Internet Appendix

A Model

A.1 Assumptions

We denote some quantities before our assumption. For $i = 1, \dots, N$, $t = 1, \dots, T$, $\delta_{N_0,N_1} = \min\{\sqrt{N_0}, \sqrt{N_1}\}$, and $K = \dim(\beta)$,

$$\mathbb{V}_{\theta,t} = \frac{\delta^2_{N_0,N_1}}{N_0} \bar{\Lambda}'_{\mathcal{T}} \left(\frac{\Lambda'\Lambda}{N}\right)^{-1} \Gamma_t \left(\frac{\Lambda'\Lambda}{N}\right)^{-1} \bar{\Lambda}_{\mathcal{T}} + \frac{\delta^2_{N_0,N_1}}{N_1} \sigma^2_e.$$

$$\widehat{\mathbb{V}}_{\theta,t} = \frac{\delta^2_{N_0,N_1}}{N_0} \bar{\Lambda}'_{\mathcal{T}} \left(\frac{\widehat{\Lambda}'\widehat{\Lambda}}{N}\right)^{-1} \Gamma_t \left(\frac{\widehat{\Lambda}'\widehat{\Lambda}}{N}\right)^{-1} \bar{\Lambda}_{\mathcal{T}} + \frac{\delta^2_{N_0,N_1}}{N_1} \widehat{\sigma}^2_e.$$

$$\mathbb{V}_{\theta,j,t} = \frac{\delta^2_{N_0,N_1}}{T_0} F'_t \left(\frac{F'F}{T}\right)^{-1} \Phi_j \left(\frac{F'F}{T}\right)^{-1} F_t + \frac{\delta^2_{N_0,N_1}}{N_0} \bar{\Lambda}'_j \left(\frac{\Lambda'\Lambda}{N}\right)^{-1} \Gamma_t \left(\frac{\Lambda'\Lambda}{N}\right)^{-1} \bar{\Lambda}_j + \delta^2_{N_0,N_1} \sigma^2_{e,t}.$$

$$\widehat{\mathbb{V}}_{\theta,j,t} = \frac{\delta^2_{N_0,N_1}}{T_0} F'_t \left(\frac{F'F}{T}\right)^{-1} \Phi_j \left(\frac{F'F}{T}\right)^{-1} F_t + \frac{\delta^2_{N_0,N_1}}{N_0} \bar{\Lambda}'_j \left(\frac{\widehat{\Lambda}'\widehat{\Lambda}}{N}\right)^{-1} \Gamma_t \left(\frac{\widehat{\Lambda}'\widehat{\Lambda}}{N}\right)^{-1} \bar{\Lambda}_j + \delta^2_{N_0,N_1} \widehat{\sigma}^2_{e,t}.$$

$$\widehat{\sigma}^2_e = \frac{1}{TN_0 - r(T + N_0) + r^2 - K} \sum_{i \le N_0} \sum_{t=1}^{T} \widehat{e}^2_{it}.$$

$$\widehat{\sigma}^2_{e,t} = \frac{1}{N-1} \sum_{i \ne j}^{N-1} \widehat{e}^2_{it}.$$

Assumption 1. *There exists a constant M such that*

1. *(Factors)*
   (a) $E\left\|F^0_t\right\|^4 \le M \le \infty$,
   (b) $T^{-1}\sum_{t=1}^{T} F^0_t F^{0\prime}_t \xrightarrow{p} \Sigma_F$, *where Σ_F is a positive definite matrix.*
2. *(Factor Loadings)*
   (a) $\left\|\Lambda^0_i\right\| \le M$,
   (b) $N^{-1}\sum_{i=1}^{N} \Lambda^{0\prime}_i \Lambda^0_i \xrightarrow{p} \Sigma_\Lambda$, *where Σ_Λ is a positive definite matrix,*
   (c) *The eigenvalues of $\Sigma_F \Sigma_\Lambda$ are distinct.*

