

# JUE Insight: Learning Epidemiology by Doing: The Empirical Implications of a Spatial-SIR Model with Behavioral Responses<sup>§</sup>

Alberto Bisin                  Andrea Moro

June 8, 2021

First version: June 10, 2020

## Abstract

We simulate a spatial behavioral model of the diffusion of an infection to understand the role of geographic characteristics: the number and distribution of outbreaks, population size, density, and agents' movements. We show that several invariance properties of the SIR model concerning these variables do not hold when agents interact with neighbors in a (two dimensional) geographical space. Indeed, the spatial model's local interactions generate matching frictions and local herd immunity effects, which play a fundamental role in the infection dynamics. We also show that geographical factors change how behavioral responses affect the epidemic. We derive relevant implications for estimating the effects of the epidemic and policy interventions that use panel data from several geographical units.

---

<sup>§</sup>Bisin: New York University, [wp.nyu.edu/albertobisin/](http://wp.nyu.edu/albertobisin/), [alberto.bisin@nyu.edu](mailto:alberto.bisin@nyu.edu). Moro: Vanderbilt University, [andreamoro.net](http://andreamoro.net), [andrea@andreamoro.net](mailto:andrea@andreamoro.net). We thank Pedro Sant'Anna, Giorgio Topa, Maxim Pinkovskiy, the editor and two anonymous referees for their helpful comments on earlier drafts of this paper, and Gianluca Violante for suggestions about the calibration.

# 1 Introduction

The SARS-CoV-2 epidemic has diffused at very different rates across countries and cities.<sup>1</sup> Plots of case statistics over time by location are common with media and opinion leaders to compare the dynamics of the epidemic across geographical units, often intending to evaluate the effect of different policy interventions. But, how can we compare the United States to Ireland or New York to Miami, given their differences in population size, density, and other geographic and socio-economic characteristics? How do we export parameter estimates about the epidemic obtained from the city of Vo', a small town near Padua, in Italy, or from the Diamond Princess cruise ship, to understand the diffusion of the epidemic in New York city?<sup>2</sup>

This paper proposes a spatial model of epidemic diffusion, the Spatial-SIR model, to study how the dynamics of an epidemic scales in relevant geographical characteristics: the number and distribution of outbreaks, population size, density, and agents' movements.<sup>3</sup> To this end, we adopt an Agent-Based methodology, that is, we simulate populations of agents characterized by micro-level rules of behavior over time and space.<sup>4</sup> Research in epidemiology has extended SIR, allowing for detailed descriptions of the geographic, socio-economic, and demographic characteristics of the population of interest and its environment. In social epidemiology, Agent-Based models generally incorporate detailed, granular assumptions on human behavior and the social and physical environment, with the aim of forecasting with accuracy and precision the dynamics of an epidemic (as, say, meteorological models of weather dynamics).<sup>5</sup> In this paper instead we aim at identifying the stylized effects of geographic characteristics on the dynamics emerging from an abstract spatial-SIR model.

The Spatial-SIR model imposes restrictions on the epidemic dynamics that depend on each location's geographical characteristics, informing comparisons across locations. We show that it is not possible to uncover these restrictions from the workhorse model of epidemic diffusion, the SIR model (Kermack and McKendrick (1927, 1932)). These restrictions are consequential for empirical analysis using longitudinal infection data. Most of the recent wealth of contributions to the study of the SARS-CoV-2 epidemic in economics has restricted its epidemiology component to SIR and does not account for the geographic characteristics

---

<sup>1</sup>See Desmet and Wacziarg (2020) and Fernandez-Villaverde and Jones (2020).

<sup>2</sup>See Lavezzo et al. (2020) and Mizumoto et al. (2020) respectively.

<sup>3</sup>Duranton and Puga (2020) overview the importance of geographic factors in economic modeling.

<sup>4</sup>See, e.g., El-Sayed et al. (2012); Hunter et al. (2018) for surveys and methodological discussions. More generally, for an introduction and survey of Agent-Based models in the social sciences, see, e.g., Billari and Prskawetz (2012); Bruch and Atwell (2015). For spatially explicit Agent-Based models in the social epidemiology tradition of detailed forecasting models, see Dunham (2005); Grefenstette et al. (2013); and Hunter et al. (2017) for a survey.

<sup>5</sup>See, e.g., Eubank et al. (2004), the research at GLEAM project, mobs-lab, and the Imperial college MRC Centre for Global Infectious Disease Analysis available, respectively, at <https://covid19.gleamproject.org>, <https://www.mobs-lab.org/projects.html>, and <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis>.

that we focus on in this paper.<sup>6</sup> Several exceptions introduce spatial dimensions to SIR, but focus on how connections between geographical units affect the epidemic diffusion.<sup>7</sup> In this paper instead, we focus on the comparative dynamics of the epidemic with respect to different geographical characteristics of (closed) units. The spatial dimensions we account for in the present paper introduce local interactions in the contact process between agents.<sup>8</sup>

Section 2 begins by highlighting the relevant invariance properties of SIR with respect to the geographic characteristics we focus on in this paper. Section 3 introduces the Spatial-SIR model. Individuals are placed in a two-dimensional space and travel in this space at a given speed. When infected, they can only infect their neighbors with a probability that we interpret as the strength of the virus. Spatial-SIR determines the diffusion rate of infection depending on epidemiological and geographic factors that are confounded in one parameter of the standard SIR. Section 4 shows how distinguishing these factors is crucial in Spatial-SIR because the local interactions arising in the model generate matching frictions across agents. What we call “local herd immunities” arise from the constrained movement of people in space. Local herd immunities break several invariance relationships which hold in the SIR model (enumerated in Section 2), where susceptible individuals match with infected individuals randomly.

Section 5 presents several simulations of a calibrated Spatial-SIR to study the roles of geographic characteristics. Section 6 incorporates behavioral responses, not accounted for in the SIR model, into Spatial-SIR to highlight how their effects on the infection diffusion depend on geographic factors.<sup>9</sup>

Section 7 presents five implications for empirical analysis we learn from our model simulations. We note that research exploiting geographic variation in longitudinal data to study

---

<sup>6</sup>See e.g., [Atkeson \(2020\)](#), [Eichenbaum et al. \(2020\)](#), [Brotherhood et al. \(2020\)](#), and [Jarosch et al. \(2020\)](#)

<sup>7</sup>See, e.g., [Argente et al. \(2020\)](#), [Antràs et al. \(2020\)](#), [Birge et al. \(2020\)](#), [Bognanni et al. \(2020\)](#), [Brady et al. \(2020\)](#), [Cuñat and Zymek \(2020\)](#), [Fajgelbaum et al. \(2020\)](#), [Giannone et al. \(2020\)](#), [Glaeser et al. \(2020a\)](#).

<sup>8</sup>Relatedly, [Acemoglu et al. \(2020a\)](#), [Alfaro et al. \(2020\)](#), [Azzimonti et al. \(2020\)](#) extend SIR to explicitly model the epidemic dynamics in networks; [Ellison \(2020\)](#) allows for heterogeneity of the contact process between subpopulations.

<sup>9</sup>Models of rational agents limiting contacts to reduce the risk of infection are relatively scarce in epidemiology; [Fenichel \(2013\)](#) [Weitz et al. \(2020\)](#) are prominent examples, see also [Funk et al. \(2010\)](#) and [Verelst et al. \(2016\)](#) for surveys. Most importantly, the formal modeling of behavioral responses has not yet broken into the large forecasting models which represent the core of the discipline as e.g., [Balcan et al. \(2009\)](#), [Balcan et al. \(2010\)](#), [Chinazzi et al. \(2020\)](#), and [Ferguson et al. \(2020\)](#). Not surprisingly, behavioral responses are instead central to epidemiological models in economics. Early contributions in this respect include [Geoffard and Philipson \(1996\)](#) and [Goenka and Liu \(2012\)](#); while recent work includes [Acemoglu et al. \(2020b\)](#), [Aguirregabiria et al. \(2020\)](#), [Argente et al. \(2020\)](#), [Bethune and Korinek \(2020\)](#), [Farboodi et al. \(2020\)](#), [Fernandez-Villaverde and Jones \(2020\)](#), [Gans \(2020\)](#), [Greenwood et al. \(2019\)](#), [Keppo et al. \(2020\)](#), [Toxvaerd \(2020\)](#), as well as several of the papers cited above modeling spatial extensions of SIR. See also [Bisin and Moro \(2020a\)](#) for a formal introduction to SIR with forward-looking rational-choice agents. None of these papers discuss the effects of the interaction between behavioral and spatial factors in the spread of an infection which we show has important implications for the dynamics of herd immunity and possibly for the effects of Non-Pharmaceutical Interventions (NPI).

the effect of policy interventions or other covariates on the epidemic outcomes must deal with time-varying heterogeneity across locations that is hard to control for without imposing a specific structure. We simulate data from Spatial-SIR models to highlight how structural estimates of SIR models fail to incorporate restrictions imposed by the spatial structure. We quantitatively assess the implied error of the effect of a lockdown policy that this mis-specification generates.

