GRAPE Working Paper # 48

Local containment policies and country-wide spread of Covid-19 in the United States: An Epidemiologic Analysis

Jacek Rothert, Ryan Brady, and Michael Insler

FAME | GRAPE, 2020

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Local containment policies and country-wide spread of Covid-19 in the United States: An Epidemiologic Analysis

Jacek Rothert
U.S. Naval Academy
and FAME|GRAPE

Ryan Brady
U.S. Naval Academy

Michael Insler
U.S. Naval Academy

Abstract

We analyze spatial diffusion of new Covid-19 cases and country-wide impact of state-specific containment policies during the early months of the Covid-19 pandemic in the United States. We first use spatial econometric techniques to document direct and indirect spillovers of new infections across county and state lines, as well as the impact of individual states' lock-down policies on infections in neighboring states. We find consistent statistical evidence that new cases diffuse across county lines, holding county level factors constant, and that the diffusion across counties was affected by the closure policies of adjacent states. We then develop a spatial version of the epidemiological SIR model where new infections arise from interactions between infected people in one state and susceptible people in the same or in neighboring states. We incorporate lock-down policies into our model and calibrate the model to match both the cumulative and the new infections across the 48 contiguous U.S. states and DC. Our results suggest that, had the states with the less restrictive social distancing measures tightened them by one level, the cumulative infections in other states would be about 5% smaller. In our spatial SIR model, the spatial containment policies such as border closures have a bigger impact on flattening the infection curve in the short-run than on the cumulative infections in the long-run.

Keywords:

Diffusion, Spatial Model, Covid-19, Epidemics

JEL Classification

R15, H77, I19

Corresponding author

Jacek Rothert, jacek.rothert@gmail.com

Acknowledgements

We thank the participants at the virtual meetings of the Urban Economic Association and the Econometric Society Dehli Winter School for helpful comments. All errors are ours. Jacek Rothert acknowledges financial support of the National Science Center, Poland, Grant no. 2019/35/B/HS4/00769. The views expressed here are those of the authors and do not necessarily represent the views of the United States Naval Academy, the Department of Defense, or the Federal Government.

Published by: FAME | GRAPE
ISSN: 2544-2473

Foundation of Admirers and Mavens of Economics
ull. Mazowiecka 11/14
00-052 Warszaw
Poland

W | grape.org.pl
E | grape@grape.org.pl
TT | GRAPE_ORG
FB | GRAPE.ORG
PH | +48 799 012 202

1 Introduction

In this paper we assess the spatial diffusion of Covid-19 in the United States and the effect that state-level lock-down policies have on that diffusion during the early months of the pandemic. Our analysis is motivated by the idea that, if there were substantial spillovers of new infections between states, then the uncoordinated responses at the state level may have exacerbated the initial outbreak of the disease. But, the magnitude of those inter-region spillovers is uncertain, as is the extent to which the relatively lax policies of one state contributed to new infections in surrounding states. Indeed, as is now well-documented, the virus spread quickly throughout the United States, with notable variations in state-level policy to follow. By March 6, a majority of U.S. states had at least one confirmed case of the virus, and by March 17 the last state (West Virginia) reported its first case. While the Center for Disease Control (CDC) and and other federal entities issued guidance on appropriate measures to mitigate the spread of the virus, the final decisions regarding the timing and the extent of restrictions were made by individual states, and sometimes even counties. The first state-wide "shelter-in-place" order was issued in California on March 19, but ultimately only 24 additional states followed suit over the next two weeks. The compliance with social distancing measures also varied greatly across regions (Painter and Qiu, 2020; Simonov et al., 2020). The goal of this paper is to assess the impact of such a scattered policy response on the country-wide spread of the virus, focusing on the early months of the pandemic.

Our analysis follows a two-pronged approach. First, we estimate spatial econometric models to measure the extent of the spatial diffusion of new cases across regions in the United States. For each model the dependent variable is the seven-day average growth rate of county-level cases,

and our primary covariate is a measure of the number of restrictions put in place in the state in which the county resides. The spatial econometric results provide useful descriptive insight into the regional spread of the virus.

Second, we develop a spatial version of the standard epidemiological SIR (Susceptible-Infected-Recovered) model based on Kermack and McKendrick (1927), which has been popularized in the economics literature by Atkeson (2020b). In our model individuals can be infected by people from their own states and from other states. Those inter-state contacts endogenously create a spatial diffusion of the infections, with the speed of such diffusion depending on the model parameters that measure the relative frequency of connections across state lines, potentially altered by social distancing measures. We calibrate the model parameters by minimizing the distance between the data and the model generated series. We then use the model to simulate the impact of lock-down policies implemented in the states with the most restrictive and most lax policies.[1] Our main results in that section are twofold. First, if the individual states had the ability to restrict the travel across their borders, infections would be smaller.[2] Specifically, cutting the value of the calibrated inter-state spillover parameter by 25% results in the reduction of country-wide infections by almost 40% in the first 3 months, and by almost 7% in the long-run. Second, if the states with the more lenient lock-down policies tightened them by one level, the cumulative cases in the remaining states would be reduced by 2% in the first three months, and by more than 5% over the 21-month period.

Our analysis contributes to a large literature on the economics of Covid-19. First, we expand the empirical literature that focuses on the spatial aspects of the outbreak. A few studies analyzed drivers of spatial heterogeneity in the scope or severity of the Covid-19 pandemic: Desmet

and Wacziarg (2020) and Gerritse (2020) looked at US counties, Verwimp (2020) at Belgian municipalities, and Ginsburgh et al. (2020) at French regions. Very few papers seemed to focus on understanding the geographic spread of the infections. Kuchler et al. (2020) analyze the correlation between the growth in new cases and the degree of social connectedness with the Covid hotspots, using an aggregated data from Facebook. Cuñat and Zymek (2020) analyze geographical spread of the virus in the U.K. by incorporating individual's location and mobility decisions with the SIR model. The most closely related studies are those that focus specifically on the spatial spillovers of cases and cross-regional impact of local policies.[3] Our paper has been probably the first attempt to estimate the extent of spatial diffusion of Covid-19 in the United States during the early months of the outbreak. Our main objective is to quantify the extent of inter-state spillovers and the impact of one state's containment measures on outcomes in surrounding states, with close attention paid both to containment measures and possible non-compliance with them.

Second, our results are important for the discussion of policy coordination. It is quite well known that in the presence of inter-state spillovers, an uncoordinated policy response may lead to sub-optimal outcomes. This may be because restrictions are too lenient and the virus spreads to other regions or countries (Beck and Wagner, 2020; Rothert, 2020, 2021). It may also be because the restrictions are too harsh, and the recession engineered in one region or industry spills over to other regions or industries (Crucini and O'Flaherty, 2020; Acharya et al., 2020). Our paper is important in the sense that it offers empirical insights into the actual magnitude of inter-regional epidemiological spillovers in the United States. Our findings suggest that those spillovers are substantial and therefore emphasize the importance of a coordinated policy response.

Third, the variation in state-level restrictions plays a key role in our analysis. This associates us with a number of papers that focus on the effectiveness of various social distancing measures or on the compliance with the official rules.[4] Painter and Qiu (2020) and Simonov et al. (2020) show that compliance with social distancing rules in the U.S. is correlated with party affiliation, and with exposure to certain opinion-forming programs on Fox News. Briscese et al. (2020) show that the compliance can vary over time and that people can become "tired of" restrictions. In our analysis we use a measure of state-imposed restrictions and we also allow for imperfect compliance with them. Our results indicate that stricter social distancing measures introduced by individual states, and better compliance with them, limit the spread of the disease not only within those states, but also in the neighboring states. Conversely, the lack of such restrictions makes it harder for the state's neighbors to contain the virus.

Finally, following Atkeson (2020b), a number of papers have contributed to modelling the spread of the pandemic. The SIR model has become the standard in that literature with different papers suggesting different modifications, depending on the paper's focus.[5] The closest papers to ours are Bisin and Moro (2020) and Acemoglu et al. (2020). The former builds a theoretical framework that formalizes aspects such as local travel and changes in individuals' behavior, but their focus is on the local diffusion around the hot-spot of the outbreak. The latter develops a multi-group version of the SIR model where infection risks differ across population groups (e.g., nursing homes, schools, etc.) and allows for the transmission of infections between population subgroups. Our main contribution is to develop a spatial version of the benchmark SIR model that allows us to quantify the spillover effects of local infections as well as local lock-down policies on the spread of

the virus in other parts of the country.

2 Covid-19 outbreaks and policy responses across the U.S.

We start by documenting some stylized facts about the time and spatial dimensions of the spread of Covid-19 and containment measures in the United States.

2.1 Data

We utilize three data sources in this paper:

1. Daily county-level data on confirmed cases and deaths are from the Covid-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (Dong et al. (2020)).

2. Daily state-level data on business closures and mandated social distancing measures are from the Institute for Health Metrics and Evaluation (IHME). Specifically, we utilize IHME-compiled information on five such metrics including the date on which a state proceeded as follows: forbade mass gatherings, introduced an initial round of business closures, closed schools, closed all non-essential businesses, and adopted a stay-at-home order.[6] For each day in our time series, we sum the number of currently-imposed restrictions within a state to generate a measure of government-imposed behavior restrictions at the state level. This metric is thus a count variable taking values 0 through 5, which we call r-score.

3. County-level data on socioeconomic, demographic, and geographic characteristics from the Bureau of Economic Analysis (BEA). In the analysis below we incorporate county-level information on population density for the year 2018 (BEA and U.S. Census Bureau), the share of

the county-level population over age 60 for the year 2018 (U.S. Census Bureau), and partisan voting share from the 2016 presidential election.[7]

2.2 Covid-19 cases across time and space

In this section we conduct some basic visual analysis to demonstrate the extent of the Covid-19 epidemic in the United States. We do so to provide a descriptive look at the dynamics of the inter-state and inter-county spillover of the virus.

Figure 1 displays the spread of Covid-19 across the whole country over time. The solid line plots the proportion of counties with confirmed cases over time, demonstrating the breadth of the epidemic. The dotted line reveals its depth, displaying how the share of the population living in a county with at least one confirmed case quickly rose from close to zero to close to one during the month of March. It is clear that from March through June, Covid-19 transitioned from a fairly sparse outbreak to a widespread epidemic among counties along both the extensive and intensive margins.

New York was the first state in the U.S. to experience a very significant outbreak, with its epicenter in New York City (NYC). A simple inspection of the dynamics of case numbers in and around NYC is suggestive of the presence of interstate spillovers (within the CT-NJ-NY-PA areas). Figure 2 shows a rapid expansion of per-capita case numbers in New York during the second half of March, followed by all other states surrounding the NYC metropolitan area in late March and the first half of April. In the case of New Jersey, per-capita case counts caught up to—and began to outpace—those of New York in mid-April.

