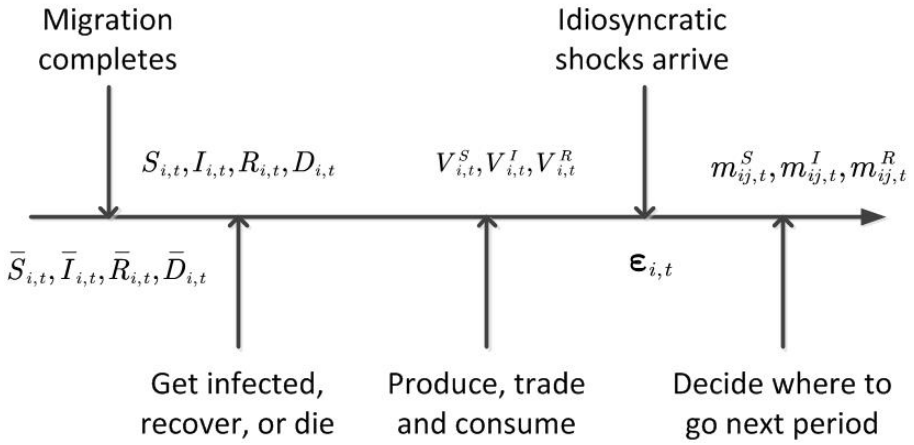

Figure 3: Model time line

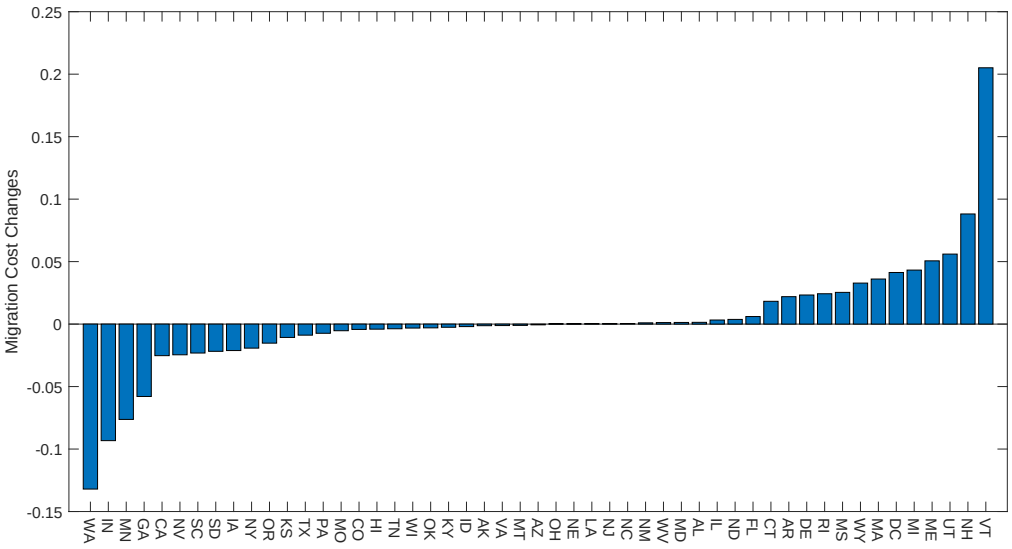

Figure 4: Calibrated Changes in Migration Cost

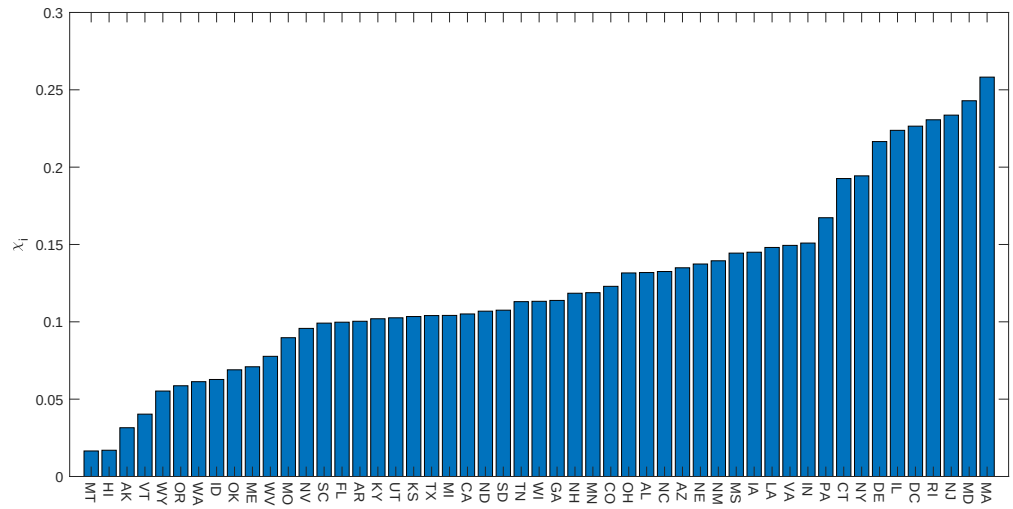

Figure 5: Calibrated Scale Parameter of the Matching Function χ_i

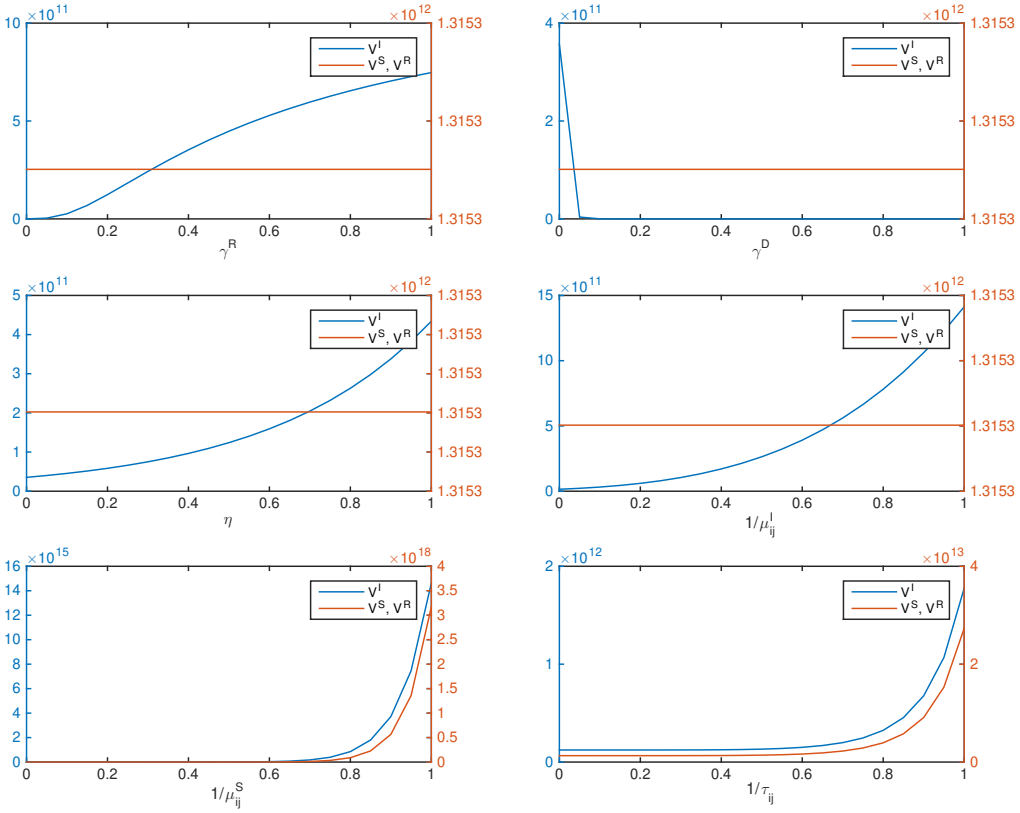

Figure 6: Comparative Statics of Welfare in Steady State

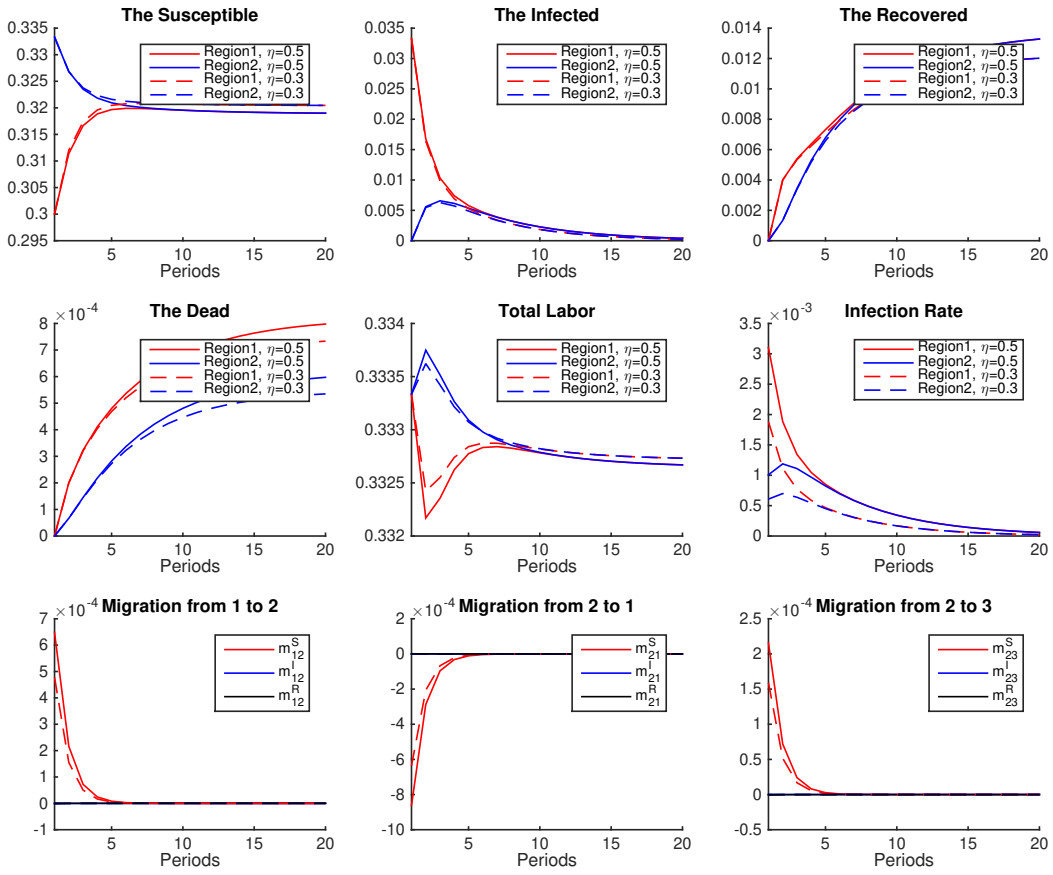

Figure 7: Effective labor supply η and demographic dynamics

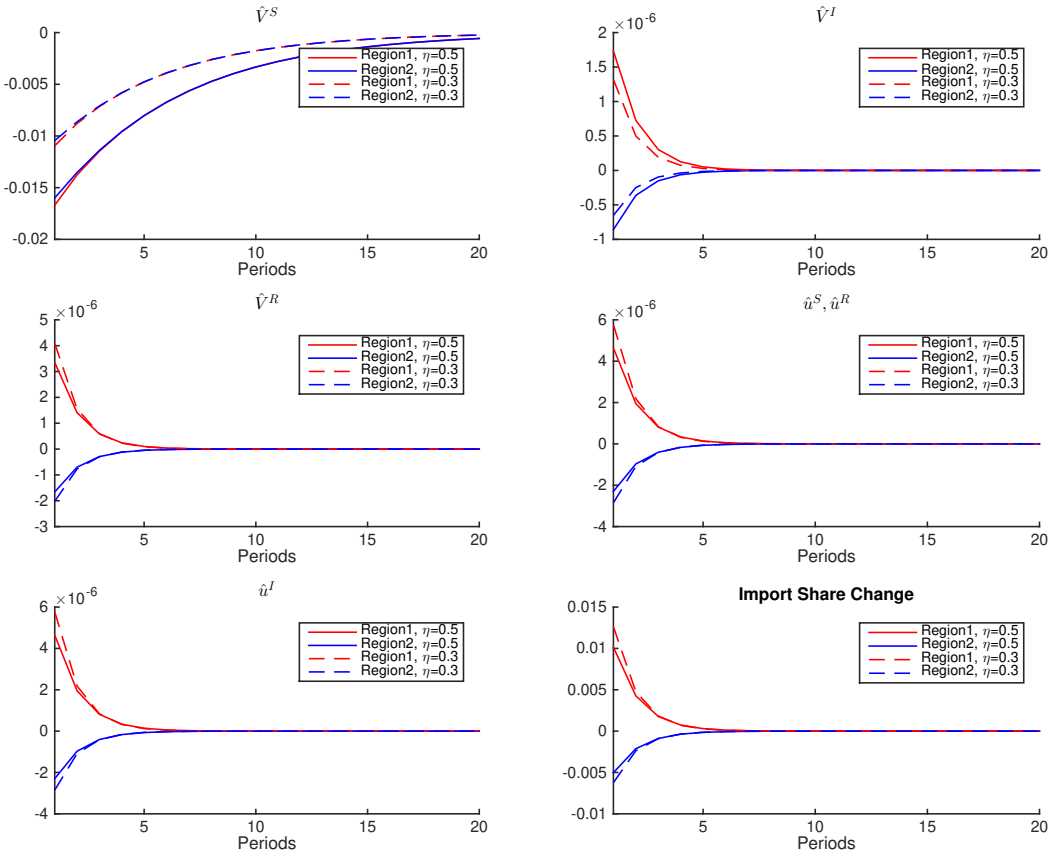

Figure 8: Effective labor supply η and the dynamics of the real economy

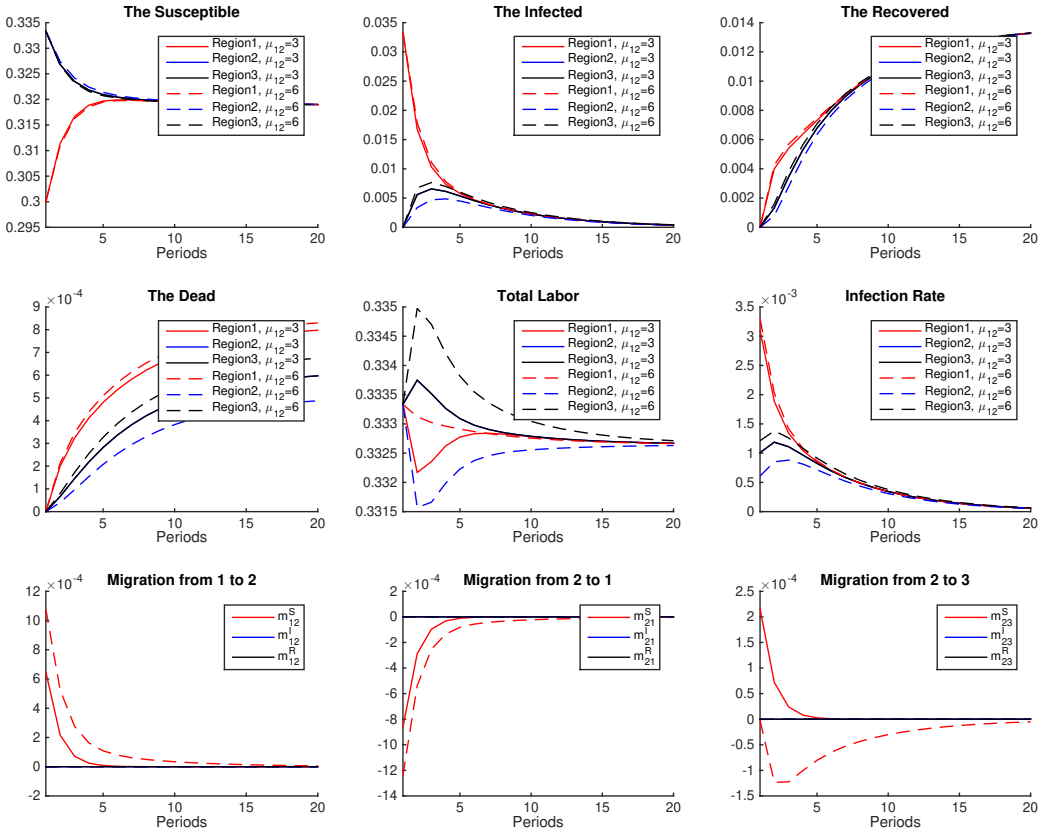

Figure 9: Migration cost μ_{12}^I and demographic dynamics

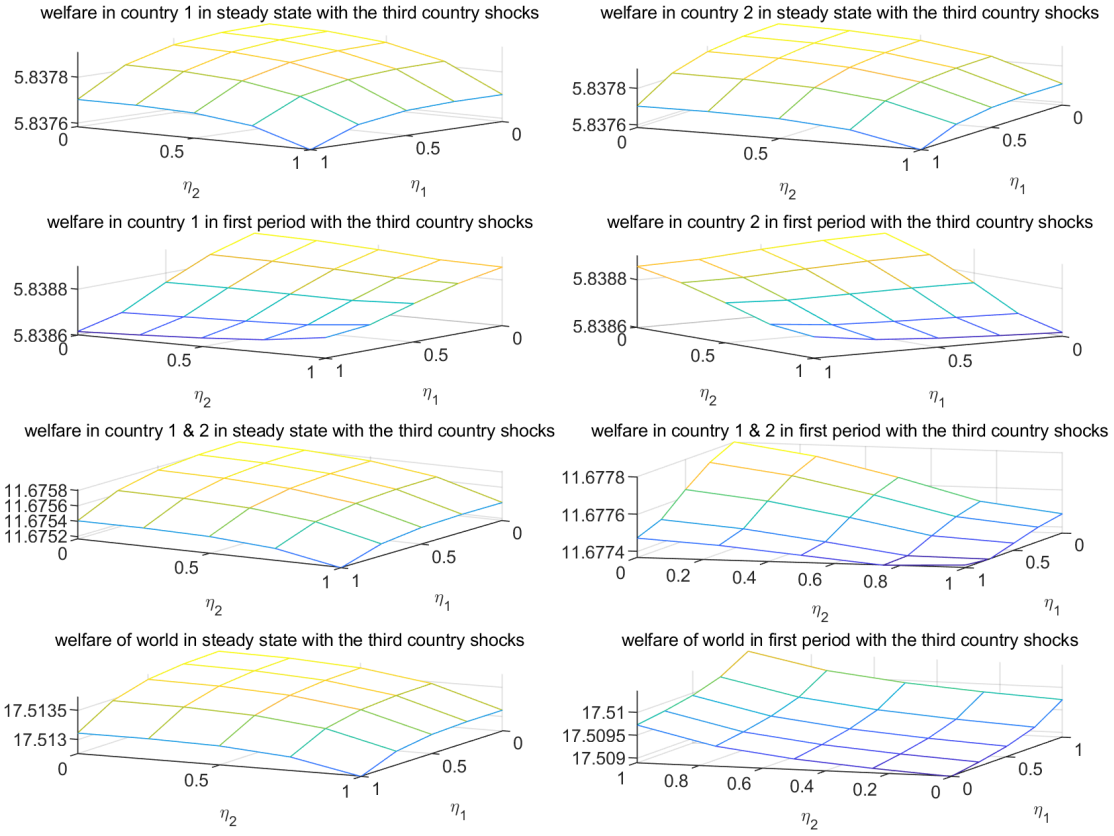

Figure 10: Policy interactions for a pandemics shock in region 3