Reduced-form methods exploiting the different time and location of the implementation of the intervention have been adopted to study the effect of policy interventions or other covariates on the epidemic outcomes, often using a Difference-in-differences design.<sup>10</sup> We run Difference-in-differences regressions on the simulated data to highlight how such an approach fails to capture the heterogeneity of the effects by time and locations. Perhaps not surprisingly, it also fails to identify the average treatment effect of the policy if the data does not include both treated and untreated units at locations with different geographic characteristics, or if treatment occurs at different times in different locations.<sup>11</sup>

Our paper highlights that empirical research can gain from imposing the cross-location restrictions implied by the epidemiological models. Spatial and time heterogeneity render the reduced-form analysis of these data problematic. Short of directly estimating structural models, simulated data can be exploited in an integrated structural/reduced-form empirical analysis.

## 2 Invariances in the SIR Model

We first introduce the standard SIR model as a benchmark to evaluate the role of adding spatial structure. The society is populated by  $N$  agents that are ex-ante identical. Let  $\mathcal{S} = \{S, I, R\}$  denote the individual state-space, indicating Susceptibles, Infected, and Recovered.<sup>12</sup> Let  $h_t = [S_t, I_t, R_t]$  denote the distribution of the population across the state-space at time  $t$ . The following transitions govern the dynamics of  $h_t$ : i) a Susceptible agent becomes infected upon contact with an infected, with probability  $\beta I_t/N$ ; ii) an agent in-

---

<sup>10</sup>Without any attempt at being exhaustive, see, e.g., Allcott et al. (2020b), Allcott et al. (2020a), Bartik et al. (2020), Borri et al. (2020), Borsati et al. (2020), Breen and Ermisch (2021), Chernozhukov et al. (2020), Chetty et al. (2020), Couture et al. (2020), Courtemanche et al. (2020a), Courtemanche et al. (2020b), Crucini and O’Flaherty (2020), Dave et al. (2020), Fang et al. (2020), Gapen et al. (2020), Glaeser et al. (2020b), Goolsbee and Syverson (2020), Gupta et al. (2020), Hsiang et al. (2020), Juranek and Zoutman (2020), Kong and Prinz (2020), Maloney and Taskin (2020), Mangrum and Niekamp (2020), Pepe et al. (2020), Sears et al. (2020), Wilson (2021), Wright et al. (2020), Yilmazkuday (2020), Ziedan et al. (2020).

<sup>11</sup>Goodman-Bacon and Marcus (2020) describes some threats to the validity of DiD-design in the analysis of NPIs to fight the spread of COVID-19. See also Callaway and Sant’Anna (2020) and Callaway and Li (2021) for approaches to mitigate the problems arising from these threats.

<sup>12</sup>In Bisin and Moro (2020b) we expand the state space to better capture some relevant aspects of the SARS-CoV-2 epidemic by adding Symptomatics and Dead. This expansion of the state space is inconsequential for studying the effects of geographical characteristics but adds realism, helping to study its policy implications.

fectured at  $t$ , can recover at any future period with probability  $\rho$ ; iii) a Recovered agent never leaves this state (this assumes that Recovered agents are immune to infection).

The SIR can be solved analytically.<sup>13</sup> The equations describing its dynamics in discrete time are

$$\Delta I_t = \beta S_t \frac{I_t}{N} - \rho I_t, \quad \Delta R_t = \rho I_t, \quad S_t + I_t + R_t = N. \quad (1)$$

Parameter  $\beta$  is to be interpreted as the infection rate and is related to  $\mathcal{R}_0 = \beta/\rho$ , the number of agents a single infected agent infects, on average, at an initial condition  $R_0 = 0$ ,  $I_0 > 0$ . The infection rate  $\beta$  can be decomposed as the infection rate per-contact between a Susceptible and an Infected,  $\pi$ , and the number of contacts per unit of time,  $c$ :  $\beta = \pi c$  (in the continuous time limit). Distinguishing the roles of the number of contacts and the contagion rate is conceptually important to avoid interpreting  $\mathcal{R}_0$  and  $\beta$  as structural parameters of the model. In the Spatial-SIR we introduce in the next section, they are the result of virological, geographical and, in Section 6, behavioral factors.

We highlight three invariance properties of the dynamics of the SIR model, with respect to initial conditions and to the spatial and virological parameters driving the dynamics. We study the robustness of these invariances to the introduction of a spatial structure.

**Stationary state invariance to initial conditions.** Given any  $I_0 > 0$ ,<sup>14</sup> the dynamic system converges to a unique stationary state. Namely, the size of the initial outbreak,  $I_0$ , does not affect the stationary state. This stationary state of Infected is  $I_\infty = 0$ , while the stationary state of Recovered (the fraction of the population infected in the course of the epidemic),  $0 < R_\infty/N < 1$ , is characterized uniquely in terms of  $\mathcal{R}_0 = \beta/\rho$ , as the solution of the following fixed point equation:

$$\frac{R_\infty}{N} = -\frac{1}{\mathcal{R}_0} \ln\left(1 - \frac{R_\infty}{N}\right). \quad (2)$$

**Transitional dynamics invariance to initial conditions (in the limit  $I_0/N \rightarrow 0$ ).** The dynamics of  $h_t/N$  depends on initial conditions only via  $I_0/N$ . It is then invariant as the fraction of infected at the initial condition converges to zero,  $I_0/N \rightarrow 0$ . In particular, the peak of infected cases in this limit is

$$\frac{I^{peak}}{N} = 1 - \frac{1}{\mathcal{R}_0} (1 + \log \mathcal{R}_0). \quad (3)$$

**Transitional dynamics invariance to contacts and probability of contagion, keeping  $\beta$  constant.** The dynamics of  $h_t/N$  depends on the number of contacts  $c$  and probability of contagion,  $\pi$ , but is invariant to changes in  $c$  and  $\pi$  that leave  $\beta = \pi c$  constant.

<sup>13</sup>See e.g., [Hethcote \(2000\)](#), [Moll \(2020\)](#), [Neumeier \(2020\)](#).

<sup>14</sup>Initial conditions are uniquely represented by  $I_0$ , since  $R_0 = 0$  and  $S_0 = N - I_0$ .

If the epidemic is governed by SIR, these invariances provide restrictions of the model which are testable with cross-city data i) when different cities have different initial conditions (infection outbreaks)  $I_0$ ; and/or ii) when differences in the number of contacts and in the probability of contagion map into differences in  $\beta = \pi c$  and  $\rho$  across cities.<sup>15</sup>

### 3 The Spatial-SIR model

We add a spatial dimension to SIR by locating agents in a 2-dimensional space, which we call the “city.” Agents are ex-ante identical in terms of demographics and symmetric in terms of location. Agents are randomly located initially, and every day  $t = [0, T]$ , they travel distance  $\mu$  toward a random direction. By doing so, they potentially meet new individuals every day, therefore  $\mu$  is an abstraction of the speed in which agents find new contacts, potentially in a different state than their previous neighbors. Two agents come into contact when they are at a geographical distance closer than  $p$ .<sup>16</sup>

Spatial-SIR is represented by the following transitions: i) a Susceptible agent in a location within distance  $p$  from the location of an Infected becomes infected with probability  $\pi$ ; ii) an Infected agent can Recover at any period with probability  $\rho$ ; iii) Recovered agents never leave these states. The resulting dynamical system is difficult to characterize formally.<sup>17</sup> We turn then to simulations.

Table 1 reports the calibrated parameters we use in the baseline model. We calibrate transitions between states,  $[S, I, R]$ , to various SARS-CoV-2 parameters from epidemiological studies, notably, e.g., [Ferguson et al. \(2020\)](#). We calibrate  $\beta$  (in its components  $\pi$  and  $c$ ) and the agents’ daily travel distance  $\mu$  to data on average contacts from [Mossong et al. \(2008\)](#) and to match estimates of initial (prior to policy interventions) growth rates of the epidemic in Lombardy, Italy.<sup>18</sup> The calibration is performed as follows.

---

<sup>15</sup>Variation across virological characteristics can, in principle, be studied with data across different epidemic. In this paper, we concentrate mainly on geographical variation across cities.

<sup>16</sup>The spatial behavior of agents as postulated in Spatial-SIR is mechanical and it abstract from the network structure, e.g., home, work, city, which characterizes real world behavior. This is to focus more clearly and directly on highlighting the fundamental effects of spatial behavior in determining the dynamics of the epidemic, as well as their stylized dependence on various geographical characteristics (outbreaks, population size, density, agents’ movements); see [Bisin and Moro \(2020b\)](#) for an extension of Spatial-SIR allowing for a network structure.

<sup>17</sup>In Appendix A we write it as a Markov chain on configurations in space, along the lines of interacting particle-system models ([Kindermann and Snell, 1980](#); [Liggett, 2012](#)). Some properties are obtained by analogy to the physics of percolation on lattices; see [Grassberger \(1983\)](#), [Tomé and Ziff \(2010\)](#). For local-interaction models in Economics see [Blume et al. \(2011\)](#), [Conley and Topa \(2007\)](#), [Glaeser and Scheinkman \(2001\)](#), [Özgür et al. \(2019\)](#). In Appendix A we also discuss the mathematical formulation of this class of models in continuous time and space, as reaction-diffusion equations systems (see e.g. [Chinviriyasit and Chinviriyasit \(2010\)](#) and [Wu et al. \(2017\)](#)).

<sup>18</sup>We acknowledge the substantial uncertainty in the literature with respect to even the main epidemiological parameters pertaining to this epidemic. As we noted in the introduction, this is less damaging when aiming at understanding mechanisms and orders-of-magnitude rather than at precise forecasts.