3. *(Error terms):*
   (a) $e_{it} \sim \mathcal{N}(0, \sigma^2_{e,t})$, $\forall i$,
   (b) $E\left(N^{-1}\sum_{i=1}^{N} e_{it}e_{is}\right) = \gamma_N(s,t)$, *where* $\sum_{t=1}^{T} |\gamma_N(s,t)| \le M, \forall s$,
   (c) $E(e_{it}e_{jt}) = \tau_{ij,t}$, *where* $|\tau_{ij,t}| \le |\tau_{ij}|$ *for some* τ_{ij}, *and* $\sum_{j=1}^{N} |\tau_{ij}| \le M, \forall i$,
   (d) $E(e_{it}e_{js}) = \tau_{ij,st}$ *and* $(NT)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} |\tau_{ij,ts}| < M$,
   (e) *For every* (t,s), $E\left(N^{-1/2}\sum_{i=1}^{N} [e_{is}e_{it} - E(e_{is}e_{it})]^4\right) \le M$,
   (f) $E\left(N^{-1}\sum_{i=1}^{N} \left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} F^0_t e_{it}\right\|^2\right) \le M$.
4. *(Central Limit Theorems)*
   (a) *For each i, as $N \to \infty$,* $N^{-1/2}\sum_{i=1}^{N} \Lambda^0_i e_{it} \xrightarrow{d} N(0, \Gamma_t)$,
   (b) *For each t, as $T \to \infty$,* $T^{-1/2}\sum_{t=1}^{T} F^0_t e_{it} \xrightarrow{d} N(0, \Phi_i)$.

Assumption 2. *Order conditions:*

1. *The conditions $T \times N_0 > r(T + N_0)$ and $N \times T_0 > r(N + T_0)$ are satisfied for any N, T, N_0, T_0,*
2. $\frac{\sqrt{N}}{\min\{N_0,T_0\}} \to 0$ *as $N_0, N \to \infty$ and $T_0, T \to \infty$,*
3. $\frac{\sqrt{T}}{\min\{N_0,T_0\}} \to 0$ *as $N_0, N \to \infty$ and $T_0, T \to \infty$.*

Assumption 3. *Factors and factor loadings:*

1. $\frac{\Lambda^{0\prime}_0 \Lambda^0_0}{N_0} \xrightarrow{p} \Sigma_{\Lambda,0} > 0$, $\frac{\Lambda^{0\prime}_m \Lambda^0_m}{N_m} \xrightarrow{p} \Sigma_{\Lambda,m} > 0$, $\frac{1}{\sqrt{N_0}}\sum_{i=1}^{N_0} \Lambda^0_i e_{it} \xrightarrow{d} N(0, \Gamma_{0t})$,
2. $\frac{F^{0\prime}_0 F^0_0}{T_0} \xrightarrow{p} \Sigma_{F,0} > 0$, $\frac{F^{0\prime}_m F^0_m}{T_m} \xrightarrow{p} \Sigma_{F,m} > 0$, $\frac{1}{\sqrt{T_0}}\sum_{s=1}^{T_0} F^0_s e_{is} \xrightarrow{d} N(0, \Phi_{0i})$.

Part (1) and (2) of Assumption 1 impose identification conditions on the factor structure and the low-rank component. Part (3) of Assumption 1 assumes the error terms to be weakly correlated in cross-sectional and time-series dimensions. Part (4) of Assumption 1 assumes the large sample distribution of factor estimates. Assumption 2 imposes that the order condition holds for both TALL and WIDE block matrices. Assumption 3 requires the subsample matrix's identification condition such that the subsample moment matrices for the factor and factor loadings to be positive definite.