Closer inspection at the county level in other metropolitan areas reveals similar patterns. Figure

Figure 1: Affected Counties and Population Over Time

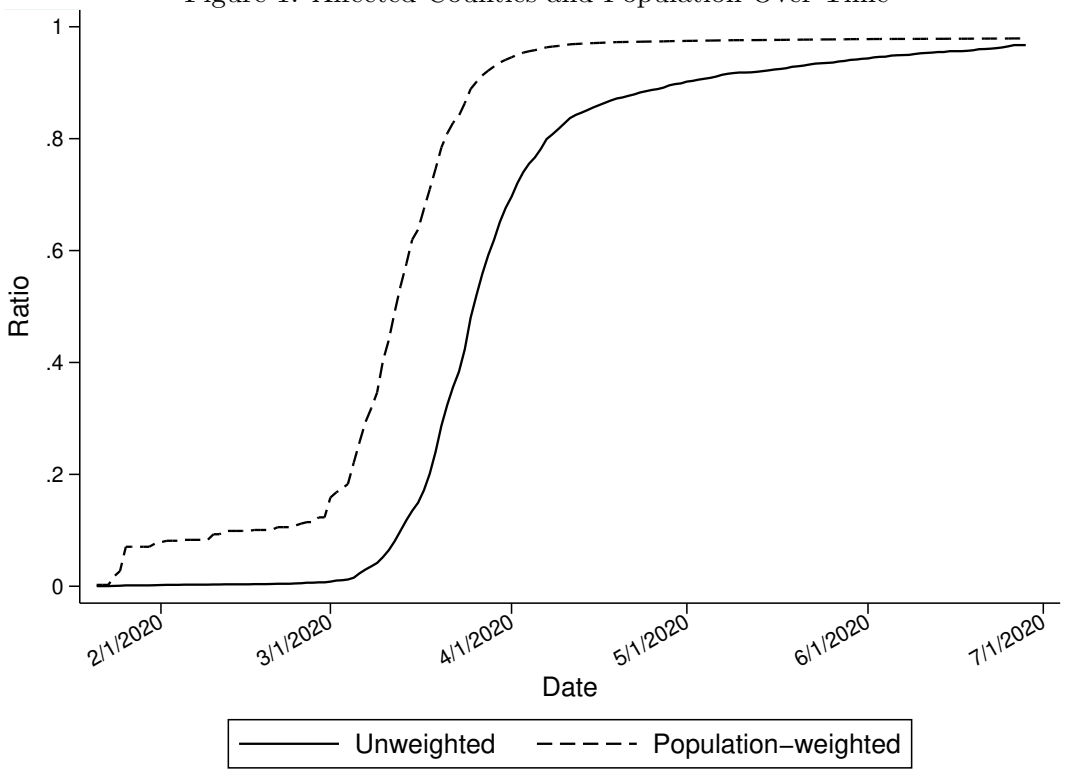


Figure 2: Cases in New York and Surrounding States

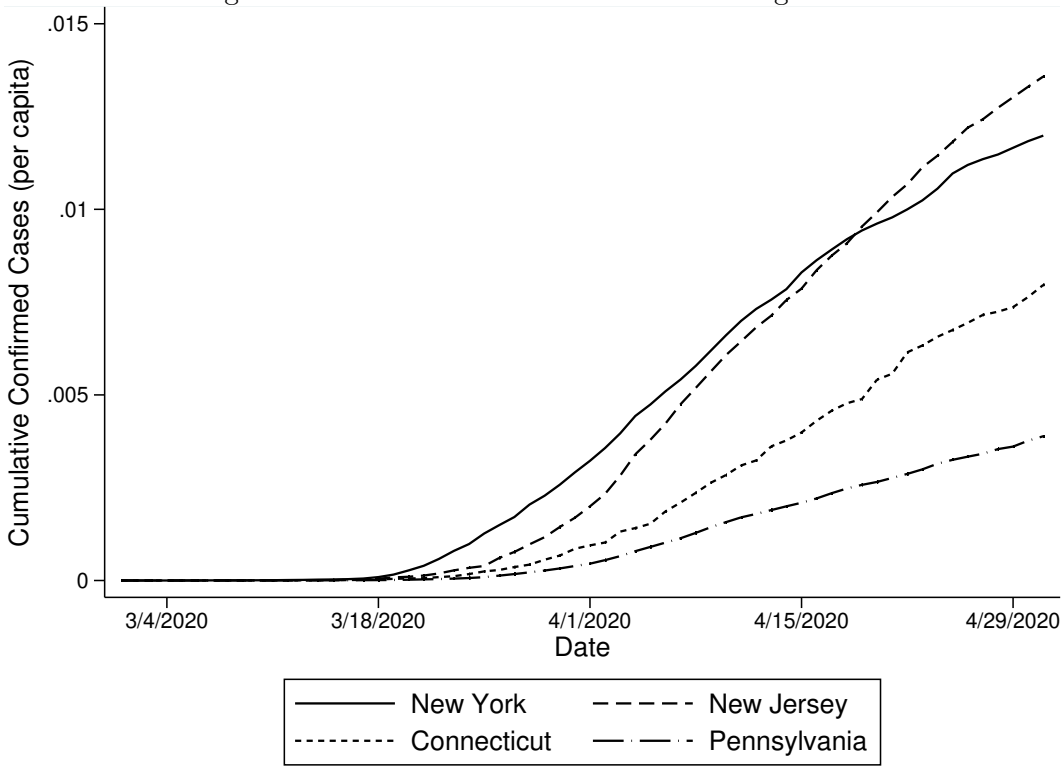
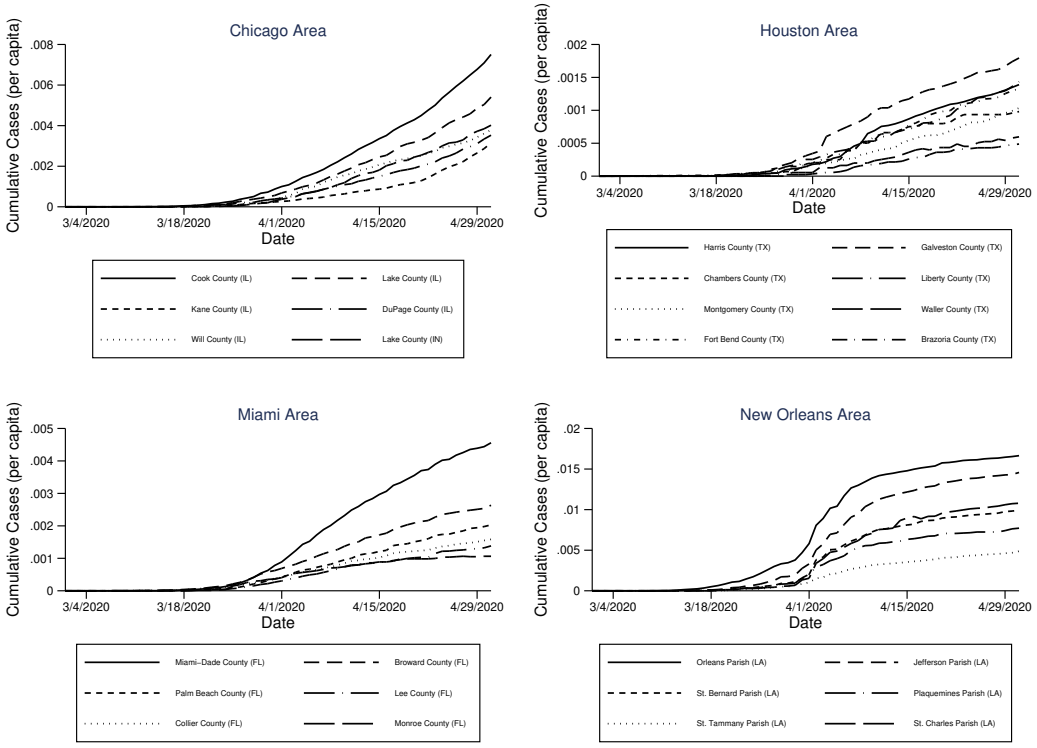
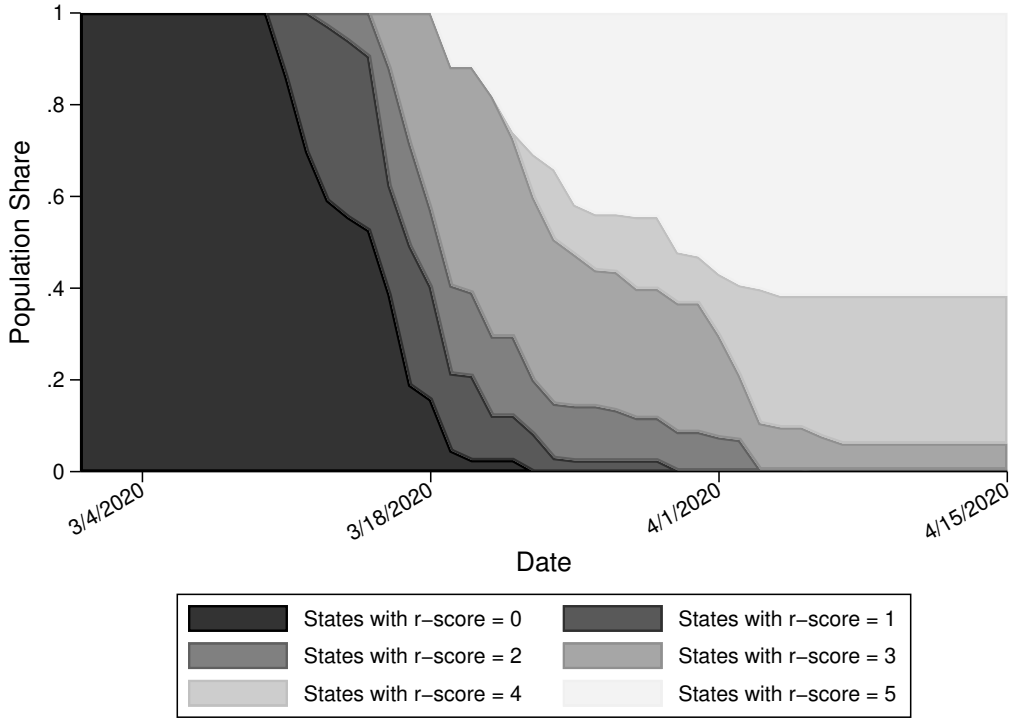


Figure 3: County-level Cases in Major Metropolitan Areas



3 plots confirmed county-level cases per-capita over time for four particular metropolitan areas: Chicago, IL (top left); Houston, TX (top right); Miami, FL (bottom left); New Orleans, LA (bottom right). In each case, the county containing the urban center appears to trigger the area outbreak (respectively: Cook County, IL; Harris County, TX;[8] Miami-Dade County, FL; Orleans Parish, LA). In the case of Chicago, which lies on the Illinois-Indiana border, the spillover appears to extend to Lake County, IN. Together, these figures provide preliminary evidence of transmission dynamics in which Covid-19 cases emanate from major urban centers into the surrounding areas, and across state lines.

Figure 4: Containment Measures Over Time

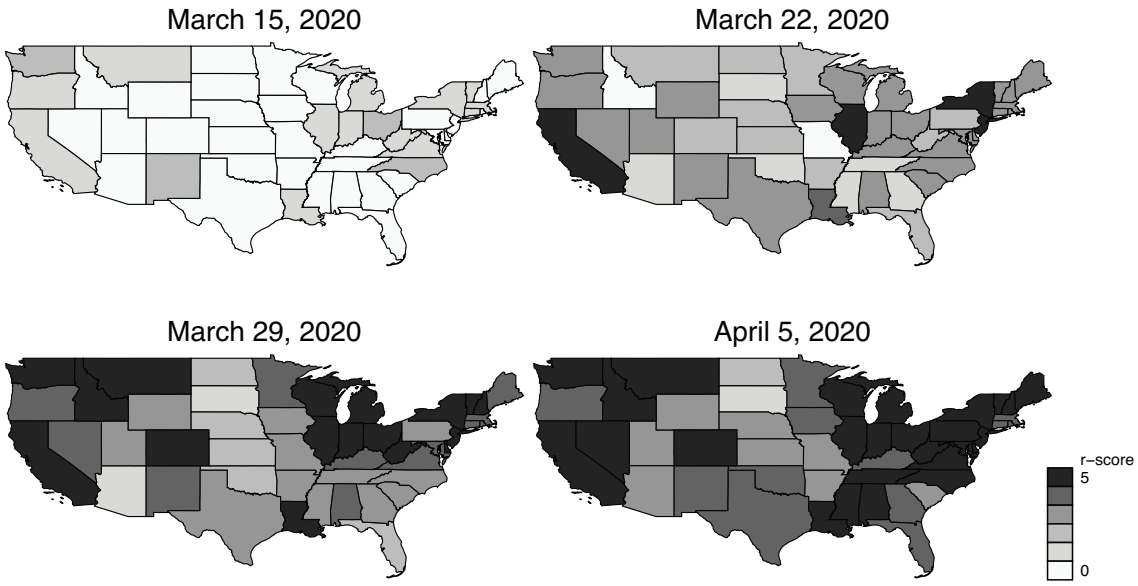


2.3 Containment measures across time and space

There is also clear evidence that during the initial shutdown of March and April, government-led containment measures varied geographically and over time. Figure 4 displays the transition from no official containment measures (early March) to universal adoption (of at least some measures) by all states by early April. The most action occurred between mid-March and early April. 60% of the U.S. population lived in a state with no containment measures on March 15; however, by April 1 nearly 60% of the population lived in a state that had adopted all five factors described by our r-score variable.