**Table 1: Calibrated parameter values: baseline model**

Parameter	Notation	Value
(1) number of people	$N$	26,600
(2) initially infected	$I_0$	30
(3) prob. of recovery	$\rho$	0.154
(4) average contacts per day	$c$	13.5
(5) contagion radius	$p$	0.013
(6) contagion probability	$\pi$	0.054
(7) mean distance traveled	$\mu$	0.034

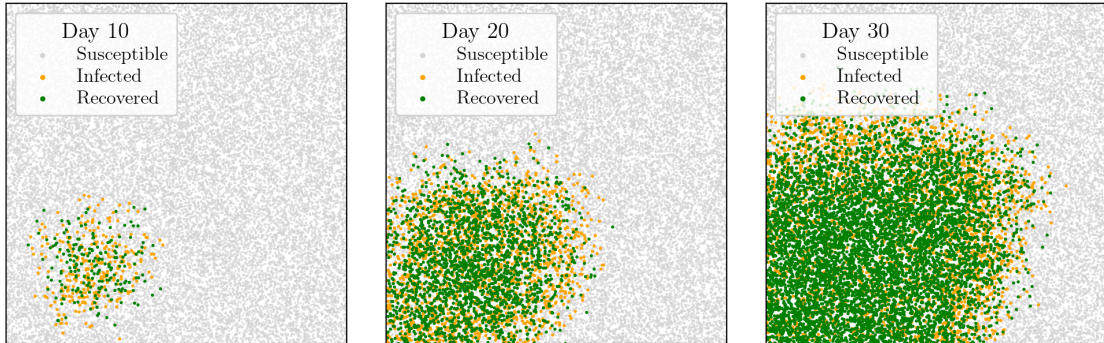
**(1)-(2) Population geography, initial conditions** We choose  $N$  so that our simulations converge in a reasonable time (see the beginning of Section 5.1 for a description of how the model scales in size). We place people initially on a unit square drawing their  $x$  and  $y$  coordinates independently from a Uniform distribution  $\sim U[0, 1]$ <sup>19</sup>. At all  $t > 0$ , individuals are relocated at distance  $\mu$  from their location at  $t - 1$ , in a direction randomly drawn from a Uniform distribution  $\sim U[0, 2\pi]$ . When individuals get close to the boundary, the movement is constrained to point to a direction opposite to the boundary. At time  $t = 0$  we set 30 individuals in Infected state; all others are Susceptibles. In all specification excepts those reported in Figures 3(b) (4) and 8 the Infected at  $t = 0$  are those initially located closest to location  $[x = 0.25, y = 0.25]$ .

**(3) Transition away from the infected state,  $I$ .** The probability any agents transitions away from state  $I$  is  $\rho$ , hence the average time an agents stays in state  $I$  is  $T_{\text{inf}} = 1/\rho$ . We set  $\rho$  to match a theoretical moment which holds exactly at the initial condition in SIR. Recall  $\mathcal{R}_0$  denotes the number of agents a single infected agent at  $t = 0$  infects, on average. Let  $g_0$  denote the growth rate of the number of infected agents at  $t = 0$ . Then, in SIR,  $(\mathcal{R}_0 - 1)/T_{\text{inf}} = g_0$  for  $t \rightarrow 0$ . For SARS-CoV-2,  $\mathcal{R}_0$  is reasonably estimated between 2.5 and 3.5. (Huang et al. (2020), Remuzzi and Remuzzi (2020), Zhang et al. (2020), Paules et al. (2020)). The daily rate of growth of infections  $g$  is estimated to be between 0.35 and 0.15 by Kaplan et al. (2020), Alvarez et al. (2020), and Ferguson et al. (2020). This implies, from the equation above for  $g_0$ , that  $T_{\text{inf}}$  is between 4 and 7 days (respectively for  $\mathcal{R}_0$  between 2.5 and 3.5). Ferguson et al. (2020) use 6.5 days, which we use to set  $\rho = 1/6.5$ .

**(4)-(5) Contagion circle radius.** The contagion radius,  $p$ , is not separately identified from the average number of contacts,  $c$ . We set it to match the the average number of contacts observed in demographic surveys. Mossong et al. (2008) suggests an average of

<sup>19</sup>In the simulation with heterogeneous density we set initial locations at a distance from the center drawn randomly from a Normal distribution  $\sim N(0, 1)$  and direction drawn from a Uniform distribution  $\sim U[0, 2\pi]$ .

**Figure 1: Geographic progression of infections and recoveries, baseline model**



Note: Position of individuals in the city at day 10, 20, and 30 since the start of the infection, color-coded by the agent’s state, model calibrated with baseline parameters

13.5 contacts every day.

**(6)-(7) Infection and contact rates.** After setting parameters (1)-(5), we calibrate the remaining parameters  $\pi$  (hence  $\beta = \pi c$ ) and  $\mu$  to match the daily growth rates of the dynamics of infections observed in the first 35 days of epidemic using data for Lombardy, Italy. Since the number of infections is not observed, we match the growth rates of deaths in the data. This is justified when, as we assumed, the case fatality rate is constant, and Death follows infection after a constant lag on average. Appendix Figure B.1 illustrates goodness of fit.

The dynamics of an epidemic in the Spatial-SIR model cannot be analyzed in closed form, but Figure 1 illustrates this dynamics over time and space at the calibrated parameters. The epidemic spreads exponentially from the location of the outbreak.<sup>20</sup> In the next sections, we compare the dynamics under SIR and Spatial-SIR.

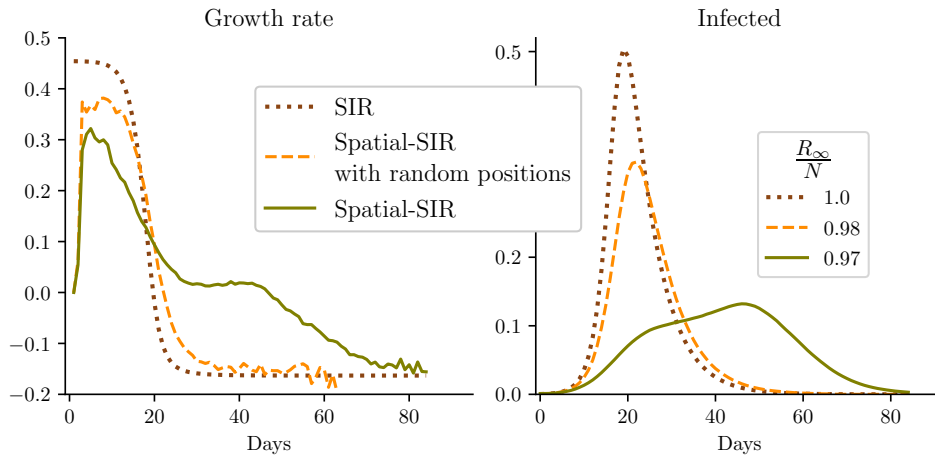
## 4 Local Herd Immunity

To understand how Spatial-SIR differs from SIR, we simulate the evolution over time of the infection growth rates and the fraction of active cases (that is, infected agents,  $I_t/N$ ) in both models. We show that geography and people’s movements in Spatial-SIR generate local interactions and matching frictions creating a form of *local herd immunity*, absent in SIR. Formally, in SIR, random matching implies that the probability that any Susceptible

<sup>20</sup>All our simulations converge to a unique distribution over the state space  $[S, I, R]$ .



**Figure 2: SIR and Spatial-SIR comparison of infection dynamics**



*Note:* Growth rate of Infected (left panel) and Infected at date  $t$  as a fraction of the population (right panel). The legend of the right panel indicates the steady state fraction of Recovered. Solid green lines: spatial-SIR baseline calibration; dashed orange lines: baseline Spatial-SIR, except with individuals’ geographic locations drawn randomly every day with the same rules used to draw initial locations; brown dotted lines: SIR with infection and recovery rates equivalent to the ones implied by the baseline calibration of Spatial-SIR.

agent is infected at time  $t$  is  $\beta I_t/N$ .<sup>21</sup> In Spatial-SIR, this probability is not common across susceptible agents as it depends on the distribution of agents and their states across space, say  $H_t$ .<sup>22</sup> We can then describe the probability that a susceptible agent is infected *on average* in the Spatial-SIR model as  $\beta\lambda(H_t)$ , for a well-defined function  $\lambda$  which encodes the effects of local herd immunity.<sup>23</sup>

More specifically, in Figure 2 we compare simulations of Spatial-SIR, with the parameters we calibrated for our baseline model, and of SIR, with  $\beta$  equal to our calibrated value of the contagion rate multiplied by the average number of daily contacts, implied our calibrated city’s population density and contagion radius. The Spatial-SIR displays lower growth rates than SIR initially. “Local herd immunities” slow down the diffusion of infection in the early stages and accelerate it afterward (as aggregate herd immunity is delayed); in other words,

<sup>21</sup>We exploit the continuous-time approximation for ease of exposition. In discrete-time (hence in the simulations), the probability that a susceptible agent is infected per unit of time after  $c$  contacts is  $1 - (1 - \pi I/N)^c$ .

<sup>22</sup>Formally,  $H_t$  is a stochastic process whose realization at  $t$  maps any agent (or any location, in an equivalent formulation) into a state  $h \in \{S, I, R\}$ .

<sup>23</sup>In Spatial-SIR, the probability that a susceptible agent is infected is not linear in  $I_t/N$  as in SIR. As a consequence, the Spatial-SIR cannot be mapped into a SIR model for some  $\beta$  depending on  $t$  (see the Appendix for a more formal description of Spatial-SIR). We discuss this point more explicitly when we draw the empirical implications of our analysis in Section 7.