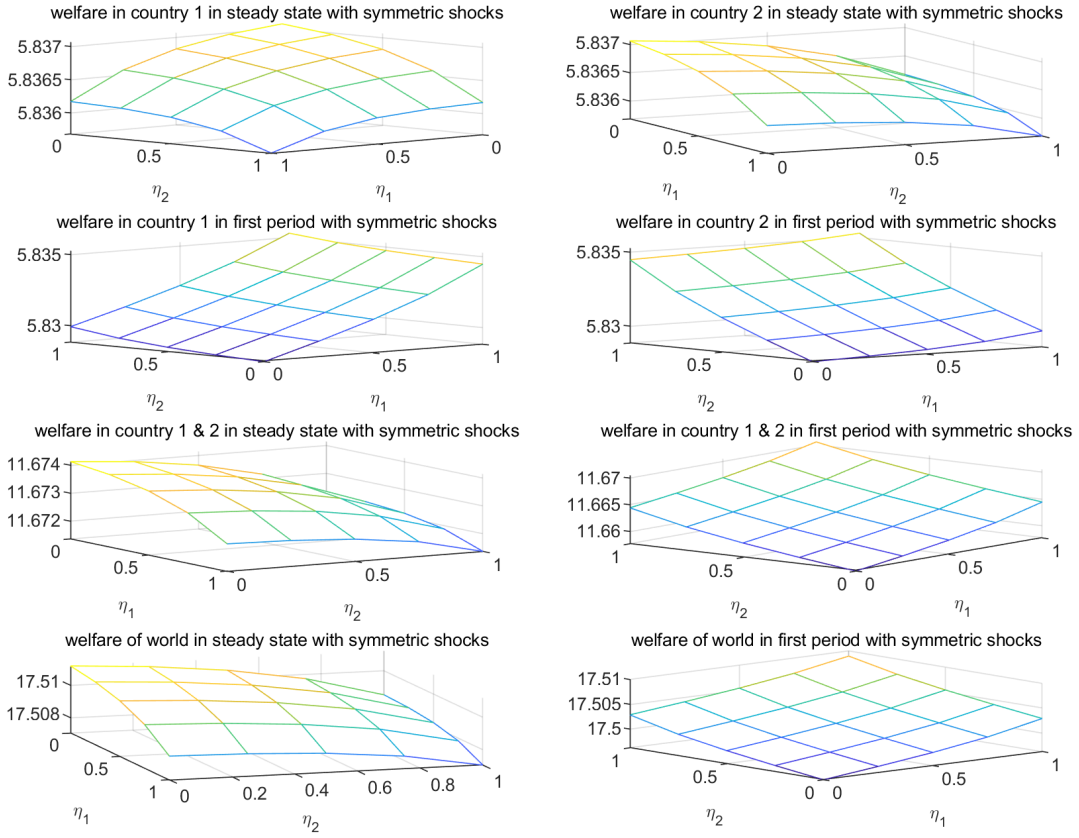

Figure 11: Policy interactions for pandemics shocks in region 1 and 2

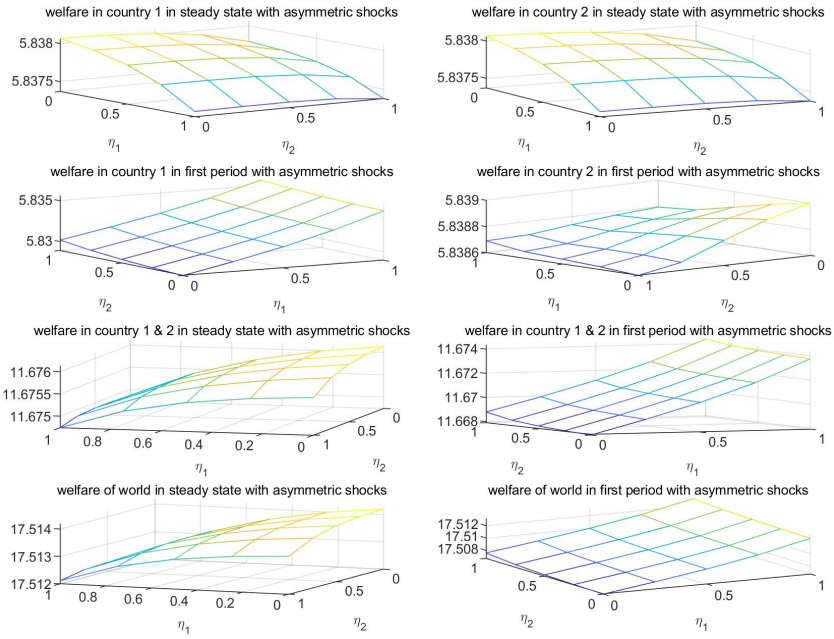

Figure 12: Policy interaction for asymmetric pandemics shocks in region 1

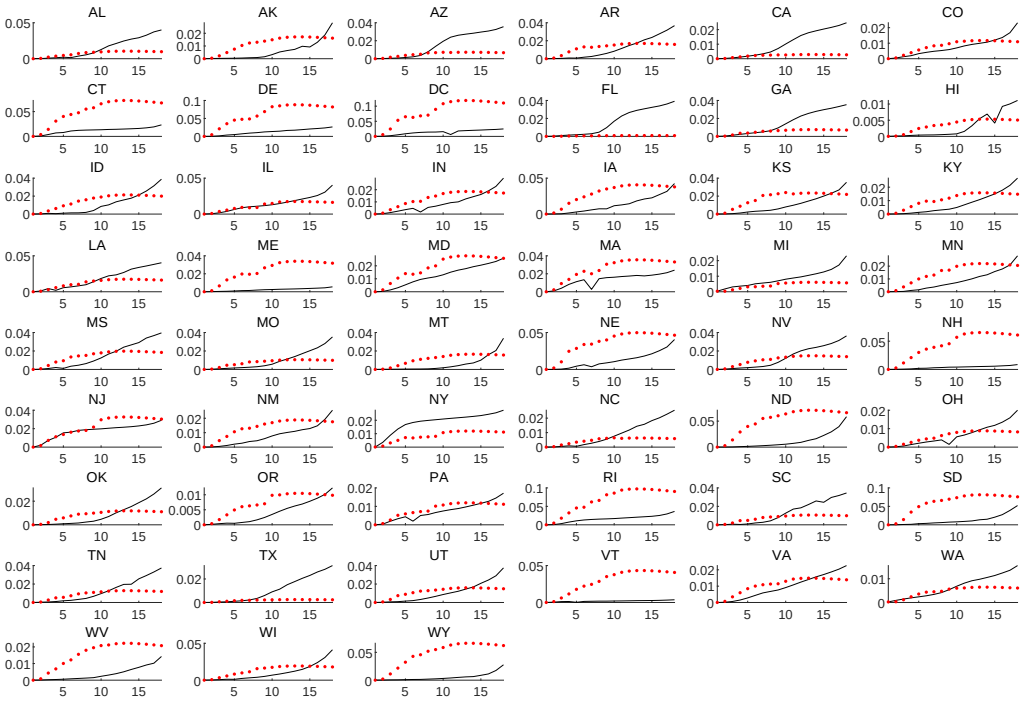

Figure 13: Model Predicted Curves of US States

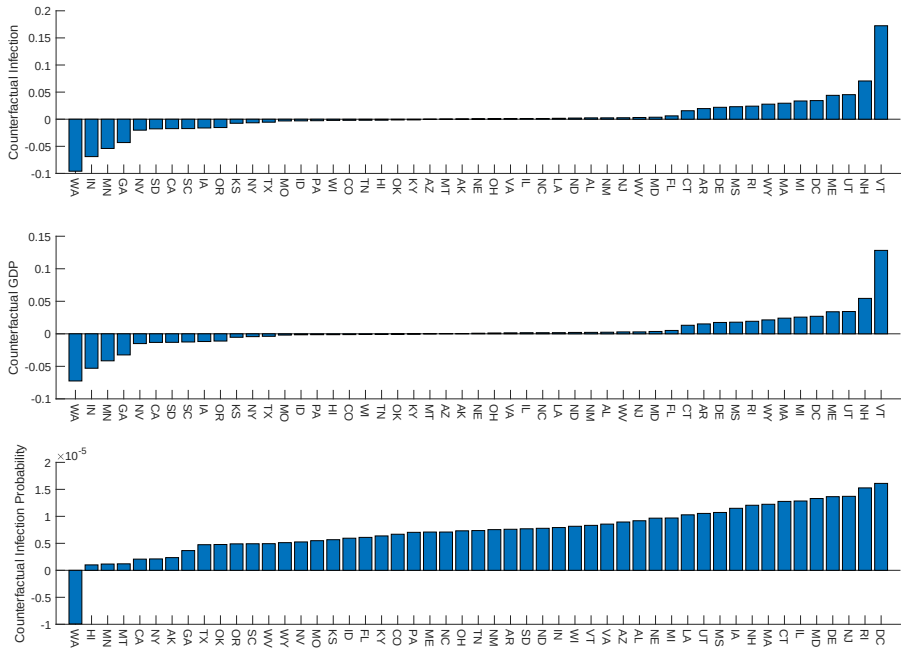

Figure 14: Counterfactual Infection with No Migration Cost Change

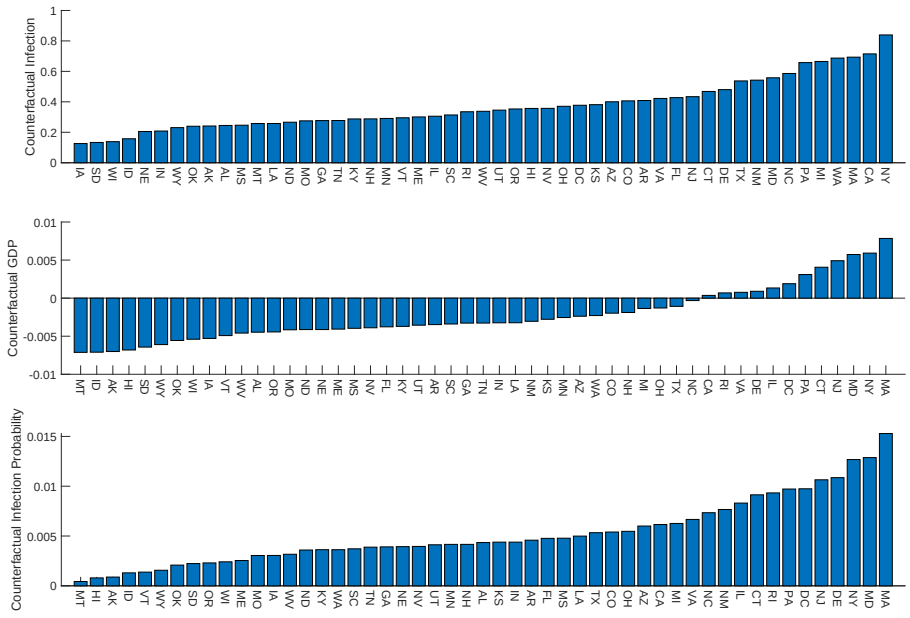

Figure 15: Counterfactual Infection with No Containment Policy

Tables

Table 1: Summary Statistics

	Definition	Obs	Mean	Std.Dev	Min	Median	Max
Panel A. Cross State Human Mobility							
Max migration	Max of bilateral migration flow between states at bi-weekly frequency	30600	0.011	0.032	0	0.003	0.706