In the sections below, our spatial-econometric analysis and our spatial model and calibration

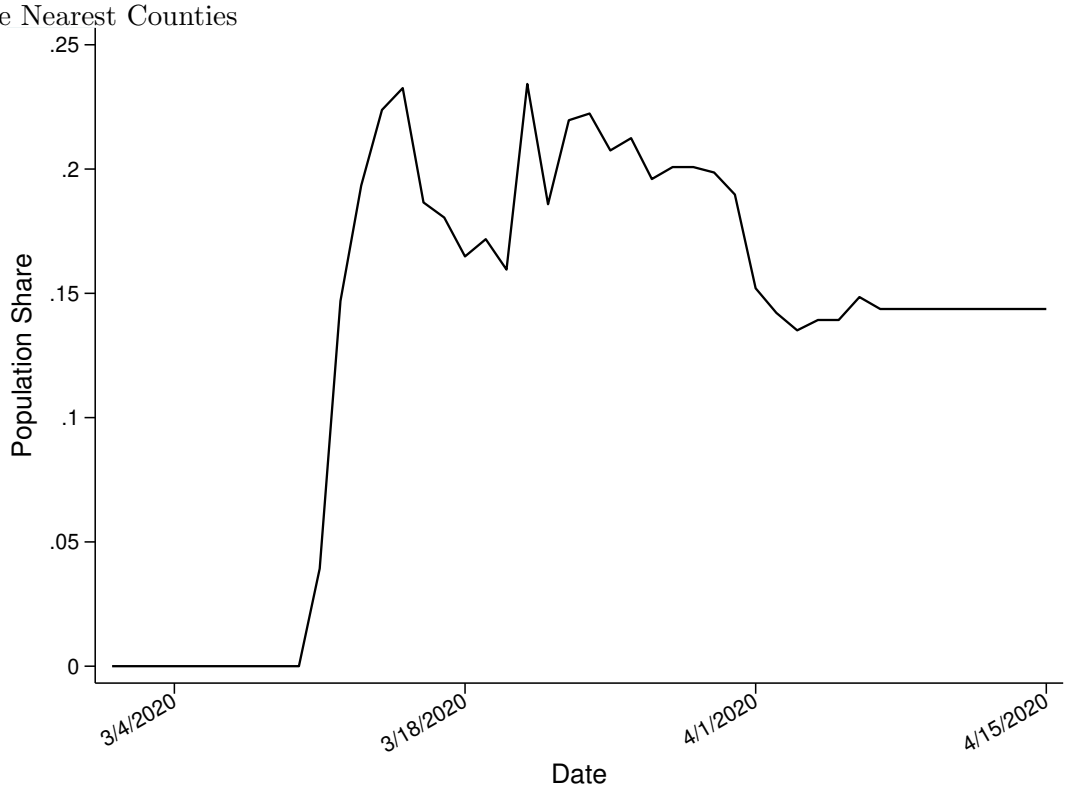
Figure 5: R-scores by State and Time



exercise seek to understand further the geographical spillovers across county and state lines. It is important that there was sufficient geographic variation in the early weeks of the pandemic. Figure 5 shows r-scores in each state from four snapshots spanning mid-March through early April. While it is clear that there are regional trends, most importantly we see that many states have neighbors with different r-scores and those relative differences change over time.

To drill down on the variation in containment measures even further, for spillover effects to be meaningful and identifiable, there must be substantial variation in government-led Covid-19 responses, specifically near state borders. Figure 6 displays the share of the U.S. population living in a county with a different level of containment measures (i.e., r-score value) than at least one of its five nearest counties, over time. During the containment action period of mid-March through

Figure 6: Share of U.S. Population Living in a County with Different Containment Measures than its Five Nearest Counties



early April, this share quickly rose to about 23% of the population, fluctuated between 15% and 23% for the next two weeks, and then settled at 15% for the remainder of April.

3 Estimating the Spatial Diffusion of Cases

In this section we estimate the spatial characteristics of cases. To do so, we estimate the spillovers of cases from county to county in the United States using oft-used spatial econometric models. In this section, the variable new cases is measured as the seven-day average of the growth rate of new cases.

3.1 Estimation of Spatial Correlation

We estimate the following standard spatial models: the Spatial Durbin Model (SDM), the SDM model with spatially correlated errors (SDM Error), the SAR Model, and the SLX model. We start with SDM which, in a panel setting, can be written as:

$$Y_t = \rho W Y_t + \beta X_t + \theta W X_t + \delta + \epsilon_t,$$

where $Y_t = Y_{1,t}, ..., Y_{N,t}$ is a $NT \times 1$ vector of the dependent variable; WY_t is the spatial lag term; X_t is an $NT \times r$ matrix of r exogenous variables; $\delta = \delta_1, ..., \delta_N$ is the vector of region-fixed effects; and WX_t is the spatial lag term for the exogenous variables, which here we refer to as the "exogenous spatial interaction" term (to distinguish this from the spatial lag variable).[9] The spatial term WY_t captures the direct effect of the spatial correlation in determining the dependent variable (capturing what otherwise might be an omitted variable). The SDM error model amends the typical SDM to include a spatially correlated error. The SAR model imposes a restriction that $\theta = 0$, and the SLX model imposes a restriction that $\rho = 0$. In all cases, we employ a contiguity matrix (row-normalized) since this version of the spatial-weighting matrix is the most common in the spatial literature though we considered other versions for robustness (such as an inverse-distance-based matrix), but do not report those results here for brevity.

For the models above, Y_t is the county level cases, and X_t is the r-score. We include a county-level fixed effect to control for county-level features such as population density, relative industry, employment shares, and so on (variables which, given the data sources and time period covered, are fixed over the time period). We report the baseline spatial results for each model, along with the direct and indirect spatial effects. We also consider additional specifications, including with

the r-score lagged 14 days, and with the lag of the dependent variable included as a regressor.

Table 1: Baseline results for Spatial Models of New Cases (measured as a growth rate)

	SAR	SAR	SLX	SLX	SDM	SDM	SDM Error	SDM Error
r-score	-2.299***		0.816**		-0.33		-0.283	
	(0.067)		(0.32)		(0.27)		(0.243)	
r-score 14 day lag		-2.189***		0.911***		-0.174		-0.0419
		(0.0527)		(0.249)		(0.212)		(0.189)
W × cases	0.584***	0.580***			0.584***	0.578***	0.865***	0.862***
	(0.00188)	(0.00189)			(0.00188)	(0.00190)	(0.00131)	(0.00134)
W × r-score			-6.766***		-2.150***		-0.545**	
			(0.34)		(0.29)		(0.25)	
W × (r-score 14 day lag)				-6.577***		-2.200***		-0.757***
				(0.26)		(0.22)		(0.20)
W × error							-0.756***	-0.753***
							(0.00)	(0.00)
N	279549	279549	279549	279549	279549	279549	279549	279549

Notes: Spatial models estimated from April 1 through June 28, controlling for county-level fixed effects. Dynamic refers to the lag of the dependent variable included in the model. 14 day lag of r-score is the value from 14 days prior. The Wald test for spatial correlation is statistically significant in all models. Standard errors are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 2: Dynamic Version of Spatial Models

	SAR	SAR	SLX	SLX	SDM	SDM	SDM Error	SDM Error
cases (t-1)	0.864***	0.864***	0.893***	0.892***	0.864***	0.864***	0.868***	0.868***
	(0.000844)	(0.000844)	(0.000717)	(0.000720)	(0.000844)	(0.000844)	(0.000863)	(0.000863)
r-score	-0.163***		0.364***		0.224*		0.217*	
	(0.0310)		(0.125)		(0.124)		(0.123)	
r-score 14 day lag		-0.198***		0.250**		0.125		0.119
		(0.0242)		(0.0972)		(0.0965)		(0.0959)
W × cases	0.0789***	0.0781***			0.0787***	0.0778***	0.0671***	0.0660***
	(0.00124)	(0.00124)			(0.00124)	(0.00124)	(0.00141)	(0.00142)
W × r-score			-0.854***		-0.423***		-0.457***	
			(0.132)		(0.131)		(0.131)	
W × (r-score 14 day lag)				-0.758***		-0.353***		-0.388***
				(0.103)		(0.102)		(0.102)
W × error							0.0651***	0.0658***
							(0.00347)	(0.00348)
N	279549	279549	279549	279549	279549	279549	279549	279549

Notes: Spatial models estimated from April 1 through June 28, controlling for county-level fixed effects. Dynamic refers to the lag of the dependent variable included in the model. 14 day lag of r-score is the value from 14 days prior. The Wald test for spatial correlation is statistically significant in all models. Standard errors are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Tables 1 and 2 report the results for the four models. Table 2 displays results with a lag of the dependent variable included in each model; we refer to this as the dynamic version.[10] We report baseline results shown in Tables 1 and 2 for purposes of comparison with "traditional" approaches to spatial estimation, of which these models represent. Across the model versions, the most robust result is that the coefficient on the spatial lag (new cases interacted with the spatial weighting matrix) is statistically significant. The effect of the r-score depends on the model. Only in the SAR model is the r-score negative and statistically significant. The interaction of the r-score with the spatial weighting matrix is statistically significant and negative in the SLX, SDM, and SDM error models.

In addition to the models shown in the tables referenced above, we included time-fixed effects in each model (in addition to the county-level effects). Chung and Hewing (2015) note that omitted common shocks may bias estimates of spatial correlation. With time-fixed effects included in each model, the spatial lag is still statistically significant, suggesting the spatial correlation of new cases is robust to including both county and time-fixed effects. However, neither the results for the r-score, nor the results for the r-score interacted with the spatial weighting matrix, are robust to the inclusion of time fixed effects.[11]

Overall, the spatial econometric models provide a general picture of the spatial correlation of the virus. The clearest evidence of the spatial spillover is via the statistical significance of the coefficient on the spatial lag in each model. To further understand the spatial characteristics of the growth rate of new cases across counties, we consider a SIR model.

4 Spatial SIR and Counterfactual Experiments

The previous sections provide evidence of a substantial degree of inter-state spillovers of Covid-19 across regions. In this section, motivated by that evidence, we construct and calibrate a structural SIR model, similar to the one described in Atkeson (2020b), Eichenbaum et al. (2020), Glover et al. (2020), or Fernández-Villaverde and Jones (2020), but with a few modifications. Most importantly, the model allows for infections across state boundaries. We then use the model to evaluate the extent to which the presence of such spillovers contributed to the spread of the infections in the U.S. We also evaluate how lock-down policies implemented in one state impact the rest of the country.

To the extent possible, we try to account for important real-life features, such as the impact of changes in social distancing measures and the state-specific effectiveness of those. Additionally, as pointed out by Fernández-Villaverde and Jones (2020), the identification of parameters in the compartmental models such as the SIR model can be challenging (this is mostly discussed by Atkeson (2020a) in the context of estimating the fatality rate, which is not the central point of our analysis). We partially address this last issue by gradually increasing the complexity of the model and exploring how that increased complexity affects different parts of the model fit. We believe our analysis can still provide useful insights into both the nature and the potential magnitude of inter-state spillovers during the early stages of the outbreak.

While the previous section estimated a spatial model of a daily panel of 3000+ counties, in this section we will consider a state-level version of the SIR model. The reason is that with 3000+ cross-sectional units, the calibration of the highly non-linear SIR model would become computationally infeasible.