$\lambda(H_t)$  is initially smaller and then larger than  $I_t/N$ . In the figure, the dashed yellow lines report growth rates for a version of Spatial-SIR with agents placed in a random location in the city *every day*, mimicking the random matching aspect of SIR and therefore minimizing the formation of local herd immunities. As expected, the effect of local herd immunity is much weaker in this model.<sup>24</sup>

The effect of local herd immunity indicates that the “reduced-form” nature of SIR models is missing a potentially important role of matching frictions and, more generally, of local dynamics. Similar considerations can be obtained looking at  $\mathcal{R}_0$  (a random variable in Spatial-SIR because the number of contacts of an individual is random). Replicating simulations of our baseline Spatial-SIR, we estimate  $\mathcal{R}_0$  as the average number of people infected by those infected during the first five days. This estimate is within the range used to calibrate transition rates in many studies (between 2.5 and 3.5) but is highly volatile. In 20 random replications of the model, the average  $\mathcal{R}_0$  is 3.17, with a standard deviation of 0.58. However, in Spatial-SIR this volatility does not translate into similarly different aggregate outcomes as predicted by standard SIR. The fraction of people ever infected in steady-state averages to 0.97 in the 20 replications, with a standard deviation of 0.001. This suggests that, in our model,  $\mathcal{R}_0$  loses its role as the fundamental driving parameter of the epidemic: outcomes are also highly sensitive to characteristics of the initial cluster of infection.<sup>25</sup> While the infection rate during the initial stages is uniquely determined by the structural parameters  $\mathcal{R}_0$  and  $\rho$ , which are (relatively) independent of the spatial structure of the model, the infection dynamics rests on the spatial local interaction structure. The growth rate of the infection declines early on following a form of local herd immunity. Indeed, this is what we observe in the data, and we set parameters to match.

## 5 Outbreaks, Population Size, Density, and Agents’ Movements

In this section, we study the comparative dynamics of the spread of an epidemic as it depends on various relevant geographical characteristics across cities, which determine matching frictions. We compare the effects of these geographical characteristics on both stationary and transitional dynamics in Spatial-SIR with those in SIR. More specifically, we capture the properties of the stationary states by the fraction of recovered,  $R_\infty/N$  (as  $I_\infty/N = 0$  and  $S_\infty/N = 1 - R_\infty/N$ ). We capture the properties of the transitional dynamics, on the other hand, by the time it takes to for an outbreak to reach the peak of active cases, a measure

---

<sup>24</sup>Local herd immunity is determined by accessibility to susceptible population. Therefore, it is not limited to geography as in our model, but also by the network structure of contacts; see [Bisin and Moro \(2020b\)](#) for a Spatial-SIR model where heterogeneous agents are structured in a network. The role of local herd immunity in both SIR and Spatial-SIR outcomes is also evident when comparing the effects of Non-Pharmaceutical Interventions such as lockdown policies (see Section 7).

<sup>25</sup>Outcomes are also somewhat sensitive to the precise location of the initially infected; therefore the simulation dynamics we display in simulations of Spatial-SIR report averages of 20 random replications of the models

of the speed of the epidemic, and the height of the peak of active cases as a fraction of the population,  $I^{peak}/N$ , a measure of the intensity of the epidemic.

The geographical characteristics we concentrate on are outbreaks, population size, density, and agents' movements. The number of initial outbreaks and the population size appear directly in both Spatial-SIR and SIR. In Spatial-SIR, however, it is the entire distribution of initial outbreaks, not just the number  $I_0$ , which affects the dynamics of the epidemic. As a consequence, the effects of population size in Spatial-SIR are mediated by the initial conditions regarding the distribution of outbreaks. City density instead appears directly in Spatial-SIR, but has only an indirect counterpart in SIR. City density in fact can be thought of as affecting parameter  $\beta$  proportionally in SIR, because density affects the number of contacts  $c$  by  $c = d\Psi$ , where  $\Psi$  is the contagion area of any (susceptible) individual (which is maintained constant in all simulations), and  $\beta = \pi c$ . Finally, agents' movements appears directly in Spatial-SIR, as the average distance traveled every day, but is absent in SIR.

Generally, we show that all these geographic characteristics affect the epidemic differently in Spatial-SIR and SIR. In particular, we highlight that the simulated dynamics of Spatial-SIR do not satisfy some of the invariance properties of the SIR dynamics we have delineated in Section 2. Fundamentally, the effect of local herd immunity on the dynamics is a function of geographic characteristics which, abusing notation, we denote as  $g = (I_0, N, d, \mu)$ . We represent this by writing  $\lambda(H_t; g)$ .<sup>26</sup> This analysis of the effects of geography on the dynamic of the epidemic has some clearcut implications that empirical cross-city studies of epidemic dynamics should account for; we discuss these in some detail in Section 7.

## 5.1 Outbreaks

The epidemic dynamics in SIR only depends on initial conditions through  $I_0/N$  (first and second invariance in Section 2). This implies scaling in size, that is, the invariance of the dynamics over  $x$ -times larger cities with  $x$ -times as many initial infected. In Spatial-SIR, however, as we noted, the space of the initial conditions is larger, as it includes the spatial distribution of infections. Indeed, the spatial distribution matters greatly in Spatial-SIR. Scaling in size obtains in Spatial-SIR only if initial infection outbreaks are appropriately homogeneously distributed across space.<sup>27</sup>

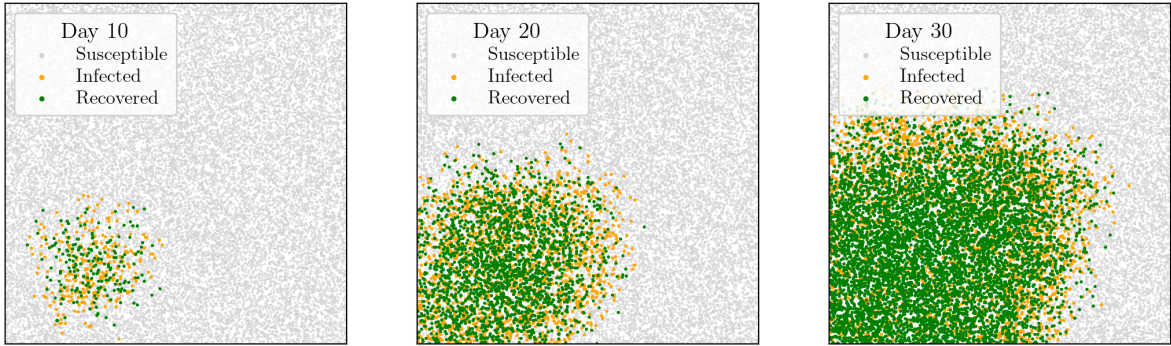
To understand the role of the distribution of outbreaks in Spatial-SIR, in Figure 3 we compare the progression of the contagion in the baseline city (panel 3(a), reproduced from Figure 1) with the progression in a city in which the infected agents are not placed on an

---

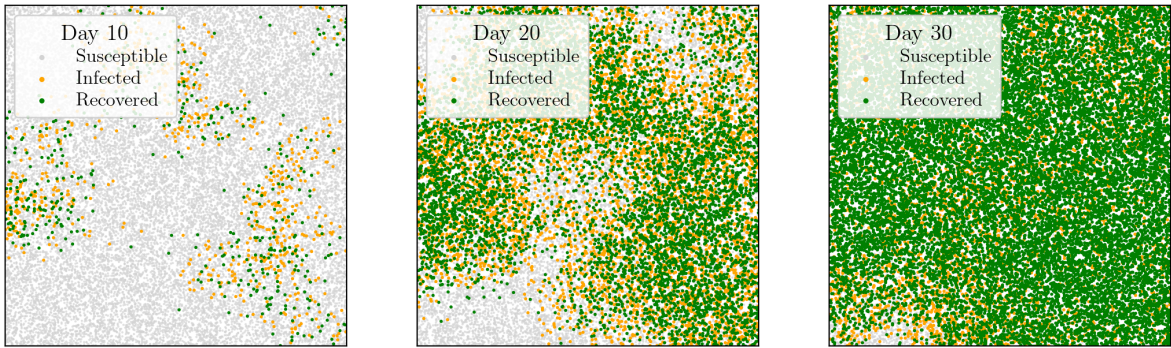
<sup>26</sup>As we noted, the dependence of  $\lambda$  on  $H_t$  changes over time and cannot be expressed in closed form; the same holds for its dependence on  $g$ . In other words, the dependence of Spatial-SIR on geographic characteristics is structurally distinct from what we would obtain in a SIR model allowing for  $\beta$  to depend on  $g$ . We discuss this point more explicitly in Section 7.

<sup>27</sup>In Appendix Figure B.2 we compare the progression of the contagion between the baseline city and a city with four times the population and the area (so that density is constant), and with four initial clusters of the same size as in the baseline located in symmetric locations. The progression of the infection is nearly identical, barring minor effects due to the randomness of people's locations and movement.