Mean migration	Average of bilateral migration flow between states at bi-weekly frequency	30600	0.010	0.030	0	0.003	0.696
Quarantine policy at destination	Quarantine policy of destination state	30600	22.60	38.62	0	0	100
Infection rate at origin	Infection rate measured at origin state level	30600	0.0009	0.0011	0	0.0005	0.0073
Infection rate at destination	Infection rate measured at destination state level	30600	0.0009	0.0011	0	0.0005	0.0073
Distance	bilateral distance between states	30600	1981.04	1461.31	37.36	1649.41	8176.1
Border	whether two states share border	30600	0.08	0.28	0	0	1
Panel B. Pandemic, Closure, and Exposure to pandemics via migration							
Closure	US state-level policy index	663	36.80	31.20	0	43.75	100
Number of new cases	Number of new cases at state bi-weekly level	663	5828.81	14794.2	1	1227	156041
ln (Number of new cases)	Ln (number of new cases) at state bi-weekly level	663	5.79	3.62	0	7.11	11.96
$\sum_{j\neq i} m_{ji}^{max} ln(newcases_j)$	sum of ln (new cases) across other states, weighted by max migration in the initial period	663	3.84	3.43	0	3.28	12.46
$\sum_{j\neq i} m_{ji}^{mean} ln(newcases_j)$	sum of ln (new cases) across other states, weighted by mean migration in the initial period	663	3.58	3.21	0	3.03	11.99
$\sum_{j\neq i} m_{ji}^{max} ln(cumulativecases_j)$	sum of ln (cumulative cases) across other states, weighted by max migration in the initial period	663	4.48	4.09	0	3.71	15.56
$\sum_{j\neq i} m_{ji}^{mean} ln(cumulativecases_j)$	sum of ln (cumulative cases) across other states, weighted by mean migration in the initial period	663	4.16	3.82	0	3.43	14.85