We further explain the common factors may capture differential time trends with the following

example. Consider the special factor and factor loading in the form of

$$F_t = \begin{bmatrix} 1 \\ \xi_t \end{bmatrix} \quad \text{and} \quad \Lambda_i = \begin{bmatrix} \alpha_i \\ 1 \end{bmatrix} \tag{A.1}$$

for all i and t. Then, Equation (A.1) can be further rewritten as

$$\begin{aligned} Y_{it} &= \theta_{it}\mathcal{D}_{it} + X'_{i,t-m}\beta + Z'_{i,t}\gamma + \Lambda'_i F_t + e_{it} \\ &= \theta_{it}\mathcal{D}_{it} + X'_{i,t-m}\beta + Z'_{i,t}\gamma + \alpha_i + \xi_t + e_{it}, \end{aligned}$$

where the interactive effect $\Lambda'_i F_t$ becomes the individual effects α_i and the time effects ξ_t entering our model additively. Consequently, we believe that our common factors effectively capture the differential time trends across states.

A.2 Estimation

In this section, we present the computation algorithm to estimate both average treatment effects and individual treatment effects that allow the pre-treatment period for each state in the treated group to be different.

Algorithm. We introduce the following algorithm to calculate the average treatment effects and individual treatment effects:

1. We estimate the coefficient of β and γ using the observations in the control group with interactive fixed effects method (Bai, 2009). We then estimate the residual R_{it} for the whole observations where $R_{it} \equiv Y_{it} - X'_{i,t-14}\widehat{\beta} - Z'_{it}\widehat{\gamma}$.
2. We estimate the latent factor F using the observations in the TALL block matrix of residual R_{TALL} with asymptotic principal components method.
3. We estimate the factor loadings Λ using two-step approach:
   (a) For $i = 1, \dots, N$, create the new matrix $\widetilde{F}_i = [\mathbb{I}_i \ F_i]$, where F_i is the submatrix of F with the row from $1 : T_{0,i}$ and $\mathbb{I}_i$ is $T_{0,i} \times 1$ matrix of one.
   (b) For $i = 1, \dots, N$, estimate the $\widetilde{\Lambda}_i = (\widetilde{F}'_i\widetilde{F}_i)^{-1}\widetilde{F}'_i R_{it}$.

4. We estimate the missing value with $\widehat{Y}_{it}(0) = X'_{it}\widehat{\beta} + \widetilde{C}_{it}$, where $\widetilde{C} = [\mathbb{I}_T \ \widetilde{F}] \times \widetilde{\Lambda}'$, $\mathbb{I}_T$ is $T \times 1$ matrix of one.
5. We compute the individual treatment effects $\widehat{\theta}_{it} = Y_{it}(1) - \widehat{Y}_{it}(0), \ i \in \mathcal{T}, \ t = T_{0,i} + 1, \dots, T$.
6. We compute the average treatment effects on treated $\widehat{\theta}_t = \frac{1}{N_1}\sum_{i \in \mathcal{T}} \widehat{\theta}_{it}$.
7. We compute the average treatment effects on treated for the individuals in the group j, $\widehat{\theta}_{j,t} = \frac{1}{N_j}\sum_{i \in \mathcal{J}} \widehat{\theta}_{it}$, where $i \in \mathcal{J}$, and N_j is the sample size of the group j.

Note that in Steps 2 and 3, we estimate both latent factors F and factor loadings Λ for both TALL block and WIDE block with asymptotic principal components method and the number of factors for these two blocks r_{TALL} and r_{WIDE} can be consistently estimated using the method developed in Bai and Ng (2002, 2019). When the number of factors of the two blocks does not coincide, we choose the number $r = \max(r_{\text{TALL}}, r_{\text{WIDE}})$ and use r for the estimation in Steps 2 and 3. Therefore, we can treat the number of factors known for the TALL and WIDE blocks in the estimation procedure.

A.3 Large Sample Properties

In this section, we establish the asymptotic distribution results for the average treatment effects on treated $\widehat{\theta}_t$ and the heterogenous treatment effects $\widehat{\theta}_{j,t}$ for a group of j that belongs to the treated group $\mathcal{T}$ under the assumptions in Appendix A.1.