**Figure 3: Geographic progression of infections and recoveries**



**(a) Baseline model**



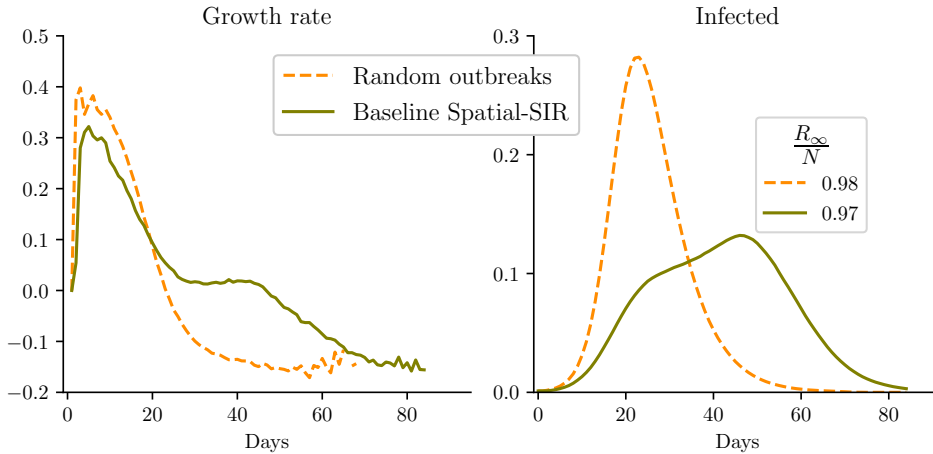
**(b) Initial clusters of contagion randomly located**

Note: Position of individuals in the city at day 10, 20, and 30 since the start of the infection, color-coded by the agent's state

initial cluster but are split in random locations (panel 3(b)). While in the baseline model contagion is relatively concentrated by day 30, contagion is much more widely spread when the initially infected are randomly located. Figure 4 summarizes the infection dynamics in these two simulations: the progression of active cases is faster when the initial cluster is randomly located, reaching a higher peak of active cases (28% rather than 13%) earlier (on day 23 rather than on day 46). However, the fraction of Recovered at the stationary state is about the same, 97%.<sup>28</sup>

<sup>28</sup>Please note that in this and other figures, the scales of the x and y axes may differ to improve the visualization. The x-axes are scaled to the number of days it takes for the epidemic to converge to steady-state; the y-axes are scaled to a level slightly greater than the maximum value of the displayed variables

**Figure 4: Infection dynamics: initial outbreaks in random locations**



*Note:* Growth rate of Infected (left panel) and Infected at date  $t$  as a fraction of the population (right panel). The legend of the right panel indicates the steady state fraction of Recovered. Solid green lines: spatial-SIR baseline calibration; dashed orange lines: Spatial-SIR with initial outbreaks in random locations

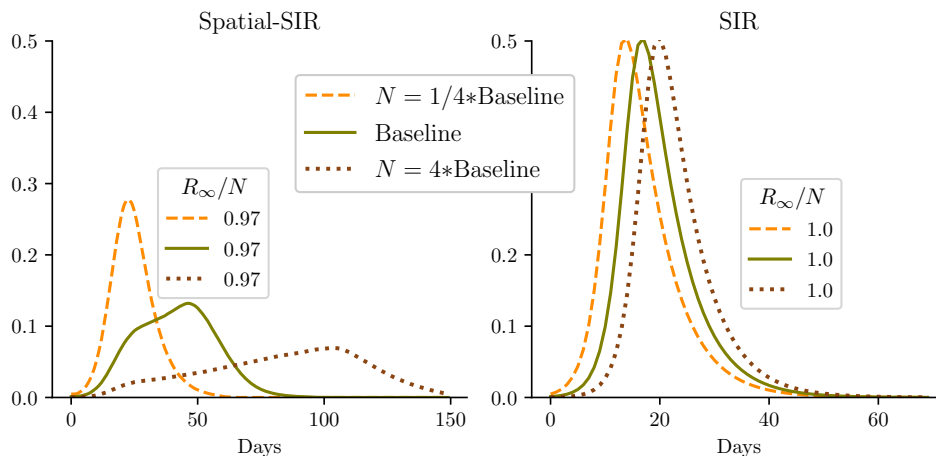
## 5.2 Population size

In this section, we study the effects of changing population size  $N$  and city area proportionally so as to keep the city density constant, fixing the number of initial infections,  $I_0$ , and their spatial distribution (assumed homogeneous to highlight the effects of  $N$ ).

In the SIR model, these changes have no effect on the stationary state (first invariance in Section 2) and a vanishing effect on the transitional dynamics for small enough  $I_0$  (second invariance in Section 2). More generally, increasing size  $x$ -times (for a given  $I_0$ ) increases the peak by only  $-1/\mathcal{R}_0 \ln x$  percentage points in SIR. On the other hand, while population size has no effect on the stationary state of the epidemic in Spatial-SIR as well, it has an important effect on transitory dynamics. In Figure 5 we report infections, as a fraction of the population, for both models in three cities: the baseline, a city with one fourth, and one with four times the baseline population. As noted, changing population size does not change the stationary state fraction of infected, approximately equal to 97 percent of the population, independently of city size, in both models.

With regards to the transitional dynamics, their dependence on size is minimal in SIR, hardly visible in fact in our simulation (see right panel — more populated cities take longer to reach the peak only because we keep the initial conditions constant). In Spatial-SIR instead, the curve displaying the fraction of active cases is flatter in larger cities (left panel). More specifically, in Spatial-SIR, with respect to SIR, the same difference in population size reduces the peak to about one quarter (from .28 to .07 active cases). Relatedly, the

**Figure 5: Infection dynamics: population size in Spatial-SIR (left panel) and SIR (right panel)**



*Note:* Infected at date  $t$  as a fraction of the population. Spatial-SIR (left panel) SIR (right panel) Solid green solid lines: spatial-SIR baseline calibration; dashed orange line: one quarter the baseline population; brown dotted lines: four times the baseline population, keeping density constant. The legend inside each panel indicates the steady-state fraction of Recovered.

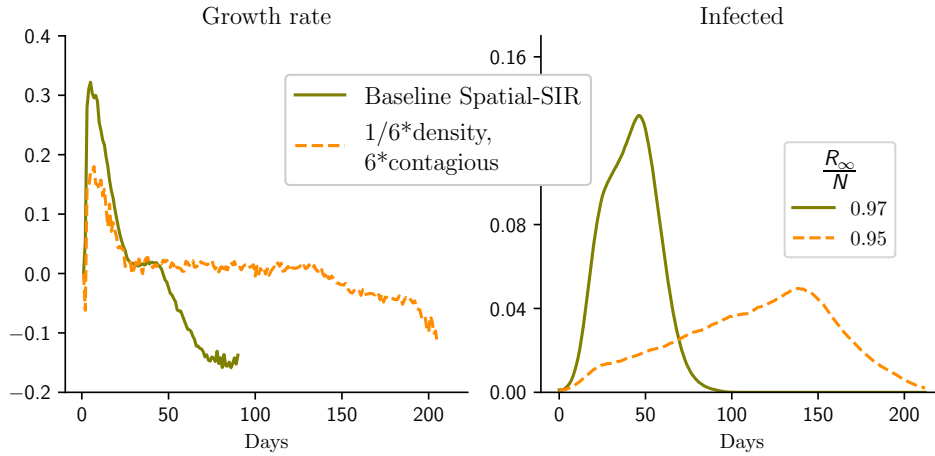
time (in days) to reach the peak in larger cities goes from 15 to 22 days in SIR, and from 23 to 104 days in Spatial-SIR. Note that the peak is lower in larger cities *as a fraction of the population*, which suggests that resources such as hospital beds, ventilators, etc. . . , distributed proportionally to population size are less likely to be binding in large cities.

### 5.3 City density

In this section, we describe the role of city density on the dynamics of the epidemic. We first show that city density - which is proportional to contacts - plays a distinct role in Spatial-SIR from the inverse of the probability of infection, breaking the third SIR invariance in Section 2. This is shown in Figure 6 where the baseline calibrated Spatial-SIR is compared with an environment with six times the probability of infection and 1/6th the density: the effect on the dynamic of infection is different both qualitatively and quantitatively.

Density is a crucial determinant of the dynamics of the epidemic because, together with the contagion rate, it determines the average number of infections occurring on a given date. Increasing density while keeping the contagion radius the same increases the number of contacts that each infected individual has on a given day. In fact, in Spatial-SIR, changing city density while keeping the contagion rate and the population size constant has important effects on both the stationary state and the transitional dynamics of the

**Figure 6: Infection dynamics: density, constant  $\beta$**



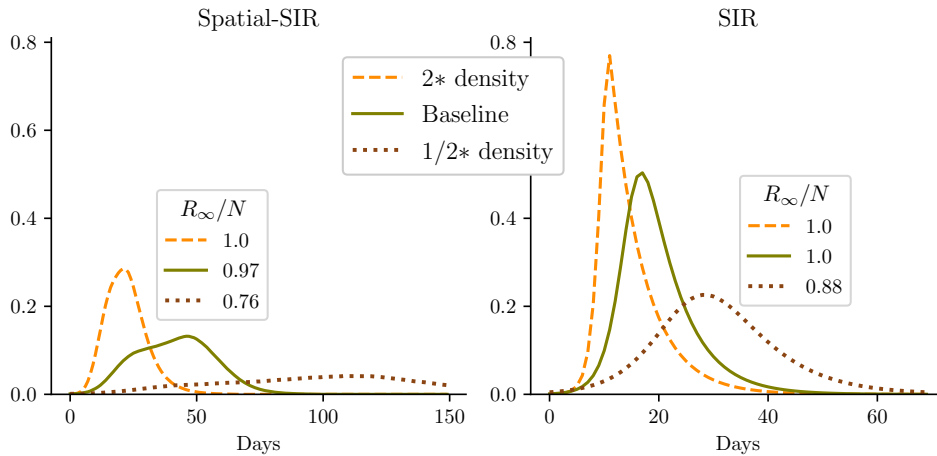
*Note:* Growth rate of Infected (left panel) and Infected at date  $t$  as a fraction of the population (right panel). The legend of the right panel indicates the steady state fraction of Recovered. solid green lines: spatial-SIR baseline calibration; dashed orange lines: baseline Spatial-SIR with one-sixth the density and six times the contagion rate as the baseline model.

epidemic, as illustrated in Figure 7 (left panel). We see that indeed the peak of active cases is increasing in, and very sensitive to, density: halving density relative to the baseline dramatically flattens the peak of the infection (by more than one-half, after more than twice as many days from the outbreak).