Notes: Panel A summarizes the data for the regressions to generate Table 2 which is at the state-pair bi-weekly level. Panel B summarizes the data for the regressions to generate Table 3 which is at the state bi-weekly level.

Table 2: The COVID-19 pandemic and US state-to-state migration

	(1)	(2)	(3)	(4)
VARIABLES	Max migration		Average migration	
	PPML		PPML	
Lagged quarantine policy at destination j	0.000	0.000	0.000	0.000
	(0.256)	(0.227)	(0.195)	(0.170)
Lagged infection rate at origin i	-1.567	-0.116	-2.647	-1.299
	(-0.182)	(-0.013)	(-0.297)	(-0.140)
Lagged infection rate of destination j	-19.435***	-19.567***	-15.172***	-15.123***
	(-3.625)	(-3.654)	(-2.778)	(-2.788)
Ln(distance)	-1.334***		-1.363***	
	(-41.422)		(-40.346)	
Shared State Border =1	1.021***		1.053***	
	(16.433)		(16.529)	
Constant	4.452***	-3.204***	4.499***	-3.256***
	(19.795)	(-438.794)	(19.227)	(-428.667)
Observations	30,600	30,600	30,600	30,600
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 the effect of infection rate in the destination on US state-to-state migration flows, conditional on quarantine policy at the destination state. We aggregate the daily cross-state mobility data from PlaceIQ to bi-weekly level. Columns (1) and (2) use the max flow within the two weeks as the dependant variable, while Columns (3) and (4) uses the average within the two weeks. The numbers in the parentheses are robust t-statistics with standard error clustered at the bi-weekly level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table 3: Cross-state mobility, containment policies and COVID-19 pandemic in the US

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln (number of new cases)				number of new cases			
Closure	-0.012**	-0.012**	-0.013**	-0.013**	-0.016***	-0.016***	-0.020***	-0.021***
	(-2.627)	(-2.587)	(-2.784)	(-2.746)	(-2.684)	(-2.674)	(-2.587)	(-2.667)
$\sum_{j\neq i} m_{ji}^{max} ln(newcases_j)$	0.409***				1.436***			
	(-25.817)				(-13.992)			
$\sum_{j\neq i} m_{ji}^{mean} ln(newcases_j)$		0.433***				1.512***		
		(-25.574)				(-14.282)		
$\sum_{j\neq i} m_{ji}^{max} ln(cumulativecases_j)$			0.308***				0.344**	
			(-13.143)				(-2.519)	
$\sum_{j\neq i} m_{ji}^{mean} ln(cumulativecases_j)$				0.325***				0.399***
				(-13.269)				(-2.662)
Constant	4.659***	4.680***	4.882***	4.902***	-0.666	-0.465	7.768***	7.539***
	(-27.425)	(-27.656)	(-19.862)	(-20.075)	(-0.597)	(-0.430)	(-5.605)	(-5.316)
Observations	663	663	663	663	663	663	663	663
R-squared	0.959	0.959	0.955	0.955	-	-	-	-
BiWeek FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table estimates the effect of cross-state mobility and containment policies on COVID-19 in the US. The Closure index is based on data of US state COVID-19 containment policy from CUSP. We follow Hale et al.(2020) to construct a US state-level policy index, which ranged from 0 to 100. We aggregate the daily cross-state mobility data from PlaceIQ to bi-weekly level. m_{ji}^{max} is the maximium of the migration share from state j to i in the initial period, while m_{ji}^{mean} is the average within period 0. We use these different weights for number of cases of the origin state j, while the cases are measured in terms ln(new cases) in the previous period or the cumulative number of cases. Columns (1)-(4) are estimated in OLS while columns (5)-(8) are in PPML. The numbers in the parentheses are robust t-statistics with standard error clustered at the bi-weekly level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table 4: Additional model parameters for a three-region economy

Parameters	Definition	value
T_i	Technology	1
χ_i	Scale of infection rate	0.1
ϕ	Matching elasticity of infection rate	1
$\tau_{ij}, i \neq j$	Trade cost	3
$\mu_{ij}, i \neq j$	Migration cost for the Susceptible	3
γ_i^R	Recovery rate	0.2
γ_i^D	Mortality rate	0.01
$\eta_{i,t}$	Effective labor supply for the Infected	0.5

Appendix

A.1 Algorithm

A.1.1 The model without pandemic

System of Equations, solve for $\mathbf{w}_t$ ($\mathbf{L}_t$):

$$\pi_{ij,t} = \frac{A_i(\tau_{ij}w_{i,t})^{-\theta}}{\sum_{k=1}^{N} A_k(\tau_{kj}w_{k,t})^{-\theta}} \tag{A.1}$$

$$w_{it}L_{it} = \sum_{j=1}^{N} \pi_{ijt} w_{jt} L_{jt} \tag{A.2}$$

with expressions:

$$\begin{aligned} u_{i,t} &= \ln\left(\frac{w_{i,t}}{P_{i,t}}\right) \\ P_{it} &= \gamma(\sum_{k=1}^{N} A_k(\tau_{kj}w_k)^{-\theta})^{-1/\theta} \end{aligned}$$

And intertemporal conditions:

$$\begin{aligned} V_{i,t} &= \exp\left(\frac{u_{i,t}}{\kappa}\right) \sum_{j=1}^{N} V_{j,t+1}^{\beta} \mu_{ij}^{-1} \\ m_{ij,t} &= \frac{V_{j,t+1}^{\beta}\mu_{ij}^{-1}}{\sum_{k=1}^{N} V_{k,t+1}^{\beta}\mu_{ik}^{-1}} \\ L_{j,t+1} &= \sum_{i=1}^{N} m_{ij,t} L_{i,t} \end{aligned}$$

Hat Algebra
Denote $\widehat{x}_t = x_t/x_{ss}$, then intertemporal conditions:

$$\widehat{V}_{i,t} = \exp\left(\frac{u_{i,t} - u_{i,ss}}{\kappa}\right) \sum_{j=1}^{N} m_{ij,ss} \widehat{V}_{j,t+1}^{\beta} \tag{A.3}$$

$$m_{ij,t} = \frac{m_{ij,ss}\widehat{V}_{j,t+1}^{\beta}}{\sum_{k=1}^{N} m_{ik,ss}\widehat{V}_{k,t+1}^{\beta}} \tag{A.4}$$

$$L_{j,t+1} = \sum_{i=1}^{N} m_{ij,t} L_{i,t} \tag{A.5}$$

Algorithm

1. Exogenously given $\{L_0\}$. Evaluate T large enough.
2. Take intial guess $\left\{\widehat{V}_t^{(0)}\right\}_{t=1}^{T}$.
3. For each iteration of $\left\{\widehat{V}_t^{(k)}\right\}_{t=1}^{T}$:
   (a) Solve forward for migration matrix $\{m_{t-1}\}_{t=1}^{T}$ from (A.4), and then the labor distribution $\{L_t\}_{t=1}^{T}$.
   (b) In each period t from 1 to T, solve for temporary allocation w_t from (A.1)(A.2), thus get P_t and u_t.
   (c) Using (A.3) to solve backward to get $\left\{\widehat{V}_t^{(k)\prime}\right\}_{t=1}^{T}$
4. If $\left\{\widehat{V}_t^{(k)\prime}\right\}$ is close to $\left\{\widehat{V}_t^{(k)}\right\}$, finish. Otherwise, set next guess $\left\{\widehat{V}_t^{(k+1)}\right\} = \left\{\widehat{V}_t^{(k)\prime}\right\}$. (What if $\left\{\widehat{V}_t^{(k+1)}\right\} = \left\{\widehat{V}_t^{(k)}\right\}^{1/2} \times \left\{\widehat{V}_t^{(k)\prime}\right\}^{1/2}$? Will it converge faster?)