4.1 The model

The model is an extension of the SIR model that allows us to account for the spatial diffusion of infections. We specify the model in discrete, rather than continuous time. In each period t, the initial population of region n is divided into four disjoint sets: Susceptible (S), Infected (I), Recovered (R), and Dead (D):

$$Pop_{n,0} = S_{n,t} + I_{n,t} + R_{n,t} + D_{n,t}$$

and population at time t is: $Pop_{n,t} = S_{n,t} + I_{n,t} + R_{n,t}$. The new infections in state n result from interactions between susceptible people S_n in that state, with infected people in potentially all other states $I_{n'}$, where $n' = 1, ..., N$. The new infections in state n at time t are given by:

$$I_{n,t}^{new} = \frac{S_{n,t}}{Pop_{n,t}} \cdot \sum_{n'} \rho(n', n) \cdot \sqrt{\beta_n \beta_{n'}} \cdot \sqrt{\kappa_{n,t} \kappa_{n',t}} \cdot I_{n',t}$$

In the expression above, the whole term $\rho(n', n) \cdot \sqrt{\beta_n \beta_{n'}} \cdot \sqrt{\kappa_{n,t} \kappa_{n',t}}$ describes the average number of close contacts that a person from state n has with a person from state n' in day t. The close contact is defined as one that would result in a transmission of a virus from an infected person to a healthy person. The new infections in state n then occur when an infected person from state n' — $I_{n',t}$ — comes in a close contact with a susceptible person from state n. The probability that a person we come in a close contact with is susceptible is $\frac{S_{n,t}}{Pop_{n,t}}$.

The parameter β_n measures the average number of distinct inter-personal contacts that any person in state n has on a regular day. We allow this parameter to vary across states, given the substantial heterogeneity in the fraction of people living in densely populated areas. We expect, of course, that a typical person in New York will have more distinct inter-personal contacts than

a person living in Montana. We assume β_n is constant over time. It is certainly possible that the typical number of inter-personal contacts will vary over time in each state, and it is quite likely that this variation will differ by state (for example, the value of β would likely plummet during the Spring Break in college towns but skyrocket in the nightclubs or bars in Florida). Given how specific this time variation would be to individual states, we have decided to assume it away, and only allow the model to have a cross-sectional variation in β, which yields 49 parameters to be calibrated. We also consider a simpler specification, with three rather than 49 parameters, where β_n is a polynomial function of the population density in region n: $\beta_n = b_0 + b_1 \log(density) + b_1 \log(density)^2$.

Next, $\kappa_{n,t}$ measures the degree to which the personal interactions are reduced by the implemented lock-down policies. The actual reduction in the personal interactions results from a combination of two factors: the official lock-down policies and their effectiveness in the particular region. That effectiveness (from the perspective of the model) can capture at least two important factors. The first factor is related to individuals' compliance and the region's enforcement of social distancing measures. The second factor is related to the fact that each social distancing measure (as recorded in our data) comes with exceptions. Those exceptions may be different in different states, or the same exception can have a different coverage in different states. In general, we should not expect the same restriction that we code as a particular value of the r-score variable to have an identical impact in each state. While we cannot speak to the reasons behind that heterogeneity, we can incorporate it in a straightforward fashion into our model. In order to do that we model $\kappa_{n,t}$ as follows:

$$\kappa_{n,t} = (1 - \xi_n) + \xi_n \cdot \sum_{i=0}^{5} \kappa^i \cdot 1_{\{i\}} (\text{r-score}_{n,t})$$

where i is the value of the r-score variable (0 through 5), κ^i is the benchmark effect of restriction i in the region where restrictions are most effective, ξ_n is the relative effectiveness of restrictions in region n, and $1_{\{i\}}(\cdot)$ is a characteristic function of a singleton set with element i (essentially, $1_{\{i\}}(\text{r-score}_{n,t})$ equals 1 if $\text{r-score}_{n,t} = i$ and 0 otherwise). When $\xi_n = 1$, the effectiveness of each restriction in state n is as high as it can be, when $\xi_n = 0$ the restrictions are completely ineffective. We normalize $\xi_n = 1$ in one of the regions (determined *endogenously* - see Appendix B.1 for details) and calibrate the 48 remaining values. We also normalize $\kappa^0 = 1$ ($i = 0$ corresponds to no restrictions). We also impose a restriction that $\kappa^{i+1} \leq \kappa^i$, so a tighter restriction would never lead to *more* contacts between people. Overall, this adds $48 + 5 = 53$ additional parameters to the calibration.

Finally, $\rho(n', n)$ denotes the spillover parameter from state n' to n. We restrict the possible values for $\rho(n', n)$ so that the matrix is consistent with the row-normalized weighing matrix used in Section 3. First, we normalize $\rho(n, n) = 1$. Next, we set $\rho(n', n) = 0$ when two states n' and n are not adjacent and we require it to be positive (even if arbitrarily small) when they are. In that case, we set $\rho(n', n) = \rho \cdot \frac{1}{\sum_m 1_{\{x \in \mathbb{R}: x > 0\}}(\rho(n', m))}$, where $\rho > 0$ will be the parameter to be calibrated. In words, the spillover from state n' to state n is divided by the total number states that the state n' is adjacent to. We do that in order to ensure that if Virginia and Maryland were one state, the total spillover from DC would be the same as it is when they are two separate states.[12]

The full dynamics of the model are described by the following equations:

$$S_{n,t+1} = S_{n,t} - I^{new}_{n,t} \tag{4.1}$$

$$I_{n,t+1} = I_{n,t} - \pi_R \cdot I_{n,t} - \pi_D I_{n,t} + I^{new}_{n,t} \tag{4.2}$$

$$R_{n,t+1} = R_{n,t} + \pi_R \cdot I_{n,t} \tag{4.3}$$

$$D_{n,t+1} = D_{n,t} + \pi_D \cdot I_{n,t} \tag{4.4}$$

$$Pop_{n,t+1} = Pop_{n,t} - \pi_D \cdot I_{n,t} \tag{4.5}$$

where π_R is the daily recovery rate and π_D is the daily death rate. We set $\pi_R = 0.03267$ and $\pi_D = 0.00067$, so that the model implies a 2% mortality and a 30-day duration of an average infection.

4.2 Calibration, model fit and parameter values

We calibrate the model by minimizing the sum of squared errors between the data and the model-generated series of both the cumulative and the new confirmed cases per-capita in each region and in the entire country. Our vector of parameters has 103 elements:

$$\theta := [\rho, \beta_1, ..., \beta_{49}, \kappa_1, ..., \kappa_5, \xi_1, ..., \xi_{48}]$$

In our calibration we assume that the confirmed cases per-capita in each state lag the infections by 14 days and we start our analysis on February 1, 2020 under the assumption that the cumulative infections on that day corresponded to confirmed cases on February 14, 2020. Our two main outcome variables are then defined as:

$$y_{n,t} := \frac{I_{n,t-7}}{Pop_{n,0}} \quad \text{and} \quad Y_t := \frac{\sum_n I_{n,t-7}}{\sum_n Pop_{n,0}}, \qquad t = 15, 16, ...$$

The sum of squared errors between the model and the data is then calculated as:

$$SSE(\theta) = \sum_n \left(\sum_{t=15}^{T} \left(y_{n,t}^m(\theta) - y_{n,t}^d \right)^2 + \sum_{t=16}^{T} \left(\Delta y_{n,t}^m(\theta) - \Delta y_{n,t}^d \right)^2 \right) + \sum_{t=15}^{T} \left(Y_t^m(\theta) - Y_t^d \right)^2 + \sum_{t=16}^{T} \left(\Delta Y_t^m(\theta) - \Delta Y_t^d \right)^2$$

The vector of calibrated parameters $\hat{\theta}$ is then given as

$$\hat{\theta} := \arg\min_{\theta} SSE(\theta)$$

The results of the calibration are reported in Table 3, which displays the overall fit of the model as well as the values of selected parameters, except for the individual regions' values of β_n and ξ_n. The latter are reported in the appendix. In the appendix we also present summary of the results from calibrations that use different time windows: starting two and four weeks later (on February 14th and March 1st), or ending two weeks earlier (on June 14th). The calibrated values are not overly sensitive to those modifications.

The progression of the model fit reported in the bottom half of Table 3 helps explain how the parameters are identified. We start our calibration by assuming that (1) $\beta_n = b_0 + b_1 \cdot \text{density}_n + b_1 \cdot \text{density}_n^2$, (2) $\kappa^i = 1$ for all i, (3) $\xi_n = 1$ for all n (r-score changes make no difference in any state). We then proceed by relaxing one restriction at a time.

First, we relax the assumption that $\kappa^i = 1$ for all i (but we still assume that $\xi_n = 1$ for all n and that β_n is a quadratic function of the population density). There is a dramatic improvement in the model fit along all dimensions, most notably in country-wide levels (R^2 doubles) and first differences (R^2 almost triples). This is because the initial growth of infections (country-wide and across states) is being slowed down by the imposition of restrictions. A model without some

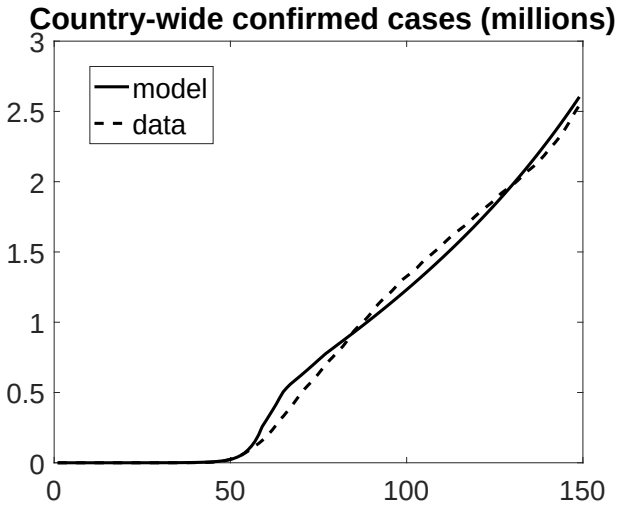
variation in the frequency of close contacts over time cannot capture that change in the dynamics of new infections.

Second, we relax the assumption that β_n is a quadratic function of the population density in state n and instead model it as a state fixed effect. Not surprisingly, the model fit improves dramatically along one important dimension: the variation in levels of infections across states. The model can now account for 96% of that variation (95% of the variation within and 99% between states). There is also a huge improvement in the model's ability to account for the between states variation in the growth rate of new infections (the R^2 more than triples, from 0.29 to 0.93).

Finally, we let ξ_n differ by state (normalizing its highest value to 1). The main improvement in the model's fit can be seen in the variation of new infections over time – both country-wide and across states. Intuitively, the model picks up the difference between states that had the same change in their r-score in the data, but experienced a different reduction in the growth rate of infections. It is worth pointing out that the calibrated values of κ's are much smaller when we allow ξ_n to vary by state, than when we assume identical, perfect effectiveness of each restriction ($\xi_n = 1$ for all n). This is not surprising at all; higher values κ's in columns two and three of Table 3 reflect the fact that in the average state the effectiveness of that restriction measure is not perfect.

While it may not seem so, our model has relatively few free parameters (even with 49 state-specific values of β and 48 state-specific values of ξ), because we have over 7,000 observations (49 regions over 150+ days). Despite that, the model does a remarkably good job in replicating the data. Figure 7 plots the total number of confirmed cases observed in the data and generated by the model for the whole country. The time paths of confirmed cases per-capita for individual states

Figure 7: Model Fit - all confirmed cases



and for DC are reported in the appendix. The R^2 between the cumulative infections for the whole country in model and in the data is 0.99. For individual states, the model accounts for 98% of the overall variation in cumulative infections, for 97% of the variation within states, and 99% of the variation between states. Naturally, the model does a poorer job in accounting for the dynamics of the new infections. For the whole country, it accounts for the 73% of the variation in the data. For individual states, it accounts for 40% of the total variation, 30% of the variation within states, but for the 97% of the variation between states.