Comparing the effects of density (through contacts, and in turn through  $\beta$ ) in SIR and Spatial-SIR reveals some interesting patterns. The effects are qualitatively similar, as can be seen comparing the simulations on the left panel (Spatial-SIR) and on the right panel (SIR) of Figure 7. Quantitatively, however, density has much larger effects in Spatial-SIR. In the stationary state limit, the fraction of recovered in the densest economy is 1.13 times larger than in the least dense economy in SIR and 1.31 times larger in Spatial-SIR. This is the case also with regards to the transition, i) the peak of active cases in the densest economy is 3.45 times larger than in the least dense economy in SIR and 7.25 times larger (resp. 5.4 times smaller) in Spatial-SIR, and ii) the number of days to the peak on infections in the densest economy is four times smaller than in the least dense economy in SIR and 5.4 times smaller in Spatial-SIR.

We also report, in Figure 8, simulations comparing the baseline city (of constant density across space) with an identical city of heterogeneous density, declining from center to periphery (and initial cluster of infection in the center - the right panel illustrates the initial condition, with each grey dot representing a susceptible individual). While the stationary states of these cities differ minimally, the city with heterogeneous density experiences a

**Figure 7: Infection dynamics: density**



*Note:* Infected at date  $t$  as a fraction of the population. Spatial-SIR (left panel) and SIR (right panel). Solid green lines: spatial-SIR baseline calibration; dashed orange lines: baseline Spatial-SIR with twice the density; brown dotted lines: baseline Spatial-SIR with one half the density, keeping population size constant. The legend inside each panel indicates the steady-state fraction of Recovered.

smaller peak substantially earlier than the baseline. Namely, heterogeneous density induces a faster-growing epidemic initially, which subsequently slows down, reaching herd immunity earlier (at about 40% infected rather than 70%). This example illustrates a more general *selection* mechanism operating when agents are heterogeneous (for example, in age, socio-economic and professional characteristics, preferences for social interactions): those more susceptible to the spread of the infection (in this simulation, those living in denser regions) are selected to achieve herd immunity earlier.<sup>29</sup>

#### 5.4 Movements in the city

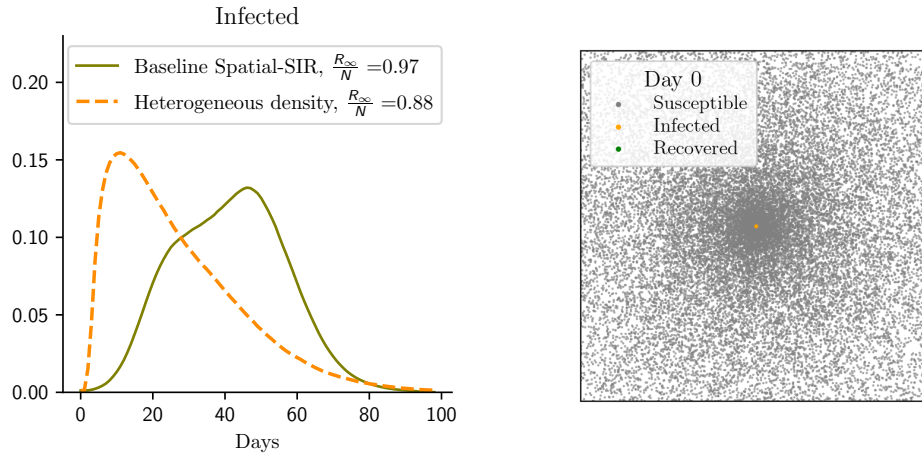
In Spatial-SIR, the parameters controlling the variation in the random movement contribute to explaining the cross-city heterogeneity in the dynamics of the epidemic (these parameters obviously do not appear in SIR). In this section we study the effects of the distance traveled every day by each agent,  $\mu$ , on the epidemic dynamics.<sup>30</sup> Changing these parameters affects the average number of contacts in the city. As we argued, the average number of contacts in the city has an effect that is similar to city density.

<sup>29</sup>See [Gomes et al. \(2020\)](#) and [Britton et al. \(2020\)](#) for related theoretical analyses.

<sup>30</sup>Given our calibrated contagion radius, if all people in the city were uniformly spaced from each other, contagion would not occur. All infections in the baseline model occur because random placement and movement generate clusters of people closer to one another than the infection radius.



**Figure 8: Infection dynamics: heterogeneous density**



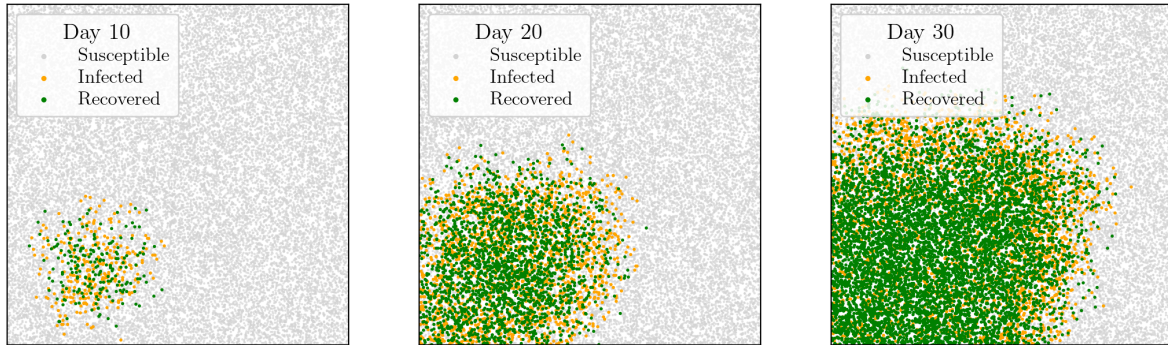
*Note:* Infected at date  $t$  as a fraction of the population. Solid green line: spatial-SIR baseline calibration; dashed orange line: Spatial-SIR with heterogeneous density. The right panel illustrates the initial spatial distribution of Susceptibles in the model with heterogeneous density. The legend on the left panel indicates the steady-state fraction of Recovered.

To provide an intuition of the dependence of the epidemic on the movement speed of agents in the city, in Figure 9 we compare the progression of contagion of the baseline model with the same progression in the extreme case when agents do not move. The infection spreads slowly. As the contagion expands, clusters of susceptible (non-infected) people are clearly visible in the rightmost panel as large white spots within the green cloud. This is less likely to occur when people move, which is why the speed of movement also affects the steady-state, as illustrated in Figure 10.

With constant density and people randomly moving around the city, the average number of contacts is constant, but local herd immunity plays a fundamental role, and the dynamic of the infection changes with speed. With faster speed, infected people are more likely to find uninfected locations, making it less likely for people in these locations to stay immune until the steady-state.

The speed of people’s movement around the city and the number of initial clusters have a very similar effect on outcomes, because if people move very fast, at the beginning of the infection they generate new clusters quickly.

**Figure 9: Geographic progression of infections and recoveries**



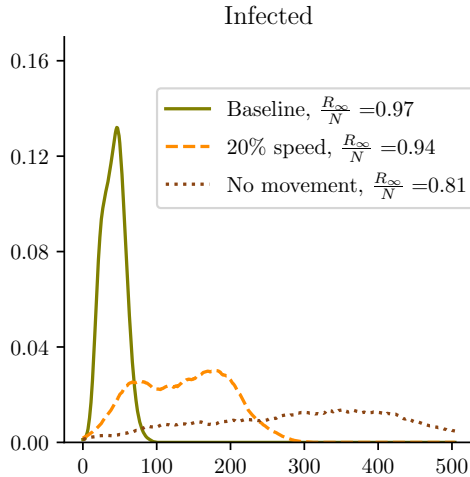
**(a) Baseline model, days 10, 20, 30**



**(b) No movement, days 10, 20, 30, 50, 150, 250**

Note: Each figure displays the position of individuals in the city at a given day since the start of the infection, color-coded by the agent's state

**Figure 10: Infection dynamics: movement speed**



*Note:* Infected at date  $t$  as a fraction of the population. Solid green line baseline Spatial-SIR; dashed orange line: Spatial-SIR with 20% speed of movement relative to baseline, and Spatial-SIR with no movement (dotted brown line). The legend indicates the steady-state fraction of Recovered.

## 6 Behavioral Spatial-SIR

As we discussed the introduction, most epidemiological models do not formally account for behavioral responses to the epidemic. In those models, as in the analysis in the previous sections, the number of daily contacts in the population,  $c$ , is a constant.