Proposition 1. *Let Assumption 1, Assumption 2, and Assumption 3 hold. As $N_0, T_0, N_1 \to \infty$, we have:*

1. *The average treatment effects on treated $\widehat{\theta}_t$ is asymptotically normal:*

$$\delta_{N_0,N_1}\left(\frac{\widehat{\theta}_t - \theta_t}{\sqrt{\mathbb{V}_{\theta,t}}}\right) \xrightarrow{d} N(0,1).$$

2. *The asymptotic variance $\widehat{\sigma}^2_{e,t}$ can be consistently estimated by:*

$$\widehat{\sigma}^2_{e,t} = \widehat{\mathbb{V}}_{\theta,t}/\delta_{N_0,N_1}.$$

Proposition 2. *Let Assumption 1, Assumption 2, and Assumption 3 hold. As $N_0, T_0, N_1, N_j \to \infty$, we have:*

1. *The heterogeneous treatment effects on group j, $\widehat{\theta}_j$ is asymptotically normal:*

$$\delta_{N_0,N_j}\left(\frac{\widehat{\theta}_j-\theta_j}{\sqrt{\mathbb{V}_{\theta,j}}}\right)\xrightarrow{d} N(0,1).$$

   where $\theta_j = \frac{1}{T_1}\sum_{s>T_0}\theta_{js}$...

2. *The asymptotic variance $\widehat{\sigma}_j^2$ can be consistently estimated by:*

$$\widehat{\sigma}_j^2 = \widehat{\mathbb{V}}_{\theta,j}/\delta_{N_0,N_j}.$$

Our estimation procedure provides a robust and reliable estimate and inference for treatment effects without relying on restrictive distributional assumptions for error terms or parallel trend assumptions for average outcomes of both treated and control units in the absence of treatment.

In the context of social distancing policies, which are endogenously enacted by health authorities, recent research has employed either a difference-in-differences approach or an instrumental variable approach to address endogeneity concerns. While these studies confirm the effectiveness of social distancing, they may overlook crucial observed and unobserved factors that can confound the relationship. Specifically, there exists a reverse causality between social distancing and the growth in new cases, as state governments introduce restrictions in response to increases in cases and deaths. Consequently, variations in policies may reflect local case numbers rather than the true causal effect of the policy. Additionally, the voluntary precaution effect, whereby people take precautions before government restrictions are announced in response to publicized COVID outbreaks, can lead to biased difference-in-difference estimates that find a spurious negative relationship between social distancing requirements and growth in new cases. Furthermore, the anticipation effect of governments announcing policies ahead of time may change residents' behavior in response to the policy information itself. Lastly, spillover effects from other states can affect the effectiveness of one state's social distancing measures on the growth of new cases in another state, biasing the difference-in-difference estimates towards zero.

In contrast, our factor model approach consistently estimates the average and heterogeneous treatment effects by identifying the causal relationship between social distancing and reductions in COVID-19 infections while accounting for factors that influence virus transmission over time and

across states. Our method can trace the policy's effect on each unit before and after it takes effect, effectively alleviating the endogeneity concerns discussed above. By considering pre-treatment and post-treatment outcomes, we capture the potential reverse causality effect on outcomes in the days leading to the policy and the voluntary precaution effect that improves outcomes for treated units before the policy implementation. Moreover, our approach considers the difference between treatment and control groups using different dimensions by controlling for both observed confounders in the regression and interactive fixed effects that remove the partial correlation that may vary across units and over time (Bai and Ng, 2021). This comprehensive approach ensures that our estimates are unbiased and accurately reflect the true causal impact of social distancing policies on the spread of COVID-19.

B Proofs

B.1 Proofs of Proposition 1

The average treatment effects on treated is defined as:

$$\theta_t = \frac{1}{N_1}\left[\sum_{i\in\mathcal{T}} Y_{it}(1) - Y_{it}(0)\right], \tag{B.1}$$

It follows that

$$\begin{aligned} Y_{it}(1) - \widehat{Y}_{it}(0) &= \theta_{it} + X_{it}'(\beta-\widehat{\beta}) + Z_{it}'(\gamma-\widehat{\gamma}) + C_{it} - \widehat{C}_{it} + e_{it},\\ Y_{it}(0) - \widehat{Y}_{it}(0) &= X_{it}'(\beta-\widehat{\beta}) + Z_{it}'(\gamma-\widehat{\gamma}) + C_{it} - \widehat{C}_{it} + e_{it}. \end{aligned} \tag{B.2}$$