4.3 Counter-factual simulations

Given the overall good fit of the structural model, we proceed with using the model to perform counterfactual simulations. In all counterfactual simulations, we use our benchmark parametrization with state-specific values of β_n and ξ_n. Naturally, any counterfactual simulation of a calibrated or estimated model that does not explicitly model people's behavior has to address the Lucas' critique

Table 3: Model Fit and Parameter Values

Model specification	$\beta_n = f(density)$; $\kappa_n = \xi_n = 1$	$\beta_n = f(density)$; $\xi_n = 1$	β_n as fixed effects; $\xi_n = 1$	ξ_n and β_n as fixed effects
Parameters				
ρ	0.136	0.144	0.098	0.156
κ^1	1.000	0.519	0.999	0.640
κ^2	1.000	0.519	0.998	0.427
κ^3	1.000	0.518	0.534	0.185
κ^4	1.000	0.322	0.315	0.055
κ^5	1.000	0.297	0.231	0.026
b_0	-4.995	-4.999	n.a.	n.a.
b_1	0.591	0.717	n.a.	n.a.
b_2	-0.028	-0.029	n.a.	n.a.
Model Fit - levels				
country-wide	0.466	0.984	0.991	0.994
states - total	0.163	0.470	0.962	0.978
states - within	0.170	0.545	0.951	0.972
states - between	0.197	0.375	0.990	0.993
Model Fit - first differences				
country-wide	0.203	0.574	0.612	0.731
states - total	0.001	0.131	0.313	0.403
states - within	0.000	0.107	0.220	0.312
states - between	0.176	0.291	0.925	0.970

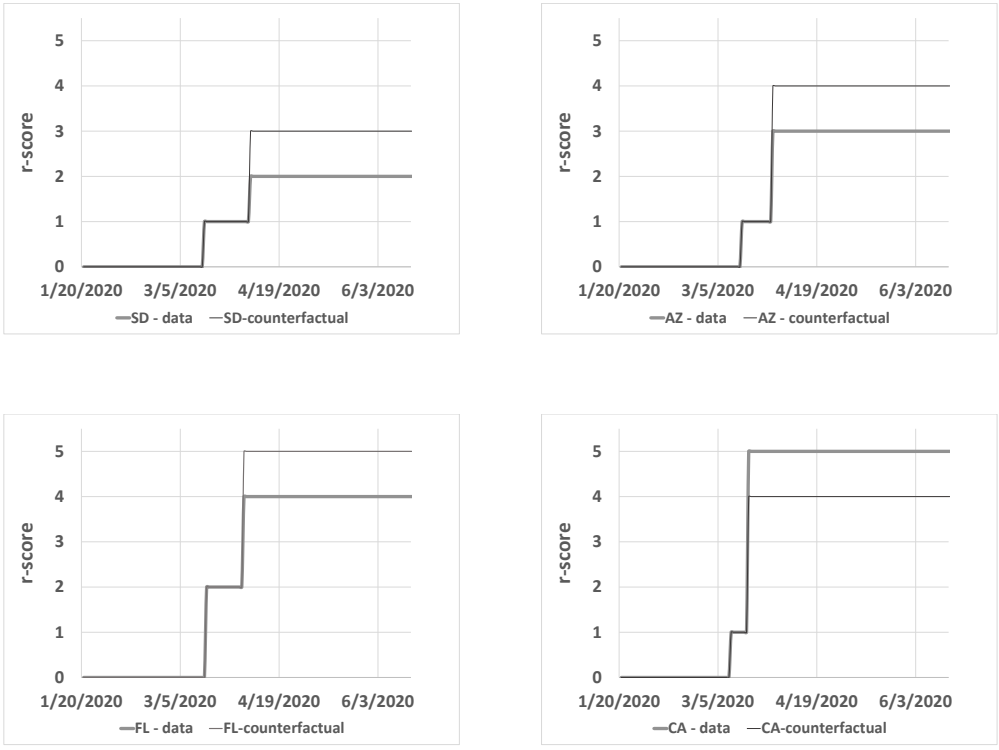
Notes: Each column corresponds to calibration results from a different model specification. We start with the first column assuming that $\beta_n = b_0 + b_1 \text{density}_n + b_2 \text{density}_n^2$, $\kappa^i = 1$ for each value i of the r-score, and $\xi_n = 1$ for each n. We then gradually ease each restriction to allow κ to vary across r-scores, and β_n and ξ_n be modelled as state-specific fixed parameters.

(Lucas, 1976). We want to point out that to some extent we already capture the differences across states in behavioral response to restrictions by calibrating a state-specific parameter ξ_n. Ideally, we would have ξ_n vary by the level of imposed restriction, but we are not able to identify that given our data. Given these considerations, our results should be interpreted as showing the impact of changes in restrictions under the assumption that compliance with them remains the same as it was before (i.e., not necessarily perfect).

Any counterfactual experiment is to some extent *ad hoc*. We present three that, given the focus of our paper, we find the most interesting and informative. First, we investigate what would happen if states changed the maximum level of r-score they implemented by one. Figure 8 plots how that counter-factual experiment is conducted. On the day when the r-score in each state reaches its maximum observed level in the data, we increase that level by one among states that chose not to impose the most aggressive policies corresponding to r-score = 5 (there are 22 of such states). That way, we are able to get a sense of the epidemiological[13] "damage" caused by more lenient policies in such states, both within that group of states, and outside of it. Conversely, among states in which the maximum r-score in the data equals five, we reduce that level by one. That way, we are able to get a sense of the epidemiological "benefit" delivered by the most aggressive policies (within that group of states, and outside of it).

Second, we investigate the role of the spillover parameter, by reducing it by 25%. Third, we combine the two: We repeat the first counterfactual under the assumption that the spillover parameter is 25% lower (of course, we use the simulated series with that lower value of ρ as a benchmark). In all experiments we assume that social distancing measures which are in place on

Figure 8: Counter-factual paths of r-scores



Notes: Each panel plots the actual (data) and counter-factual paths of r-scores in the experiment that we consider: modifying the maximum value of r-score by one. The states are (from top-left, clockwise): South Dakota, Arizona, California, and Florida.

June 28, 2020, remain in place forever.

Table 4 presents the results from counterfactual experiments. We report both percentage as well as absolute changes relative to the benchmark for both total cases (in thousands) and deaths. We first divide all states into four distinct groups, based on the maximum level that the r-score reached in the data. The results are reported in rows 1 through 4. We then combine the states where r-score never reached five into a single group that we will call "lax states". We do that because of the large discrepancy between the number of states within each four groups (there are

twenty seven states where r-score reached 5 and only two states where it reached the maximum value of 2). That way, we split the whole country into two groups, somewhat similar in size. The results for that group are reported in row 5. Next, in row 6 we show the effect of reducing the value of the spillover parameter by 25%. Finally, in row 7, we show the effect of increasing r-score by 1 under the lower spillover parameter. We consider both short-run effects (on June 28, 2020) and long-run effects (December 31, 2021). The first two columns report country-wide effects, columns three and four report own effects (within each group), columns five and six report spillover effects (outside each group).

Raising/lowering the max r-score by 1 Consider first the effects of increasing the maximum level of restrictions by 1 (or lowering it by 1 among states where max(r-score) = 5). They are reported in rows one through five in Table 4. The first thing we notice is that the own effect is an order of magnitude larger than the spillover effect. Second, the spillover effect in the long-run is stronger than in the short-run. Third, even though the relative spillover effect is much smaller than the own effect, it can be quite sizeable. For example, the combined spillover effect among states with max(r-score) < 5 (i.e., the impact on states with max(r-score) = 5) is -2.2% in the short-run and -5.5% in the long-run. While the percentages look small, the total number of confirmed cases among states with max(r-score) = 5 on 6/28/2020 was 1.7 mln, so a 2.2% change means to 39,000 fewer people infected and 790 fewer dead. The -5.5% in the long-run corresponds to reducing the cumulative confirmed cases by more than 4 million, and reducing deaths by over 84 thousands towards the end of December 2021. Even the spillover effect from the two least restrictive states (South Dakota and North Dakota) is not trivial: raising max r-score from 2 to 3 means saving more

Table 4: Counterfactual simulations - changes relative to the benchmark

% changes

Counterfactual (states; action)	country-wide		own effect		spillover	
	short-run	long-run	short-run	long-run	short-run	long-run
rscore = 2; add 1	-0.14	-0.18	-24.25	-19.72	-0.02	-0.04
rscore = 3; add 1	-3.85	-5.71	-62.06	-51.39	-0.60	-1.14
rscore = 4; add 1	-11.39	-19.91	-38.79	-47.92	-1.93	-4.16
rscore = 5; less 1	73.21	36.29	98.99	61.11	16.91	6.89
rscore = 2,3,4; add 1	-15.17	-25.87	-43.48	-49.97	-2.21	-5.52
all; $\rho \downarrow$ by 25%	-40.10	-7.17	-40.10	-7.17	n.a	n.a
rscore = 2,3,4; add 1; ($\rho \downarrow$ by 25%)	-11.81	-26.83	-40.47	-51.95	-1.10	-5.32

Δ cases (000s)

Counterfactual (states; action)	country-wide		own effect		spillover	
	short-run	long-run	short-run	long-run	short-run	long-run
rscore = 2; add 1	-3.5	-248.8	-2.9	-186.9	-0.6	-61.9
rscore = 3; add 1	-100.3	-8,065.7	-85.4	-6,608.5	-14.9	-1,457.1
rscore = 4; add 1	-296.5	-28,104.0	-259.1	-24,346.9	-37.4	-3,757.0
rscore = 5; less 1	1,906.2	51,236.8	1,767.9	46,787.9	138.2	4,448.9
rscore = 2,3,4; add 1	-395.0	-36,516.8	-355.5	-32,289.4	-39.5	-4,227.4
all; $\rho \downarrow$ by 25%	-1,044.0	-10,121.5	-1,044.0	-10,121.5	n.a	n.a
rscore = 2,3,4; add 1; ($\rho \downarrow$ by 25%)	-184.2	-35,160.2	-171.8	-31,401.2	-12.4	-3,759.0

Δ deaths

Counterfactual (states; action)	country-wide		own effect		spillover	
	short-run	long-run	short-run	long-run	short-run	long-run
rscore = 2; add 1	-70	-4,976	-58	-3,739	-13	-1,237
rscore = 3; add 1	-2,006	-161,314	-1,708	-132,171	-298	-29,143
rscore = 4; add 1	-5,930	-562,079	-5,182	-486,939	-748	-75,140
rscore = 5; less 1	38,124	1,024,737	35,359	935,758	2,765	88,978
rscore = 2,3,4; add 1	-7,900	-730,335	-7,110	-645,787	-790	-84,548
all; $\rho \downarrow$ by 25%	-20,881	-202,429	-20,881	-202,429	n.a.	n.a.
rscore = 2,3,4; add 1; ($\rho \downarrow$ by 25%)	-3,685	-703,205	-3,436	-628,024	-249	-75,180

Notes: the first column lists the group of states in which the experiment takes place, followed by the description of the experiment. For example, "rscore = i; add 1" means that we only take states in which max score in the data is $i = 2, 3, 4$, and we increase that max r-score by one (as illustrated in Figure 8), keeping the paths in other states at their benchmark (i.e., taking the empirical path). "Own effect" refers to the effect within the group of states considered, and "spillover" to the effects in all other states. All changes relative to the benchmark paths of infections and deaths. Short-run: all series cut on June 28, 2020; long-run: all series simulated up to December 31, 2021.

than 1,200 lives in *other* parts of the country.

Lower spillover The effect of reducing the value of the spillover parameter is presented in row six in each part of the table. Not surprisingly, if spillovers across states' borders are smaller, the total number of infections and deaths is smaller. Interestingly, the percentage difference is larger in the short-run that in the long-run. In other words, if U.S. states had the ability to restrict travel between them (akin to border closures between countries in the Schengen Zone), the main epidemiological benefit would operate through the flattening of the infection curve.[14]

Changing r-score with lower spillover Finally, we consider the country-wide and the spillover effects of changing local restrictions, when the spillover parameter is smaller. The results are reported in row seven in each part of the table (for ease of exposition, we only report the results for a subset of states). Our results indicate that when the spillover parameter is smaller, the impact of changing the restrictions is diminished substantially in the short-run, but only marginally in the long-run. In the short-run, the percentage spillover effect or raising restrictions in the lax states changes from -2.2% to -1.1%. In the long-run that change is much smaller: from -5.5% to -5.3% (we do no compare absolute changes between rows five and seven, because they correspond to different benchmarks).

5 Conclusion

In this paper we estimate the magnitude of inter-region diffusion of the Covid-19 infections in the early months of the pandemic in the United States. We find evidence that new cases diffuse across county lines, and that the spatial diffusion across counties is affected by the closure policies of

adjacent states. Using a spatial version of the SIR model we find that tightening restrictions in states with the less restrictive policies could have reduced the infections in *other* states by more 2% in the first 3 months, corresponding to a reduction in the number of confirmed cases by 40,000. Also, estimates from traditional spatial models show the spatial correlation is significant between counties, with some evidence that the r-score of counties in adjacent counties have an effect on the growth rate of a county's new cases.