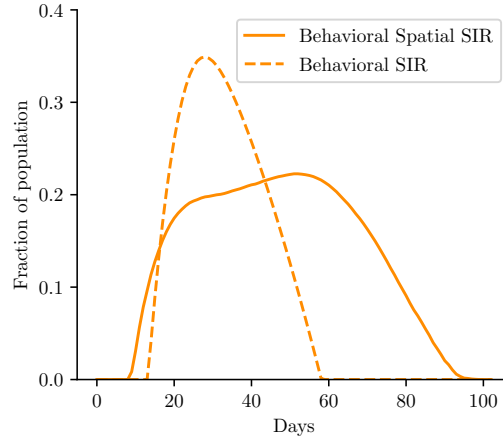
In this section, we model agents responding to the epidemic by choosing to limit their contacts. Following Keppo et al. (2020), we introduce a reduced-form behavioral response, represented by a function  $0 \leq \alpha(I_t) \leq 1$ , acting as a proportional reduction of the agent's contacts (the population density  $d$  multiplied by the contagion area of an infected individual  $\Psi$ ) as a function of the number of infected in the population:

$$c = \alpha(I_t)d\Psi, \quad \alpha(I_t) = \begin{cases} 1 & \text{if } I_t \leq \underline{I} \\ \left(\frac{\underline{I}}{I_t}\right)^{1-\phi} & \text{if } I_t > \underline{I} \end{cases}. \quad (4)$$

We calibrate the dynamics of the epidemic allowing for the behavioral response (4) in both SIR and Spatial-SIR.<sup>31</sup> In simulations of the behavioral models we set  $\phi = 0.88$  in (4) as estimated by Keppo et al. (2020) using Swine flu data, and assume people start responding to the spread of the contagion very soon by setting  $\underline{I} = 0.01$ . In simulations of the standard SIR model with behavioral responses, we use the same parameters. We rank individuals by

<sup>31</sup>We calibrated the SIR model as in the simulations in Section 4 for this comparison.

**Figure 11: Reduction in contacts in behavioral models**



Note: fraction of the population induced to avoid contact due to behavioral responses

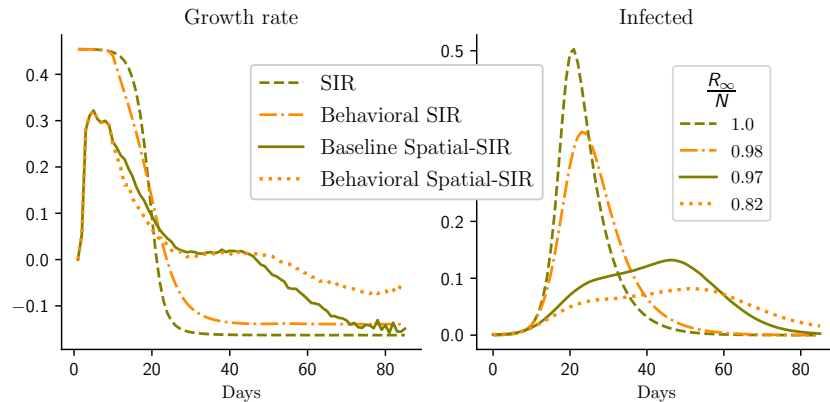
risk aversion and assume that anyone who self-isolates when  $I_t = \hat{I}_t$  also self-isolates when  $I_t > \hat{I}_t$ , inducing persistence in the identity of the individuals who respond.

The simulated reduction in the number of contacts due to the behavioral response is reported in Figure 11. In both the behavioral SIR and the behavioral Spatial-SIR, as the infection spreads, the reduction in contacts due to the behavioral reaction increases. Then, as herd immunity begins and the number of Infected declines, the reduction of contacts decreases, and contacts return towards the initial (pre-infection) state. In Spatial-SIR, however, the reduction in contact is (about a third) smaller, but its peak drags for much longer. This is because local herd immunity starts showing its effects earlier, inducing agents to stop reducing contacts earlier, but then builds up more slowly.

Figure 12 reports simulations of the effects of behavioral responses by comparing models with and without behavior (green and orange lines, respectively). The peak of the growth rate of infections in the behavioral Spatial-SIR is lower than in the behavioral SIR, but it plateaus when declining, after about 25 days. In fact, the growth rate of infections has a much longer declining plateau in the behavioral Spatial-SIR than in the Spatial-SIR, a dramatic display of the effects of the interaction of behavioral and spatial factors in the spread of infection via the dynamics of local herd immunity.<sup>32</sup>

<sup>32</sup>The interaction of behavioral and spatial factors also modulate the effects of various non-pharmaceutical interventions, e.g., lockdown rules, as studied in Bisin and Moro (2020b). In particular, substantiating a “Lucas critique” argument, the cost of naive discretionary policies ignoring the behavioral responses of agents and firms depend fundamentally on the local herd immunity effects due to the spatial dimension of the dynamics of the epidemic.

**Figure 12: Infection dynamics in SIR and Spatial-SIR, effects of behavioral responses**



*Note:* Growth rate of Infected (left panel) and Infected at date  $t$  as a fraction of the population (right panel). The legend of the right panel indicates the steady state fraction of Recovered. Solid green lines: spatial-SIR baseline calibration; dashed green lines: SIR; dotted orange lines: Behavioral Spatial-SIR; dash-dotted orange lines: Behavioral SIR.

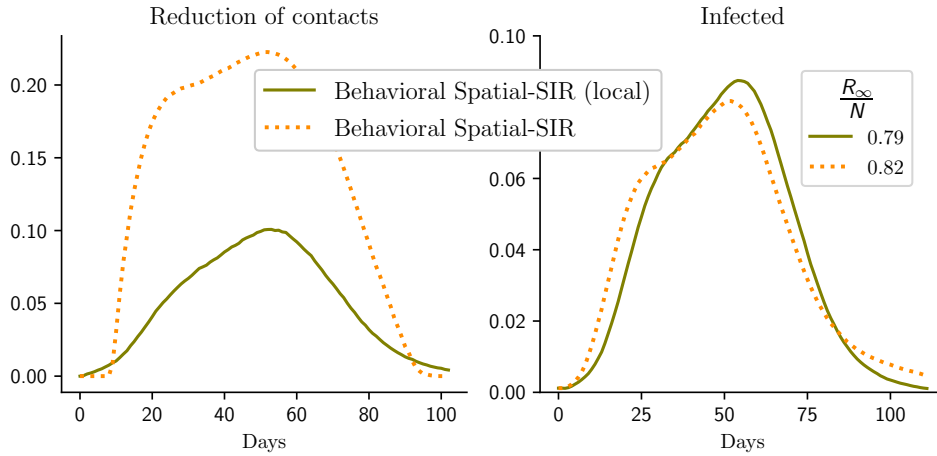
In both SIR and Spatial-SIR, not surprisingly, the qualitative effects of behavioral response is to reduce the spread of infection, lowering the peak of infected. In Spatial-SIR, however, behavior also has the effect of slowing down the operation of herd immunity. As the number of contacts returns to normal, the behavioral response has lasting effects in the stationary state, reducing total cases more in Spatial-SIR than in SIR. While we do not report simulations to this effect, we notice here the important fact that the behavioral response, when derived from the agents' choice, depends on geographical characteristics  $g$  as well, and these affect contacts. We denote the behavioral response then as  $\alpha(I_t; g)$ .

Figure 12 highlights the differential effects of behavioral responses on SIR. The behavioral response is not only much stronger in Spatial-SIR, but qualitatively different when comparing both infection growth rates and the fraction of active cases. The peak of active cases in Spatial-SIR is 1/3 with respect to SIR, but the decline of the infection after the peak is slower. This is the result of the composition of the behavioral response,  $\alpha(I_t; g)$ , and the local herd immunity factor  $\lambda(H_t; g)$ . The first acts on the number of contacts, while the second acts on the distribution of infected between the contacts.

## 6.1 Behavioral Spatial-SIR with local reactions

Spatial-SIR allows for a further dimension of the interaction between behavior and local space. Consider the case in which the behavioral reaction of an individual at time  $t$  depends only on the fraction of people infected in a circular neighborhood centered at her location

**Figure 13: The effect of local behavioral responses**



*Note:* Infected at date  $t$  as a fraction of the population. Behavioral Spatial-SIR (solid green line), Behavioral Spatial-SIR with local responses (dotted orange line) (right panel). The left panel illustrates the reduction in contacts (fraction of population). The legend of the right panel indicates the steady state fraction of Recovered.

and of radius equal to the contagion radius.<sup>33</sup>

Results are reported in Figure 13, which reproduces for comparison the results of the behavioral model (the orange dotted line in Figure 12). The heterogeneity of infection rates across neighborhoods has much larger and interesting effects when the behavioral reaction of the agents is local, that is, when it depends on the infection rate in the neighborhood. In this case, as seen in Figure 13, the higher reduction in contacts during the first and the last days of the epidemic is very short-lived. For most of the epidemic, agents reduce their contacts substantially less (by more than half) with respect to the case in which they react to the infection rate of the whole population because it occurs only in neighborhoods with infections. Interestingly, these composition effects induce relatively small quantitative effects on the population growth rate of infections over time. In other words, local behavioral reactions appear to “save” on the reduction in contacts for (almost)-given dynamics of infections.

<sup>33</sup>We induce persistence in the identity of those that self-isolate again by ranking individuals by risk aversion. If the formula predicts a reaction by  $x_i\%$  individuals in a neighborhood centered around person  $i$ , person  $i$  self-isolates if her risk aversion is below the  $x$ -th percentile among her neighbors.