Combining Equation (B.1) and Equation (B.2), we have

$$\begin{aligned} \widehat{\theta}_t - \theta_t &= \frac{1}{N_1}\sum_{i\in\mathcal{T}} X_{it}'\left(\beta-\widehat{\beta}\right) + \frac{1}{N_1}\sum_{i\in\mathcal{T}} Z_{it}'(\gamma-\widehat{\gamma}) + \frac{1}{N_1}\sum_{i\in\mathcal{T}}\left(C_{it}-\widehat{C}_{it}\right) + \frac{1}{N_1}\sum_{i\in\mathcal{T}} e_{it} \\ &= A_1 + A_2 + A_3 + A_4. \end{aligned} \tag{B.3}$$

By Assumption 1, we have:

$$A_4 = \frac{1}{N_1}\sum_{i\in\mathcal{T}} e_{it} = O_p(\frac{1}{\sqrt{N_1}}).$$

By Assumption 2, β is homogeneous across i and t, we apply the Theorem 1 in Bai (2009):

$$A_1 = \frac{1}{N_1}\sum_{i\in\mathcal{T}} X_{it}'\left(\beta-\widehat{\beta}\right) = O_p\left(\frac{1}{\sqrt{N_0T_0}}\right) \quad \text{and} \quad A_2 = \frac{1}{N_1}\sum_{i\in\mathcal{T}} Z_{it}'(\gamma-\widehat{\gamma}) = O_p\left(\frac{1}{\sqrt{N_0T_0}}\right).$$

We decompose the term A_3 as

$$\begin{aligned} A_3 =& \frac{1}{N_1}\sum_{i\in\mathcal{T}}\left(\widehat{C}_{it} - C_{it}\right) \\ =& F_t'\left(\frac{F'F}{T}\right)^{-1}\mathbf{B}_F\frac{1}{T_0N_1}\left(\sum_{i\in\mathcal{T}}\sum_{s=1}^{T_{0,i}} F_s e_{is}\right) + \bar{\Lambda}_{\mathcal{T}}\left(\frac{\Lambda'\Lambda}{N}\right)^{-1}\mathbf{B}_\Lambda\frac{1}{N_0}\sum_{k=1}^{N_0}\Lambda_k e_{kt} \\ &+ O_p\left(\delta_{N_0,T_0}^{-2}\right) \\ =& A_{31} + A_{32} + A_{33}, \end{aligned}$$

where is $\mathbf{B}_F$ defined as $\frac{T_o}{T}I_r + \frac{T_m}{T}\left(\frac{F_m^{0\prime}F_m^0}{T_m}\right)\left(\frac{F_o^{0\prime}F_o^0}{T_o}\right)^{-1}$ and $\mathbf{B}_\Lambda$ is defined as $\frac{N_o}{N}I_r + \frac{N_m}{N}\left(\frac{\Lambda_m^{0\prime}\Lambda_m^0}{N_m}\right)\left(\frac{\Lambda_o'\Lambda_o}{N_o}\right)^{-1}$.

By Assumption 1,

$$A_{31} = O_p\left(\frac{1}{\sqrt{T_0N_1}}\right).$$

We have:

$$\widehat{\theta}_t - \theta_t = \bar{\Lambda}_{\mathcal{T}}'\left(\frac{\Lambda'\Lambda}{N}\right)^{-1}\mathbf{B}_\Lambda\frac{1}{N_0}\sum_{k=1}^{N_0}\Lambda_k e_{kt} + \frac{1}{N_1}\sum_{i\in\mathcal{T}} e_{it} + O_p\left(\frac{1}{\sqrt{T_0N_0}}\right) + O_p\left(\frac{1}{\sqrt{T_0N_1}}\right) + O_p\left(\delta_{N_0,T_0}^{-2}\right). \tag{B.4}$$