The presence of inter-state spillovers significantly affected the rate of increase in the number of confirmed cases in the early stages of the outbreak. A unique feature of the United States is that its federal government cannot compel individual states to simply close their borders nor mandate state-specific lock-down policies. This only emphasizes the importance of other tools that promote coordination between states' authorities and regular citizens. First, uniform and consistent messaging on precautionary measures such as masks, or encouraging the compliance with social distancing restrictions and discouraging unnecessary inter-state travel, are examples of such tools that would impact individual behavior. Second, the evidence provided in the literature thus far (Piguillem and Shi, 2020; Berger et al., 2020) suggests that there are potentially huge benefits from implementing a country-wide testing system as early in the outbreak as possible—aimed at reducing the delay between test and result—thus revealing virus hot-spots much sooner to potential travelers. Finally, the literature on fiscal federalism may offer some insights into the role the federal government can play when the jurisdictional boundaries do not overlap with the boundaries of regions affected by local policies.[15]

This discussion highlights how our findings reveal the importance of carefully coordinated policy

responses during the early stages of an outbreak, before a virus becomes endemic, while there is still a chance to substantially slow the spread of, or even eradicate, the virus altogether. Given that by the very nature of the problem any policy implemented or not implemented in response to a viral outbreak creates external effects on surrounding regions, we believe this is a very important area for further research to inform policy makers in combating future outbreaks.

References

Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020): "Optimal Targeted Lockdowns in a Multi-Group SIR Model," Working Paper 27102, National Bureau of Economic Research.

Acharya, V. V., Z. Jiang, R. J. Richmond, and E.-L. von Thadden (2020): "Divided We Fall: International Health and Trade Coordination During a Pandemic," Working Paper 28176, National Bureau of Economic Research.

Akovali, U. and K. Yilmaz (2020): "Polarized Politics of Pandemic Response and the Covid-19 Connectedness Across the U.S. States," *Covid Economics*, 57.

Atkeson, A. (2020a): "How Deadly Is COVID-19? Understanding The Difficulties With Estimation Of Its Fatality Rate," Working Paper 26965, National Bureau of Economic Research.

——— (2020b): "What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios," Working Paper 26867, National Bureau of Economic Research.

Atkeson, A., K. Kopecky, and T. Zha (2020): "Estimating and Forecasting Disease Scenar-

ios for COVID-19 with an SIR Model," Working Paper 27335, National Bureau of Economic Research.

Beck, T. and W. Wagner (2020): "National containment policies and international cooperation," *Covid Economics*, 8, 120–134.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2020): "An SEIR Infectious Disease Model with Testing and Conditional Quarantine," Working Paper 26901, National Bureau of Economic Research.

Bisin, A. and A. Moro (2020): "Learning Epidemiology by Doing: The Empirical Implications of a Spatial-SIR Model with Behavioral Responses," Working Paper 27590, National Bureau of Economic Research.

Brady, R. R. (2014): "The spatial diffusion of regional housing prices across US states," *Regional Science and Urban Economics*, 46, 150–166.

Brinkman, J. and K. Mangum (2020): "Travel Behavior and the Coronavirus Outbreak," *Economic Insights*, 5, 23–26.

Briscese, G., N. Lacetera, M. Macis, and M. Tonin (2020): "Compliance with COVID-19 Social-Distancing Measures in Italy: The Role of Expectations and Duration," Working Paper 26916, National Bureau of Economic Research.

Chung, S. and G. Hewing (2015): "Competitive and complementary relationship between regional economies: a study of the Great Lake States," *Spatial Economic Analysis*, 10.

Crucini, M. J. and O. O'Flaherty (2020): "Stay-at-Home Orders in a Fiscal Union," Working Paper 28182, National Bureau of Economic Research.

Cuñat, A. and R. Zymek (2020): "The (structural) gravity of epidemics," *Covid Economics*, 17.

Dave, D. M., A. I. Friedson, D. McNichols, and J. J. Sabia (2020): "The Contagion Externality of a Superspreading Event: The Sturgis Motorcycle Rally and COVID-19," Working Paper 27813, National Bureau of Economic Research.

Debarsy, N., C. Ertur, and J. P. LeSage (2012): "Interpreting dynamic space–time panel data models," *Statistical Methodology*, 9, 158–171.

Desmet, K. and R. Wacziarg (2020): "Understanding Spatial Variation in COVID-19 across the United States," Working Paper 27329, National Bureau of Economic Research.

Dong, E., H. Du, and L. Gardner (2020): "An interactive web-based dashboard to track COVID-19 in real time," *Lancet Infectious Diseases*, 20, 533–534.

Eckardt, M., K. Kappner, and N. Wolf (2020): "Covid-19 across European Regions: the Role of Border Controls," Working Paper DP15178, Center for Economic and Policy Research.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020): "The Macroeconomics of Epidemics," Working Paper 26882, National Bureau of Economic Research.

Elhorst, J. P. (2012): "Dynamic spatial panels: models, methods, and inferences," *Journal of geographical systems*, 14, 5–28.

Ellison, G. (2020): "Implications of Heterogeneous SIR Models for Analyses of COVID-19," Working Paper 27373, National Bureau of Economic Research.

Favero, C. (2020): "Why is Covid-19 mortality in Lombardy so high? Evidence from the simulation of a SEIHCR model," *Covid Economics*, 4.

Fernández-Villaverde, J. and C. I. Jones (2020): "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities," Working Paper 27128, National Bureau of Economic Research.

Gerritse, M. (2020): "Cities and COVID-19 infections: population density, transmission speeds and sheltering responses," *Covid Economics*, 37.

Giannone, E., N. Paixão, and X. Pang (2020): "The Geography of Pandemic Containment," *Covid Economics*, 52.

Ginsburgh, V., G. Magerman, and I. Natali (2020): "COVID-19 and the Role of Economic Conditions in French Regional Departments," Working Papers ECARES 2020-17, ULB – Universite Libre de Bruxelles.

Glover, A., J. Heathcote, D. Krueger, and J.-V. Rios-Rull (2020): "Health versus Wealth: On the Distributional Effects of Controlling a Pandemic," Working Paper DP14606, Center for Economic and Policy Research.

Halleck Vega, S. and J. P. Elhorst (2015): "The SLX model," *Journal of Regional Science*, 55, 339–363.

Holden, R. and D. Thornton (2020): "The Stochastic Reproduction Rate of a Virus," *Covid Economics*, 41.

Hornstein, A. (2020): "Social distancing, quarantine, contact tracing and testing: Implications of an augmented SEIR model," *Covid Economics*, 18.

Jinjarak, Y., R. Ahmed, S. Nair-Desai, W. Xin, and J. Aizenman (2020): "Accounting for Global COVID-19 Diffusion Patterns, January-April 2020," Working Paper 27185, National Bureau of Economic Research.

Kermack, W. O. and A. G. McKendrick (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, 700–721.

Kuchler, T., D. Russel, and J. Stroebel (2020): "The Geographic Spread of COVID-19 Correlates with the Structure of Social Networks as Measured by Facebook," Working Paper 26990, National Bureau of Economic Research.

Lucas, R. J. (1976): "Econometric policy evaluation: A critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

McAdams, D. (2020): "Nash SIR: An economic-epidemiological model of strategic behaviour during a viral epidemic," *Covid Economics*, 16.

Oates, W. E. (1999): "An Essay on Fiscal Federalism," *Journal of Economic Literature*, 37, 1120–1149.

Pace, R. K., R. Barry, J. M. Clapp, and M. Rodriquez (1998): "Spatiotemporal autoregressive models of neighborhood effects," *The Journal of Real Estate Finance and Economics*, 17, 15–33.

Painter, M. and T. Qiu (2020): "Political Beliefs affect Compliance with COVID-19 Social Distancing Orders," *Covid Economics*, 4.

Piguillem, F. and L. Shi (2020): "Optimal COVID-19 quarantine and testing policies," *Covid Economics*, 27.

Pragyan Deb, Davide Furceri, J. D. O. and N. Tawk (2020): "The effect of containment measures on the COVID-19 pandemic," *Covid Economics*, 19.

Renne, J. P., G. Roussellet, and G. Schwenkler (2020): "Preventing COVID-19 Fatalities: State versus Federal Policies," *Covid Economics*, 56.

Rothert, J. (2020): "Optimal federal redistribution during the uncoordinated response to a pandemic," Departmental working papers, United States Naval Academy Department of Economics.

——— (2021): "Strategic inefficiencies and federal redistribution during uncoordinated response to pandemic waves," *European Journal of Political Economy, forthcoming*, 102003.

Simonov, A., S. K. Sacher, J.-P. H. Dubé, and S. Biswas (2020): "The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the Covid-19 Pandemic," Working Paper 27237, National Bureau of Economic Research.

Verwimp, P. (2020): "The Spread of COVID-19 in Belgium: a Municipality-Level Analysis," Working Papers ECARES 2020-25, ULB – Universite Libre de Bruxelles.

Weber, E. (2020): "Which measures flattened the curve in Germany?" *Covid Economics*, 24.

Notes

[1]We discuss possible limitations arising due to the Lucas' critique in Section 4.3.

[2]Of course, we do not suggest that giving the states such ability would be desirable. We are merely evaluating its potential impact on the spread of infections across the whole country.

[3]Those include Eckardt et al. (2020), Dave et al. (2020), Renne et al. (2020), Giannone et al. (2020), Akovali and Yilmaz (2020), or Brinkman and Mangum (2020).

[4]See e.g., Weber (2020), Pragyan Deb and Tawk (2020), Jinjarak et al. (2020).

[5]A non-exhaustive list of examples includes Atkeson et al. (2020), Holden and Thornton (2020), McAdams (2020), Favero (2020), Berger et al. (2020), Hornstein (2020), or Ellison (2020).

[6]While masking guidance grew in prominence over the course of the pandemic, our dataset covers the early stages of the pandemic when such guidance was inconsistent at all levels (national, state, and county). States and localities primarily took their queues from the national authorities with regard to masking requirements, which were insignificant until later in 2020. On the other hand, the most prominent containment measures, listed here, were imposed at the state-level during the initial weeks of the pandemic.

[7]Via Luis Sevillano on GitHub, but originally published in the New York Times and available here.

[8]Here the outbreak appears to stem from both Harris County (Houston) and Galveston County, which is also a fairly densely populated area.

[9]See Halleck Vega and Elhorst (2015) for a detailed discussion on the SLX model and the other spatial models employed in this paper.

[10]There are various versions of spatial models with temporal dynamics. Elhorst (2012) refers to a SAR model augmented with temporal lag of the dependent variable and a temporal lag of the spatial lag as the "time-space dynamic model." Pace et al. (1998) provide an example with their "STAR" model. Brady (2014) provides a brief overview of some of these models. See also Debarsy et al. (2012) for discussion.

[11]A table of results with time-fixed effects are available in the appendix.

[12]Allowing for $\rho(n', n)$ to have distinct value for each pair of states would yield $49 \times 24 = 1,176$ parameters to be calibrated if we assume symmetric spillovers, and double that if we do not.

[13]We cannot say anything about suboptimality of the lenient policies, since we do not have a properly specified social welfare function that would take into account the economic costs.

[14]Eckardt et al. (2020) show that border closures between European regions significantly slowed down the spread of the virus.

[15]See Oates (1999) for the literature review on that topic. Rothert (2020) discusses a few examples of federal fiscal tools that could impact local policies.