A.1.2 Model with Pandemics

System of equations:

$$\pi_{ij,t} = \frac{A_i(\tau_{ij}w_{i,t})^{-\theta}}{\sum_{k=1}^{N} A_k(\tau_{kj}w_{k,t})^{-\theta}} \tag{A.6}$$

$$w_{it}\left(\eta_i I_{i,t} + S_{i,t} + R_{i,t}\right) = \sum_{j=1}^{N} \pi_{ijt} w_{jt}\left(\eta_j I_{j,t} + S_{j,t} + R_{j,t}\right) \tag{A.7}$$

with expressions:

$$\begin{aligned} u_{i,t} &= \frac{w_{i,t}}{P_{i,t}} \\ u_{i,t}^{I} &= \eta_i \frac{w_{i,t}}{P_{i,t}} \\ P_{it} &= \gamma\Big(\sum_{k=1}^{N} A_k(\tau_{kj}w_k)^{-\theta}\Big)^{-1/\theta} \end{aligned}$$

Actually, when there is no investment, all the labor income is spent on consumption, so contemporary utility is identical across S and R groups within regions. And intertemporal

conditions:

$$
\begin{aligned}
V_{it}^{S} &= \exp(\frac{u_{it}}{\kappa}) \sum_{j=1}^{N} \left(V_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})} \left(V_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}} \left(\mu_{ij}^{S}\right)^{-1} \\
V_{it}^{I} &= \exp(\frac{u_{it}^{I}}{\kappa}) \sum_{j=1}^{N} \left(V_{j,t+1}^{I}\right)^{\beta(1-\gamma_{j}^{R}-\gamma_{j}^{D})} \left(V_{j,t+1}^{R}\right)^{\beta\gamma_{j}^{R}} \left(\mu_{ij}^{I}\right)^{-1} \\
V_{it}^{R} &= \exp(\frac{u_{it}}{\kappa}) \sum_{j=1}^{N} \left(V_{j,t+1}^{R}\right)^{\beta} \left(\mu_{ij}^{R}\right)^{-1}
\end{aligned}
$$

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

Labor evolution conditions:

$$
\begin{aligned}
\overline{S}_{i,t} &= \sum_{j=1}^{N} S_{j,t-1} m_{ji,t-1}^{S} && \text{(A.8)} \\
\overline{I}_{i,t} &= \sum_{j=1}^{N} I_{j,t-1} m_{ji,t-1}^{I} && \text{(A.9)} \\
\overline{R}_{i,t} &= \sum_{j=1}^{N} R_{j,t-1} m_{ji,t-1}^{R} && \text{(A.10)} \\
\overline{D}_{i,t} &= \overline{D}_{i,t-1} && \text{(A.11)}
\end{aligned}
$$

$$
\begin{aligned}
S_{i,t} &= \overline{S}_{i,t} - T_{i,t} && \text{(A.12)} \\
I_{i,t} &= T_{i,t} + \left(1 - \gamma_{i}^{R} - \gamma_{i}^{D}\right) \overline{I}_{i,t} && \text{(A.13)} \\
R_{i,t} &= \overline{R}_{i,t} + \gamma_{i}^{R} \overline{I}_{i,t} && \text{(A.14)} \\
D_{i,t} &= \overline{D}_{i,t} + \gamma_{i}^{D} \overline{I}_{i,t} && \text{(A.15)}
\end{aligned}
$$

With the assumption of matching function that $M(x, y) = \chi xy$, we obtain:

$$
\begin{aligned}
T_{i,t} &= \overline{S}_{i,t} \frac{\chi \overline{I}_{i,t}}{L_{i,t}} && \text{(A.16)} \\
\alpha_{j,t} &= \chi \frac{\overline{I}_{i,t}}{L_{i,t}} && \text{(A.17)}
\end{aligned}
$$

Steady State

At steady states, there is no I group, and $\alpha = 0$. So:

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

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

And labor distribution:

$$\begin{aligned}
S_i &= \sum_{j=1}^{N} S_j m_{ji}^S \\
R_i &= \sum_{j=1}^{N} R_j m_{ji}^R
\end{aligned}$$

Hat Algebra

Denote $\widehat{x}_t = x_t / x_{ss}$, then intertemporal conditions:

$$\widehat{V}_{it}^S = \exp(\frac{u_{it} - u_{i,ss}}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^S \left(\widehat{V}_{j,t+1}^S\right)^{\beta(1-\alpha_{j,t+1})} \left(\widehat{V}_{j,t+1}^I\right)^{\beta\alpha_{j,t+1}} \left(\frac{V_{j,ss}^I}{V_{j,ss}^S}\right)^{\beta\alpha_{j,t+1}} \tag{A.18}$$

$$\widehat{V}_{it}^I = \exp(\frac{u_{it}^I - u_{i,ss}^I}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^I \left(\widehat{V}_{j,t+1}^I\right)^{\beta(1-\gamma_j^R-\gamma_j^D)} \left(\widehat{V}_{j,t+1}^R\right)^{\beta\gamma_j^R} \tag{A.19}$$

$$\widehat{V}_{it}^R = \exp(\frac{u_{it} - u_{i,ss}}{\kappa}) \sum_{j=1}^{N} m_{ij,ss}^R \left(\widehat{V}_{j,t+1}^R\right)^{\beta} \tag{A.20}$$

$$m_{ijt}^{S} = \frac{m_{ij,ss}^{S}\left(\widehat{V}_{j,t+1}^{S}\right)^{\beta(1-\alpha_{j,t+1})}\left(\widehat{V}_{j,t+1}^{I}\right)^{\beta\alpha_{j,t+1}}\left(\frac{V_{j,ss}^{I}}{V_{j,ss}^{S}}\right)^{\beta\alpha_{j,t+1}}}{\sum_{k=1}^{N} m_{ik,ss}^{S}\left(\widehat{V}_{k,t+1}^{S}\right)^{\beta\left(1-\alpha_{k,t+1}\right)}\left(\widehat{V}_{k,t+1}^{I}\right)^{\beta\alpha_{k,t+1}}\left(\frac{V_{k,ss}^{I}}{V_{k,ss}^{S}}\right)^{\beta\alpha_{j,t+1}}} \tag{A.21}$$

$$m_{ijt}^{I} = \frac{m_{ij,ss}^{I}\left(\widehat{V}_{j,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)}\left(\widehat{V}_{j,t+1}^{R}\right)^{\beta\gamma_j^R}}{\sum_{k=1}^{N} m_{ik,ss}^{I}\left(\widehat{V}_{k,t+1}^{I}\right)^{\beta(1-\gamma_j^R-\gamma_j^D)}\left(\widehat{V}_{k,t+1}^{R}\right)^{\beta\gamma_j^R}} \tag{A.22}$$

$$m_{ijt}^{R} = \frac{\left(V_{j,t+1}^{R}\right)^{\beta}\left(\mu_{ij}^{R}\right)^{-1}}{\sum_{k=1}^{N}\left(V_{k,t+1}^{R}\right)^{\beta}\left(\mu_{ik}^{R}\right)^{-1}} = \frac{m_{ij,ss}^{R}\left(\widehat{V}_{j,t+1}^{R}\right)^{\beta}}{\sum_{k=1}^{N} m_{ik,ss}^{R}\left(\widehat{V}_{k,t+1}^{R}\right)^{\beta}} \tag{A.23}$$

Algorithm

1. Exogenously given $\{L_0\}$. Evaluate T large enough.
2. Take intial guess $\left\{\widehat{V}_t^{S(0)}, \widehat{V}_t^{I(0)}, \widehat{V}_t^{R(0)}\right\}_{t=1}^{T}$.
3. 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 (A.8)(A.9)(A.10)(A.11)(A.12)(A.13)(A.14)(A.15)(A.16)(A.17)(A.21)(A.22)(A.23)
   (b) In each period t from 1 to T, solve for temporary allocation w_t from (A.1)(A.2), thus get P_t and u_t.
   (c) Using (A.18)(A.19)(A.20) to solve backward to get $\left\{\widehat{V}^{S(k)\prime}, \widehat{V}_t^{I(k)\prime}, \widehat{V}_t^{R(k)\prime}\right\}_{t=1}^{T}$
4. 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 $\left\{\widehat{V}_t^{S(k+1)}, \widehat{V}_t^{I(k+1)}, \widehat{V}_t^{R(k+1)}\right\} = \left\{\widehat{V}_t^{S(k)}, \widehat{V}_t^{I(k)}, \widehat{V}_t^{R(k)}\right\}$.