$\square$

B.2 Proofs of Proposition 2

The average treatment effects on treated for group j are defined as:

$$\theta_{j,t} = \frac{1}{N_j}\left[\sum_{i\in\mathcal{J}} Y_{it}(1) - Y_{it}(0)\right], \tag{B.5}$$

It follows that

$$\begin{aligned} Y_{it}(1) - \widehat{Y}_{it}(0) &= \theta_{it} + X_{it}'(\beta-\widehat{\beta}) + C_{it} - \widehat{C}_{it} + e_{it},\\ Y_{it}(0) - \widehat{Y}_{it}(0) &= X_{it}'(\beta-\widehat{\beta}) + C_{it} - \widehat{C}_{it} + e_{it}. \end{aligned} \tag{B.6}$$

The proof of Proposition 2 are similar to that of Proposition 1, and hence, we only highlight those different parts. Recall Equation (B.5) and Equation (B.6) and consider the estimation of treatment effect on a single unit j:

$$\widehat{\theta}_{jt} - \theta_{jt} = X_{jt}'\left(\beta - \widehat{\beta}\right) + Z_{jt}'\left(\gamma - \widehat{\gamma}\right) + \left(C_{jt} - \widehat{C}_{jt}\right) + e_{jt} \tag{B.7}$$

$$\begin{aligned} &= X_{jt}'\left(\beta - \widehat{\beta}\right) + Z_{jt}'\left(\gamma - \widehat{\gamma}\right) \\ &\quad - F_t'\left(\frac{F'F}{T}\right)^{-1}\mathbf{B}_F\frac{1}{T_0}\left(\sum_{s=1}^{T_{0,j}} F_s e_{js}\right) - \Lambda_j'\left(\frac{\Lambda'\Lambda}{N}\right)^{-1}\mathbf{B}_\Lambda\frac{1}{N_0}\sum_{k=1}^{N_0}\Lambda_k e_{kt} + e_{jt} + O_p\left(\delta_{N_0,T_0}^{-2}\right). \end{aligned} \tag{B.8}$$

Then, it is similar to the proof of Proposition 1 to verify Proposition 2.

□

C Robustness Results

In this section, we conduct two robustness checks to further validate our findings. The first check uses state-level data on non-essential business closures as a measure of social distancing, while the second extends the time window beyond April 20 to examine whether cases rebounded after the lifting of stay-at-home orders.

For the first robustness test, we create a social distancing policy index $D_{i,t}$ for each state, which equals 1 if non-essential business closures are announced in state i on day t and 0 otherwise. Using non-essential business closures as the measure for social distancing, we have $N_0 = 24$, $T_0 = 28$, $N = 55$, and $T = 61$. Therefore, Assumption 2f, which states that $\frac{\sqrt{N}}{\min\{N_0,T_0\}} \to 0$ and $\frac{\sqrt{T}}{\min\{N_0,T_0\}} \to 0$ as $N_0, N \to \infty$ and $T_0, T \to \infty$, is satisfied for the state-level data we considered.

Panel (a) of Figure C.1 presents the dynamic effect of the state-level social distancing policy on the growth rate of infections with 95% confidence intervals. The social distancing policy is associated with a 2.30% reduction in the weekly growth rate of new cases after 7 days, a 7.06% reduction after 14 days, a 10.57% reduction after 21 days, and a 15.18% reduction after 28 days. Panel (b) of Figure C.1 displays the dynamic effect of state-level social distancing policies on the growth rate of COVID-related deaths with 95% confidence intervals. The policy starts to reduce the growth rate of COVID-related deaths from the fifth day after its announcement and is associated with a 1.49% reduction in the growth of weekly COVID-related deaths after 7 days, a 3.68% reduction after 14 days, a 6.84% reduction after 21 days, and an 8.74% reduction after 28 days.