A States' summary statistics and calibrated fixed effects

State	β_n	ξ_n	Density	Cases per 1 mln	max(r-score)
AL	0.20	0.54	227.8	7250.8	5
AR	0.12	0.71	169.4	6339.5	3
AZ	0.16	0.62	319.3	10313.0	3
CA	0.21	0.48	1326.5	5484.6	5
CO	0.26	0.60	909.7	5739.1	5
CT	0.17	0.59	908.9	12890.1	4
DC	0.06	0.14	10275.8	14588.8	5
DE	0.13	0.72	742.1	11576.0	5
FL	0.62	0.82	795.9	6616.4	4
GA	0.30	0.61	1021.6	6694.3	4
IA	0.16	0.54	231.8	9034.4	4
ID	0.10	0.59	196.0	3049.8	5
IL	0.27	0.62	1681.7	11150.1	5
IN	0.23	0.61	641.7	6884.6	5
KS	0.28	0.78	456.7	4814.0	3
KY	0.20	0.83	561.7	3464.3	4
LA	0.32	0.71	382.4	12065.5	5
MA	0.28	0.70	1735.3	15702.4	4

Continued on next page

State	β_n	ξ_n	Density	Cases per 1 mln	max(r-score)
MD	0.47	0.72	1698.8	11147.3	5
ME	0.10	0.69	121.8	2380.4	5
MI	0.37	0.73	940.1	6991.2	5
MN	0.27	0.63	966.2	6320.8	4
MO	0.16	0.72	812.6	3450.8	4
MS	0.14	0.46	124.7	8671.9	5
MT	0.05	0.64	28.3	812.4	5
NC	0.30	0.55	621.6	5988.1	5
ND	0.12	0.71	43.6	4598.2	2
NE	0.24	0.68	639.1	9796.5	3
NH	0.18	0.75	301.1	4235.3	5
NJ	0.40	0.75	2978.5	19350.6	5
NM	0.14	0.56	211.6	5635.6	4
NV	0.17	0.72	222.0	5664.1	5
NY	0.43	0.76	11497.0	20330.0	5
OH	0.25	0.58	825.9	4303.8	5
OK	0.19	0.86	476.9	3282.7	4
OR	0.14	0.69	531.6	1992.0	4
PA	0.22	0.69	2050.2	7010.6	5

Continued on next page

State	β_n	ξ_n	Density	Cases per 1 mln	max(r-score)
RI	0.21	0.66	1110.4	14134.9	4
SC	0.25	0.48	292.8	6553.7	4
SD	0.14	0.85	76.0	7697.7	2
TN	0.25	0.62	527.9	5900.3	5
TX	0.23	0.61	1245.4	5306.1	4
UT	0.21	0.62	680.8	6692.3	3
VA	0.45	0.77	1545.9	7560.2	5
VT	0.03	1.00	111.8	1909.6	5
WA	0.24	0.56	499.7	4424.5	5
WI	0.17	0.65	322.2	4791.7	5
WV	0.10	0.81	160.8	1568.3	5
WY	0.04	0.83	12.4	2452.7	3

B SIR model

B.1 Calibration procedure

B.1.1 General description

In this appendix we describe our calibration procedure. In general, we minimize the sum of squared errors between observed (1) confirmed infections (country-wide and state-by-state) and (2) changes in confirmed infections (country-wide and state-by-state), and those simulated in the model.

We follow a step-by-step process to have a better understanding of how each set of parameters is identified (our model has 100+ parameters, so we believe this approach is quite useful). Table 3 in the main text details how, and along which dimension the model fit improves with each step.

B.1.2 Step-by-step process

Let θ be the vector of parameters to be calibrated at a given step, and let $SSE(\theta)$ be the sum of squared errors between model simulated series and the data, as defined in Section 4.2 in the main text.

Step 1 (column 1 in Table 3)

In the first step, we exploit the cross-state correlation between the growth rate of new cases and population density, as well as the spatial correlation of new cases between states.

Assume $\kappa^i = \xi_n = 1$ for all $i = 0, 1, .., 5$ and for all $n = 1, ..., 49$.

Assume $\log \beta_n = b_0 + b_1 \cdot \ln(\text{density}_n) + b_2 \cdot [\ln(\text{density}_n)]^2$.

Define $\theta = [\rho, b_0, b_1, b_2]$.

Set initial guess to be: $\rho = 0.1$; $b_0 = -2$; $b_1 = 0.0001$; $b_2 = -0.0001$

Define $\hat{\theta} = [\hat{\rho}, \hat{b}_0, \hat{b}_1, \hat{b}_2] = \arg\min_\theta SSE(\theta)$

Step 2 (column 2 in Table 3)

In the second step, we exploit the correlation between changes in restrictions and the slow-down in the growth of new cases, relative to what would have been predicted by the model with β_n constant over time.

Normalize $\kappa^0 = 1$ and assume $\xi_n = 1$ for all $n = 1, ..., 49$.

Assume $\kappa^i = \tilde{\kappa}^1 \cdot ... \cdot \tilde{\kappa}^i \quad \forall i$, where $\tilde{\kappa}^i \in [0, 1]$. This ensures that $\forall i$ we have $\kappa^i \geq \kappa^{i+1}$.

Assume $\log \beta_n = b_0 + b_1 \cdot \text{distance}_n + b_2 \cdot \text{distance}_n^2$.

Define $\theta = [\rho, b_0, b_1, b_2, \tilde{\kappa}^1, \tilde{\kappa}^2, ..., \tilde{\kappa}^5]$.

Set initial guess to be: $\rho = \hat{\rho}$; $b_0 = \hat{b}_0$; $b_1 = \hat{b}_1$; $b_2 = \hat{b}_2$ (solutions from the previous step), and set initial guess $\tilde{\kappa}^i = 0.999$ for all i.

Define $\hat{\theta} = \arg\min_\theta SSE(\theta)$

Step 3 (column 3 in Table 3)

In the third step, we relax the assumption that the only factor behind state-specific β_n is population density.

Normalize $\kappa^0 = 1$ and assume $\xi_n = 1$ for all $n = 1, ..., 49$.

Assume $\kappa^i = \tilde{\kappa}^1 \cdot ... \cdot \tilde{\kappa}^i \quad \forall i$, where $\tilde{\kappa}^i \in [0, 1]$. This ensures that $\forall i$ we have $\kappa^i \geq \kappa^{i+1}$.

Define $\theta = [\rho, \tilde{\kappa}^1, \tilde{\kappa}^2, ..., \tilde{\kappa}^5, \beta_1, ..., \beta_{49}]$.

Set initial guesses to be:

- $\rho = \hat{\rho}$ (solution from the previous step)
- $\tilde{\kappa}^i = \hat{\tilde{\kappa}^i}$ (solutions from the previous step)

- $\ln \beta_n = \hat{b}_0 + \hat{b}_1 \ln \text{density}_n + \hat{b}_2 \left[\ln \left(\text{density}_n\right)\right]^2$

Define $\hat{\theta} = \arg\min_\theta SSE(\theta)$

Step 4 (column 2 in Table 3)

In the fourth and final step, we relax the assumption that the effectiveness of restrictions is identical across states.

Normalize $\kappa^0 = 1$.

Define $\tilde{\xi}_1 \equiv 1$ and consider a vector of non-negative values $[\tilde{\xi}_2, \tilde{\xi}_3, ..., \xi_{49}] \geq 0$.

For each $n = 1, 2, ..., 49$ define

$$\xi_n = \frac{\tilde{\xi}_n}{\max\left(1, \tilde{\xi}_2, \tilde{\xi}_3, ..., \xi_{49}\right)}$$

This ensures that $\max(\xi_1, ..., \xi_{49}) = 1$ without assuming a priori which element is set to one.

Assume $\kappa^i = \tilde{\kappa}^1 \cdot ... \cdot \tilde{\kappa}^i \quad \forall i$, where $\tilde{\kappa}^i \in [0, 1]$. This ensures that $\forall i$ we have $\kappa^i \geq \kappa^{i+1}$.

Define $\theta = [\rho, \tilde{\kappa}^1, \tilde{\kappa}^2, ..., \tilde{\kappa}^5, \beta_1, ..., \beta_{49}, \tilde{\xi}_2, ..., \tilde{\xi}_{49}]$.

Set initial guesses to be:

- $\rho = \hat{\rho}$ (solution from the previous step)
- $\tilde{\kappa}^i = \hat{\tilde{\kappa}^i}$ (solutions from the previous step)
- $\ln \beta_n = \hat{b}_0 + \hat{b}_1 \ln \text{density}_n + \hat{b}_2 \left[\ln \left(\text{density}_n\right)\right]^2$
- $\tilde{\xi}_n = 1$ for all n

Define $\hat{\theta} = \arg\min_\theta SSE(\theta)$

B.2 Stability of parameters

Table B.2 below shows the sensitivity of the model parameters to the calibration window.

Table B.2: Sensitivity of model parameters

	Calibration window			
Parameter	Feb 1 - Jun 14	Feb 1 - Jun 28	Feb 15 - Jun 28	Mar 1 - Jun 28
ρ	0.137	0.156	0.155	0.158
κ^1	0.701	0.640	0.819	0.737
κ^2	0.468	0.427	0.519	0.436
κ^3	0.249	0.185	0.448	0.405
κ^4	0.108	0.055	0.224	0.038
κ^5	0.041	0.026	0.026	0.005
$\min \beta_n$	0.031	0.030	0.018	0.023
average β_n	0.194	0.188	0.234	0.189
$\max \beta_n$	0.478	0.624	0.816	0.467
$\min \xi_n$	0.146	0.144	0.332	0.183
average ξ_n	0.691	0.664	0.802	0.664

B.3 SIR model fit

Figures B.1 - B.4 show the model fit for individual 49 contiguous regions (48 states + DC).

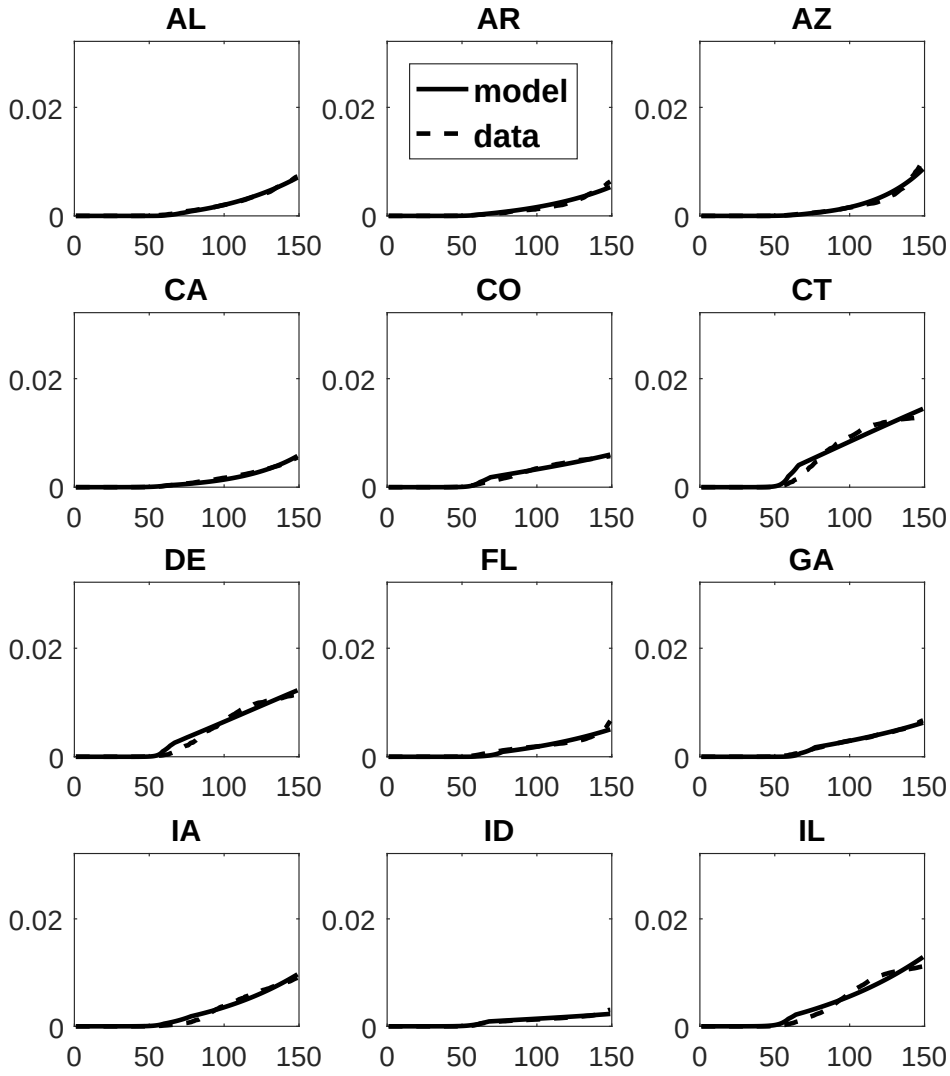


Figure B.1: Model (solid) vs. data (dashed) - individual states; Y-axis: cases per capita; X-axis: days since February 01