A.2 Extensions

A.2.1 Model with Vaccination

Governments are rolling out vaccination programs which have been shown to be effective in reducing infection (Chen et al. 2021). In a world with vaccination, 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 migrating to region j will have a chance of δ_j directly transferring to type R. Now the Bellman equation for group S people becomes:

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

Correspondingly, when calculating next period demographic dynamics, a fraction δ_i of the susceptible will transform into the recovered:

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

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

Therefore, vaccination can sharply slow down the spread of pandemics directly. If $\delta_i = 0$, then model collapses into our benchmark.

A.2.2 Model with Testing

It is more complicated when considering testing, because now comes the problem of identification. Testing only allows the testees to identify their current types. Those not tested may be susceptible, infected or recovered. Even if an agent is tested negative for now, it does not mean he/she would not get infected in the future. Moreover, agents could be either susceptible or have recovered from asymptomatic infection. These people will not get infected any more but they still think they may.

We address these issues by redefining groups. Suppose only a δ fraction of people get tested randomly in each period and are able to identify the current status. With a little bit abuse of notation, let $\alpha_{i,t}$ denote the fraction of untested people that actually are infected at the end of period t, which is also the fraction of tested people getting infected according to the independence across individuals. Suppose $U^S_{i,t}$ now denotes the value function of the tested people that are not infected, $U^N_{i,t}$ the untested people, $U^I_{i,t}$ the tested people that are infected, and $U^R_{i,t}$ those tested people getting infected but have recovered. Then we have

$$U^S_{i,t} = u_{i,t} + \max_j \left\{ \beta E_t \left[\delta \left((1-\alpha_{j,t+1}) U^S_{j,t+1} + \alpha_{j,t+1} U^I_{j,t+1} \right) + (1-\delta) U^N_{j,t+1} \right] - \widetilde{\mu}_{ijt} + \varepsilon_{ij,t} \right\}$$

$$\begin{aligned} U^N_{i,t} = u_{i,t} + \max_j \Big\{ \beta E_t \Big[& \alpha_{i,t} \left((1-\gamma^D_j)\delta U^I_{j,t+1} + (1-\gamma^D_j)(1-\delta) U^N_{j,t+1} + \gamma^D_j U^D \right) \\ & + (1-\alpha_{i,t}) \left(\delta(1-\alpha_{j,t+1}) U^S_{j,t+1} + (1-\delta) U^N_{j,t+1} + \delta \alpha_{j,t+1} U^I_{j,t+1} \right) \Big] - \widetilde{\mu}_{ijt} + \varepsilon_{ij,t} \Big\} \end{aligned}$$

Now group S includes all tested and uninfected people from previous period's S group and N group. Some untested people actually have been infected and died this period. Those entering I group this period only stands for the population who get both infected

and tested.Therefore, the law of motion for each group can be written as:

$$
\begin{aligned}
S_{i,t} &= \delta(1-\alpha_{i,t})(\overline{S}_{i,t}+\overline{N}_{i,t}) \\
N_{i,t} &= (1-\delta)(\overline{S}_{i,t}+\overline{N}_{i,t}) - \gamma_i^D \sum_{j=1}^{N} \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1} \\
I_{i,t} &= \delta\alpha_{i,t}(\overline{S}_{i,t}+\overline{N}_{i,t}) + (1-\gamma_i^D-\gamma_i^R)\overline{I}_{i,t} \\
R_{i,t} &= \overline{R}_{i,t} + \gamma_i^R \overline{I}_{i,t} \\
D_{i,t} &= \overline{D}_{i,t} + \gamma_i^D(\overline{I}_{i,t} + \sum_{j=1}^{N} \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1})
\end{aligned}
$$

The infected people can be decomposed into these groups: the newly infected, those untested but infected from the origin region, those died, tested and infected, and untested but infected from the perspective of results.

$$
\alpha_{i,t}N_{i,t} + \delta\alpha_{i,t}(\overline{S}_{i,t}+\overline{N}_{i,t}) = (1-\gamma_i^D)(\overline{I}_{i,t} + \sum_{j=1}^{N} \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1}) + M(A_{i,t}, B_{i,t})
$$

where $A_{i,t}$ and $B_{i,t}$ denote the uninfected and infected density in reality, no matter identified or not by the test. If we further denote γ_i^S as the self-healing rate, $H^N_{i,t}$ as the number of untested people who already self heal and $H^S_{i,t}$ as the number of tested people who self heal, then we have:

$$
\begin{aligned}
(\overline{S}_{i,t}+\overline{N}_{i,t}+\overline{I}_{i,t}+\overline{R}_{i,t})A_{i,t} &= \sum_{j=1}^{N}((1-\alpha_{j,t-1})N_{j,t-1} - H^N_{j,t-1})m^N_{ji,t-1} + \sum_{j=1}^{N}(S_{j,t-1} - H^S_{j,t-1})m^S_{ji,t-1} \\
(\overline{S}_{i,t}+\overline{N}_{i,t}+\overline{I}_{i,t}+\overline{R}_{i,t})B_{i,t} &= \sum_{j=1}^{N} \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1} + \overline{I}_{i,t}
\end{aligned}
$$

$$
\begin{aligned}
H^N_{i,t} &= (1-\delta)\sum_{j=1}^{N}(H^N_{j,t-1}m^N_{ji,t-1} + H^S_{j,t-1}m^S_{ji,t-1} + \gamma_i^S \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1}) \\
H^S_{i,t} &= \delta\sum_{j=1}^{N}(H^N_{j,t-1}m^N_{ji,t-1} + H^S_{j,t-1}m^S_{ji,t-1} + \gamma_i^S \alpha_{j,t-1} N_{j,t-1} m^N_{ji,t-1})
\end{aligned}
$$

It is worth noticing that when $\delta = 1$, i.e. full testing, the testing model collapse into our benchmark as well. Obviously, asymmetric information disables population from identifying their true types and will cause further welfare loss. Model with testing captures more reality, but contains more parameters and requires a lot more computation.

A.3 Additional Tables and Figures

Table A.1: The estimation of trade elasticity across US states

	(1)	(2)	(3)	(4)
VARIABLES	Trade share in value		Trade share in weight	
	OLS			
Ln(Income ratio)	-3.239**	-2.961***	-2.917**	-2.061***
	(-2.365)	(-5.190)	(-2.044)	(-3.800)
Ln(Distance)		-1.121***		-1.513***
		(-30.713)		(-36.072)
Border		0.358***		0.383***
		(6.527)		(4.735)
Observations	2,601	2,601	2,601	2,601
R-squared	0.042	0.857	0.010	0.779
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes

Notes: This table estimates the parameter of trade elasiticity across US states in equation (3). The numbers in the parentheses are robust t-statistics with standard error clustered at the state-pair level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

Table A.2: The estimation of migration elasticity across US states

	(1)	(2)	(3)	(4)
VARIABLES	Migration share			
	OLS			
Real income at next period	-0.184***	-0.184***	-0.262***	-0.323***
	(-4.749)	(-4.855)	(-5.203)	(-6.711)
Ln(Distance)		0.286***		0.587***
		(5.116)		(9.881)
Border		-0.897***		-0.564***
		(-7.449)		(-4.598)
Observations	10,200	10,200	10,200	10,200
R-sqaured	0.002	0.014	0.027	0.045
Destination FE	No	No	Yes	Yes

Notes: This table estimates the parameter of migration elasiticity across US states in equation (3). The numbers in the parentheses are robust t-statistics with standard error clustered at the state-pair level. Significance levels are indicated by *, **, *** at 0.1, 0.05 and 0.01, respectively.