For the second robustness test, we extended our main analysis to include data up to May 7, 2020, allowing us to capture the dynamics of infection rates after the relaxation of social distancing measures. As presented in Figure C.2, our findings indicate that there was indeed a rebound in cases after the lifting of stay-at-home orders. This provides further evidence supporting the effectiveness of these policies in controlling the spread of the virus. The extended analysis strengthens our study by addressing the potential longer-term impacts of stay-at-home orders and their subsequent relaxation.

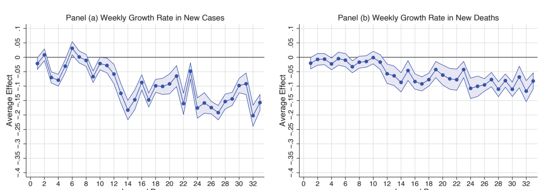

Figure C.1: Average Treatment Effects of Social Distancing on the Growth Rate of New Infections and COVID-Related Deaths Using Non-Essential Business Closures as the Policy Variable.

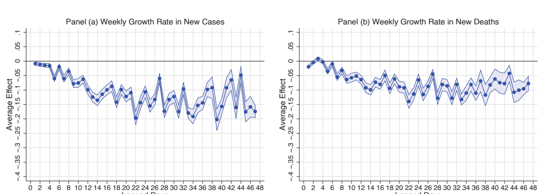

Figure C.2: Average Treatment Effects of Social Distancing on Growth Rate of New Infections and COVID-Related Deaths Using an Extended Time Window.

D SEIR Model

Following the work of Kucharski et al. (2020), we construct a modified SEIR model to investigate the transmission dynamics of COVID-19. Our model divides the total population into four distinct classes: susceptible, exposed, infectious, and removed (individuals who are isolated or recovered).

We denote S_{it} as the number of susceptible individuals in state i at time t, E_{it}^1 and E_{it}^2 as the number of individuals in state i during the first and second incubation periods, respectively, I_{it}^1 and I_{it}^2 as the number of individuals in state i during the first and second infectious periods, respectively, R_{it} as the number of recovered individuals in state i at time t, Q_{it} as the number of symptomatic individuals in state i yet to be reported at time t, C_{it} as the cumulative number of confirmed cases in state i at time t, and D_{it} as the cumulative number of confirmed deaths in state i at time t. We assume that all individuals become symptomatic and infectious simultaneously and that the population in state i, N_i, remains constant. The law of motions for state i is derived as follows:

$$\begin{aligned} S_{it+1} &= S_{it} - \beta_t S_{it}\left[I_{it}^1 + I_{it}^2\right]/N_i \\ E_{it+1}^1 &= E_{it}^1 + (1-f)\beta_t S_{it}\left[I_{it}^1 + I_{it}^2\right]/N_i - 2\sigma E_{it}^1 \\ E_{it+1}^2 &= E_{it}^2 + 2\sigma E_{it}^1 - 2\sigma E_{it}^2 \\ I_{it+1}^1 &= I_{it}^1 + 2\sigma E_{it}^2 - 2\gamma I_{it}^1 \\ I_{it+1}^2 &= I_{it}^2 + 2\gamma I_{it}^1 - 2\gamma I_{it}^2 \\ Q_{it+1} &= Q_{it} + 2\sigma E_{it}^2 e^{-\gamma\kappa} - \kappa Q_{it} \\ C_{it+1} &= C_{it} + (1-\kappa)\eta Q_{it} \\ D_{it+1} &= D_{it} + \kappa\eta Q_{it} \end{aligned}$$

where β_t represents the transmission rate at time t, σ denotes the probability of becoming symptomatic, γ is the probability of isolation, κ is the probability of death conditional on being infectious, η is the probability of recovery, and f is the fraction of cases that travel to other states.