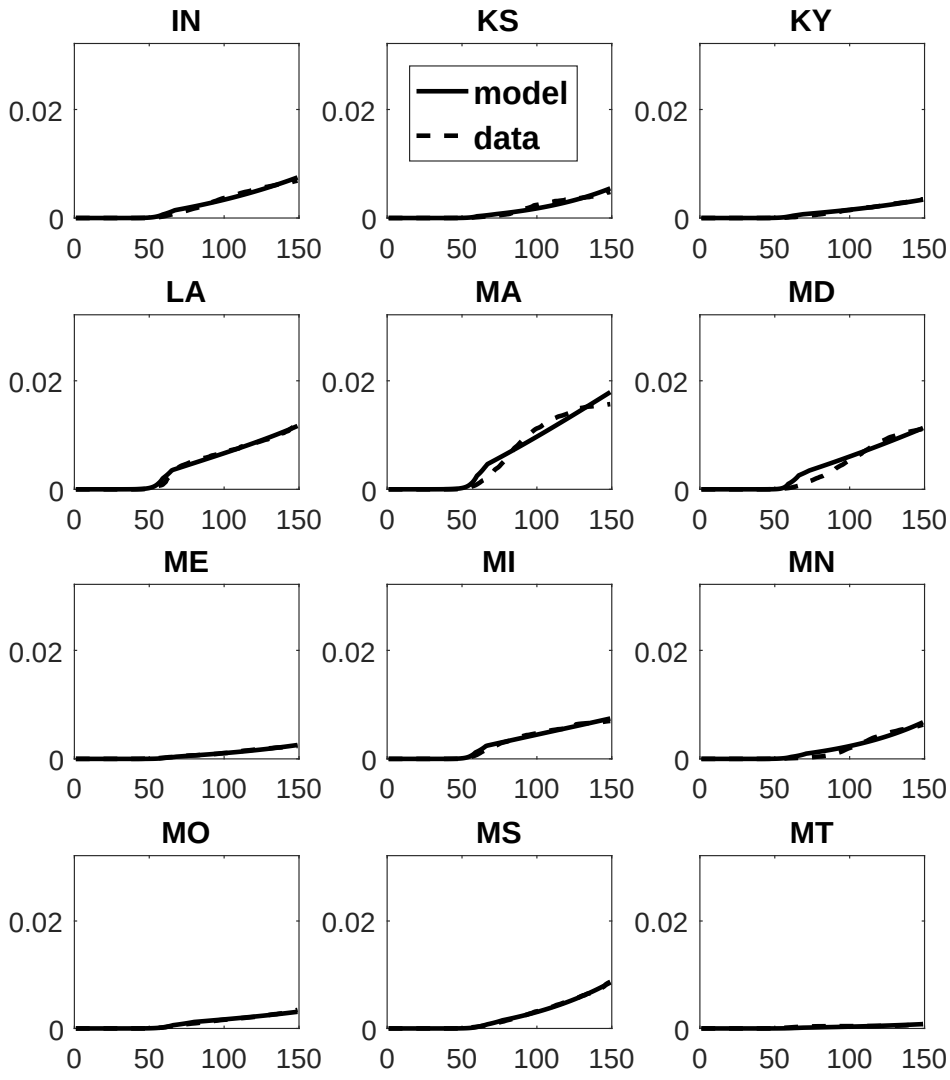


Figure B.2: Model (solid) vs. data (dashed) - individual states; Y-axis: cases per capita; X-axis: days since February 01

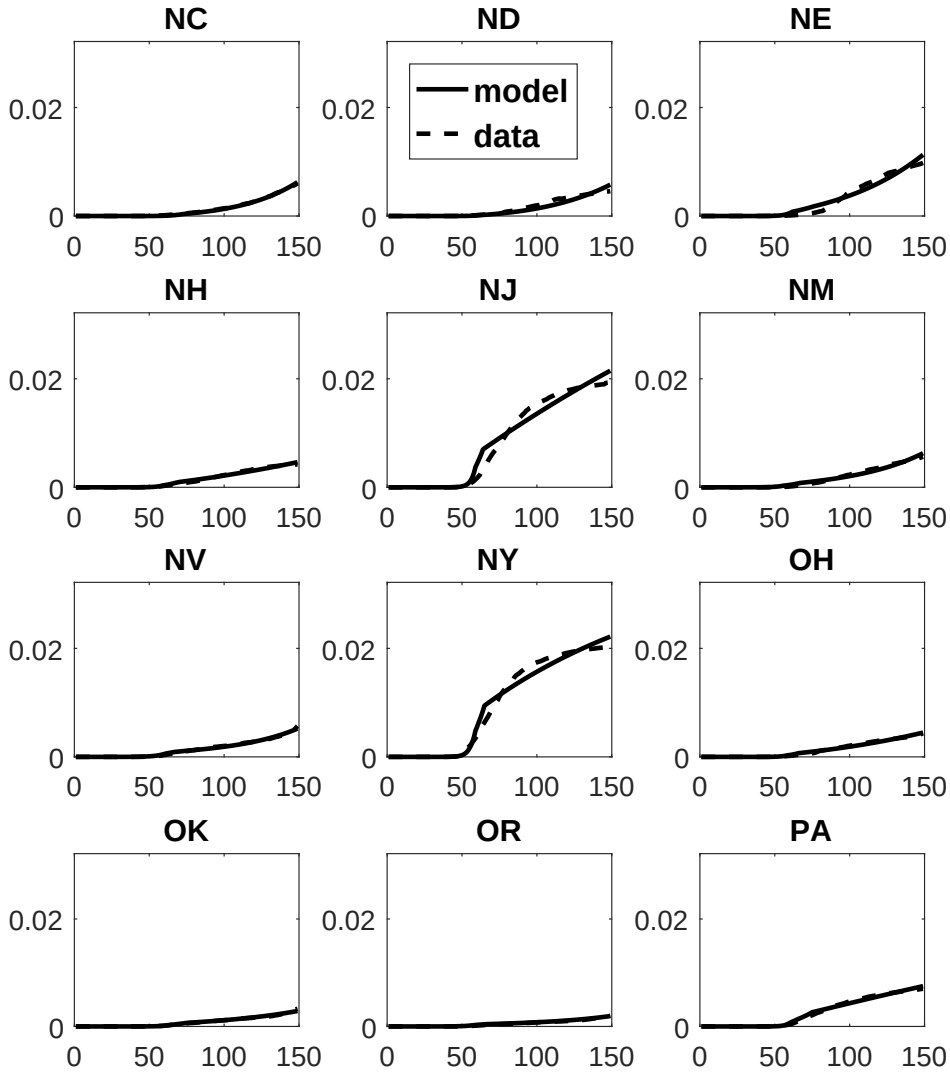


Figure B.3: Model (solid) vs. data (dashed) - individual states; Y-axis: cases per capita; X-axis: days since February 01

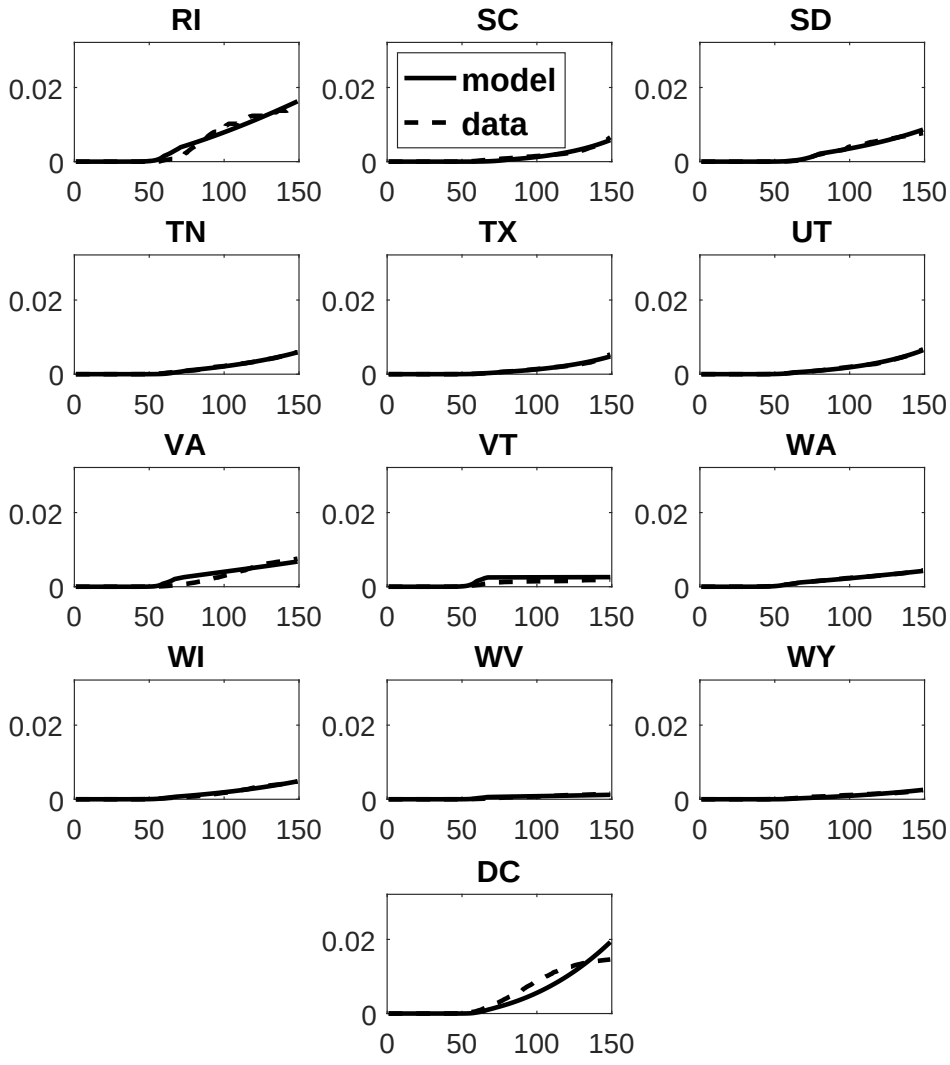


Figure B.4: Model (solid) vs. data (dashed) - individual states; Y-axis: cases per capita; X-axis: days since February 01

B.4 Spatial Models with time-fixed effects

The table below provides a comparison of the results reported in Table 1 with the spatial econometric models estimated with time-fixed effects included in addition to county-fixed effects.

Table B.3: Spatial models with time-fixed effects

	SAR	*SAR*	SLX	*SLX*	SDM	*SDM*	SDM Error	*SDM Error*
r-score	-2.299***	0.037	0.816**	-1.37***	-0.33	-1.262***	-0.283	-0.90***
	(0.067)	(0.278)	(0.32)	(0.285)	(0.27)	(0.269)	(0.243)	(0.228)
W*cases	0.584***	0.403***			0.584***	0.403***	0.865***	-0.5946***
	(0.00188)	(0.0024)			(0.00188)	(0.0024)	(0.00131)	(0.00515)
W*r-score			-6.766***	1.79***	-2.150***	1.45***	-0.545**	0.858***
			(0.34)	(0.307)	(0.29)	(0.289)	(0.25)	(0.320)
W*error							-0.756***	0.741***
							(0.00)	(0.002)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects		Yes		Yes		Yes		Yes
N	276408	276408	276408	276408	276408	276408	276408	276408

Notes: Spatial models estimated from April 1 through June 28, controlling for county-level fixed effects and with county-fixed effects *and* time-fixed effects (see column headers in italics). The Wald test for spatial correlation is statistically significant in all models. Standard errors are in parentheses.
* p<.1, ** p<.05, *** p<.01