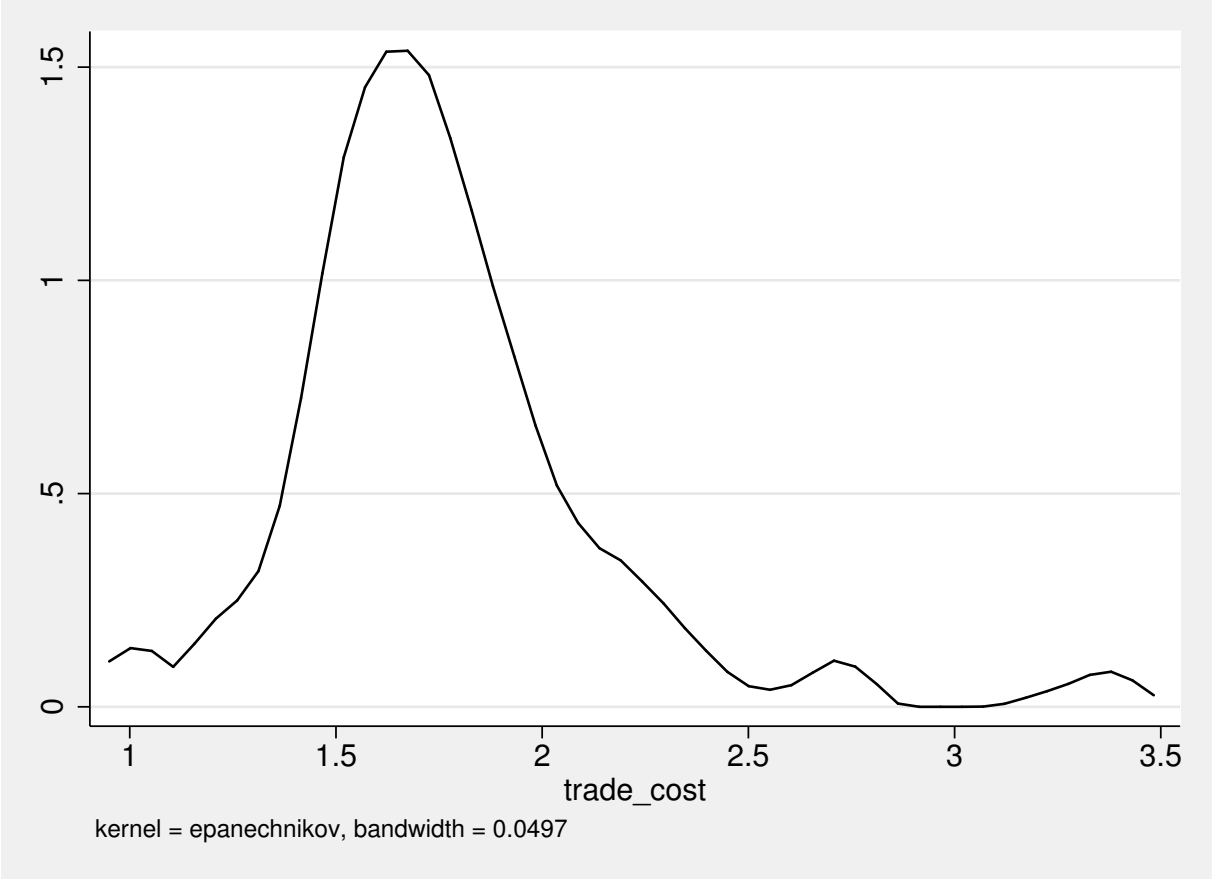

Figure A1: Distribution of estimated bilateral trade costs across US states

Notes: The figure plots the distribution of estimated bilateral trade costs, using data from 2017 Commodity Flow Survey (CFS).

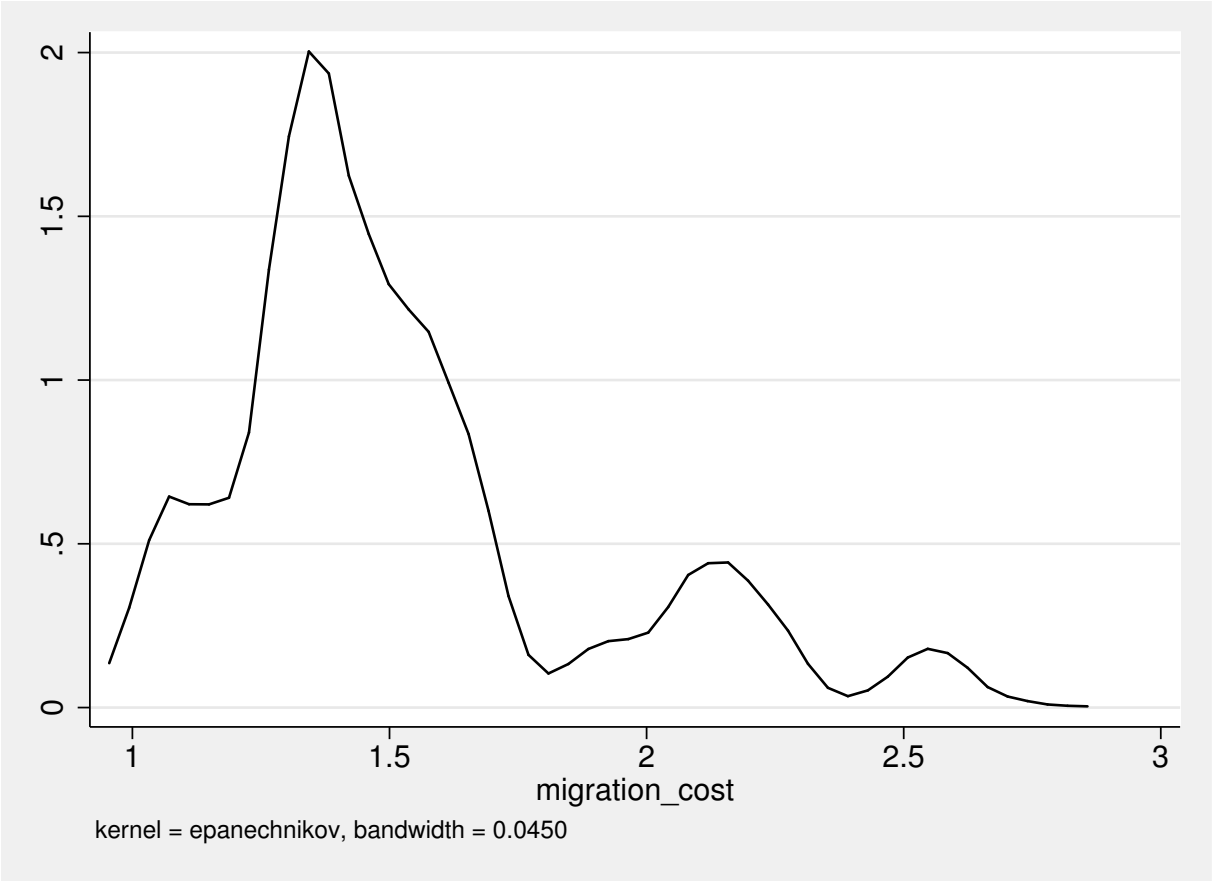

Figure A2: Distribution of estimated migration costs across US states

Notes: The figure plots the distribution of estimated bilateral migration costs, using data from IPMUS-USA database during 2014 to 2018.