Furthermore, we model the law of motions in state j as follows:

$$
\begin{aligned}
E^1_{jt+1} &= E^1_{jt} + f\beta_t S_{it}\left[I^1_{it} + I^2_{it}\right]/N_i - 2\sigma E^1_{jt}\\
E^2_{jt+1} &= E^2_{jt} + 2\sigma E^1_{jt} - 2\sigma E^2_{jt}\\
I^1_{jt+1} &= I^1_{jt} + 2\sigma E^2_{jt} - 2\gamma I^1_{jt}\\
I^2_{jt+1} &= I^2_{jt} + 2\gamma I^1_{jt} - 2\gamma I^2_{jt}\\
Q_{jt+1} &= Q_t + 2\sigma E^2_{jt}\mathrm{e}^{-\gamma\kappa} - \kappa Q_{jt}\\
C_{jt+1} &= C_{jt} + (1-\kappa)\eta Q_{jt}\\
D_{jt+1} &= D_{jt} + \kappa\eta Q_{jt}
\end{aligned}
$$

where E^1_{jt} and E^2_{jt} represent the number of individuals who traveled from state i to state j during the first and second incubation periods, respectively, I^1_{jt} and I^2_{jt} denote the number of individuals who traveled from state i to state j during the first and second infectious periods, respectively, Q_{jt} is the number of symptomatic cases among travelers from state i to state j yet to be reported at time t, C_{jt} is the cumulative number of confirmed cases in state j at time t, and D_{jt} is the cumulative number of confirmed deaths in state j at time t.

Our model operates under the assumption that the COVID-19 outbreak in the United States originated from a single infectious case in New York, after which the entire population became susceptible (Acemoglu et al., 2021, 2024; Ferretti et al., 2020). We model the transmission rate using a geometric random walk process and employ the branching process to estimate the unknown parameters of interest. Our model incorporates delays in symptom appearance and case reporting by including transitions between reporting states and disease states (Manski and Molinari, 2021; Viner et al., 2020). Furthermore, we explicitly account for the uncertainty in case and death observation by modeling the observed dynamics of new cases, symptom appearance of new cases, and reported confirmation of new cases (Springborn et al., 2016).

E Supplementary Details

We show the detailed statewide stay-at-home orders timeline in Table E.1.

Table E.1: Statewide Stay-at-Home Orders Timeline. The data, collected from state government websites, is updated as of April 20, 2020. The Governor of Pennsylvania initially issued stay-at-home orders for some counties on March 23, and then extended these orders statewide on April 1.

Statewide Stay-at-Home Orders Timeline		
State	Date Announced	Effective Date
Alabama	April 3	April 4
Alaska	March 27	March 28
Arizona	March 30	March 31
Arkansas	-	-
California	March 19	March 19
Colorado	March 26	March 26
Connecticut	March 20	March 23
Delaware	March 22	March 24
District of Columbia	March 30	April 1
Florida	April 1	April 3
Georgia	April 2	April 3
Hawaii	March 23	March 25
Idaho	March 25	March 25
Illinois	March 20	March 21
Indiana	March 23	March 24
Iowa	-	-
Kansas	March 28	March 30
Kentucky	March 22	March 26
Louisiana	March 22	March 23
Maine	March 31	April 2
Maryland	March 30	March 30

Massachusetts	March 23	March 24
Michigan	March 23	March 24
Minnesota	March 25	March 27
Mississippi	March 31	April 3
Missouri	April 3	April 6
Montana	March 26	March 28
Nebraska	-	-
Nevada	April 1	April 1
New Hampshire	March 26	March 27
New Jersey	March 20	March 21
New Mexico	March 23	March 24
New York	March 20	March 22
North Carolina	March 27	March 30
North Dakota	-	-
Ohio	March 22	March 23
Oklahoma	-	-
Oregon	March 23	March 23
Pennsylvania	March 23	April 1
Rhode Island	March 28	March 28
South Carolina	-	-
South Dakota	-	-
Tennessee	March 30	March 31
Texas	March 31	April 2
Utah	-	-
Vermont	March 24	March 24
Virginia	March 30	March 30
Washington	March 23	March 23
West Virginia	March 23	March 24
Wisconsin	March 24	March 25

Wyoming	-	-
Guam	-	-
Northern Mariana Islands	-	-
Puerto Rico	-	-
Virgin Islands	-	-