## 7 Implications for Empirical Analysis

We summarize several implications of our analysis to guide empirical research using panel data about the diffusion of an epidemic.<sup>34</sup> We first discuss structural estimates of a formal epidemic model. We then consider estimates of the causal effects of a policy (typically, a Non-Pharmaceutical Intervention (NPI), e.g., a lockdown), which in many applications adopt a Difference in Differences (DiD) design.

### 7.1 Structural Estimates

Consider panel data on the dynamics of an infection across different geographic units  $i$ . The econometrician observes the geographic characteristics  $g_i = [I_{i,0}, N_i, d_i, \mu_i]$  of each city  $i$ , and data on the dynamics of the infection,  $I_{i,t}, R_{i,t}$  (hence  $S_{i,t}$ ).

**Cross-city restrictions in the standard SIR.** With these data, consider estimating a SIR model without behavioral effects as in standard epidemiological studies. Consider the following specification of the growth rate of infections, derived directly from (1):

$$\frac{\Delta I_{i,t}}{I_{i,t}} = \beta_{i,t} \frac{S_{i,t}}{N_i} - \rho, \quad (5)$$

$$\text{where } \beta_{i,t} = \beta_i = \pi c_i, \quad c_i = d_i \Psi. \quad (6)$$

This formulation highlights how model restrictions could be exploited for empirical analysis. Indeed, Equations (5-6) impose falsifiable cross-city restrictions, e.g., along the lines of the invariances we identified in Section 2. Several other empirically testable restrictions can be obtained directly from the closed-form representation of the dynamics; e.g., the growth rate of the fraction of infected in a city is, *ceteris paribus*, proportional to the density of the city.

**Identifying local spatial factors.** If the data are generated by a Spatial-SIR, however, accounting for spatial structure becomes important to capture matching frictions through local social interactions, as shown in Section 5. The dynamics of the infection takes the form

$$\frac{\Delta I_{i,t}}{I_{i,t}} = \beta_{i,t} \lambda(H_t; g_i) S_{i,t} - \rho, \quad (7)$$

$$\text{where } \beta_{i,t} = \beta_i \text{ as in (6).}$$

---

<sup>34</sup>For illustration, we concentrate on density as the geographic factor in this analysis. As we showed in the paper, the distribution of outbreaks and the speed of movement of the agents also have systematic effects on the dynamics of an epidemic and are maybe harder to measure. Proxies like airport activity for the number of outbreaks, the distribution of socio-economic characteristics for the distribution of outbreaks, the use of public transportation for the movement of agents, could be fruitfully adopted in both reduced-form and structural estimates.

The main driver of the differential effects in Spatial-SIR is local herd immunity. Geographic characteristics  $g_i$  mediate the relationship between parameters and model outcomes without a parametric expression for the function  $\lambda$ , nor for the transition matrix of the stochastic process  $H_t$ , making it challenging to identify the effects of geography from infection strength separately. However, one can use the structure of the model to match data with model predictions using simulation methods. Alternatively, one could use simulations to estimate  $\lambda(H_t; g_i)$  which can be used as a correction to the (much faster to simulate) dynamics of the SIR model, to estimate (7-6).<sup>35</sup>

**Identifying behavioral responses.** Finally, accounting for agents’ behavioral response to the epidemic in the data is generally of first order importance, as we discussed in Section 6. In this case, the formal representation of the dynamics of the infection, Equation (5) in SIR and (7) in Spatial-SIR, are unchanged, but the number of contacts is endogenous and (6) is modified to

$$\beta_{i,t} = \pi c_{i,t}, \quad c_{i,t} = \alpha(I_t; g_i) d_i \Psi. \quad (8)$$

This amplifies the issues we highlighted so far, requiring a new identification strategy, notably, to handle the dependence of  $\beta_{i,t}$  on  $t$ . In SIR the parameters predict the infection dynamics precisely. For example, there is a one-to-one correspondence between initial infection growth rates and the peak. Deviations from such dynamics can non-parametrically identify  $\pi$  from  $\alpha(I_t; g_i)$ . Parametric identification can be achieved by assuming a functional form for  $\alpha(I_t; g_i)$  along the lines of (4). In Spatial-SIR the full specification is (7-8). Identification in this case can rely on simulation methods, as needed for the identification of spatial factors.<sup>36</sup> Evidence of agents’ movements, using “Big-Data” from Google, Safegraph, and Cuebig could also provide useful empirical strategies for identifying behavioral responses from infection dynamics exploiting restrictions imposed by Spatial-SIR.<sup>37</sup>

To examine the importance of local spatial factors and behavioral responses in empirical research exploiting structural models of epidemic, we present some simulation experiments. In particular, we report on the biases occurring when estimating (5) on data generated by a Spatial-SIR, with and without behavioral responses of agents to the epidemic.

Figure 14 reports the values of the estimates of the coefficients on variable  $S_{i,t}/N_i$  for cities with different densities from 0.5 to 1.5 relative to the baseline specification density (we do not report confidence intervals because they are very small). We simulate each city 5 times and stop the simulations after the first 80 days, for a total of 400 observations per density level. Interestingly, the mis-specification bias in the estimation is small in cities with small density (where spatial and behavioral factors are less relevant); but very large for

---

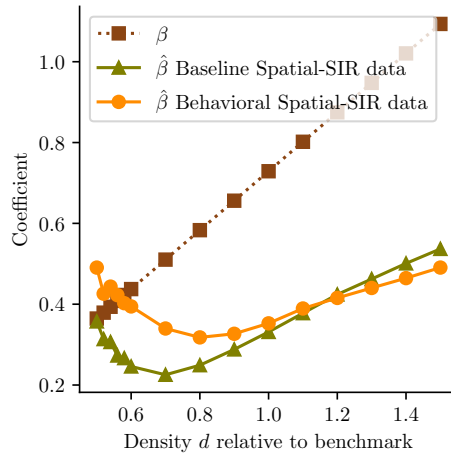
<sup>35</sup>Note that, besides density, various distributions of outbreaks can be easily included in the estimation of a Spatial-SIR (but not in the estimation of the SIR).

<sup>36</sup>Fernandez-Villaverde and Jones (2020) adopt simulation methods to estimate parameters separately for each location without imposing geographic restrictions.

<sup>37</sup>See, e.g., Farboodi et al. (2020).



Figure 14: Estimates of  $\beta$  from simulated data

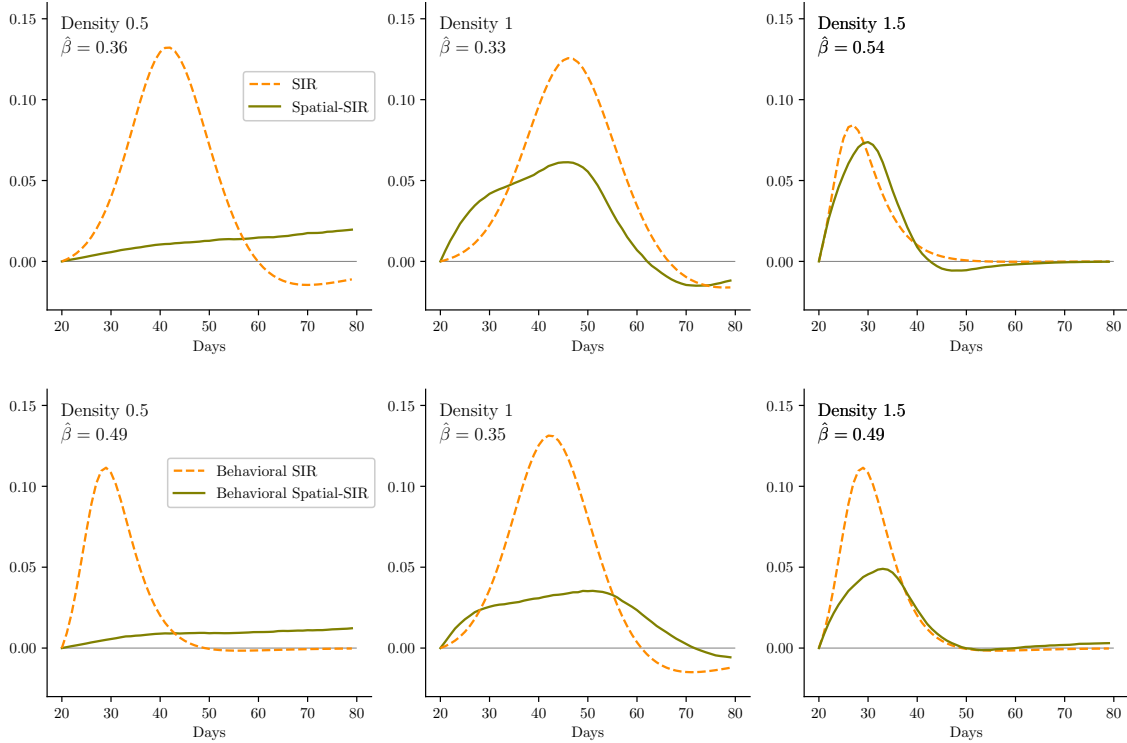


Note: Estimates of  $\beta$  using simulated data from models with baseline parameters, except for different levels of population density from 50% to 150% of the baseline density, obtained by varying the city size. Green line: Spatial-SIR, Orange line: Behavioral Spatial-SIR, The brown line reports the corresponding SIR infection parameter  $\beta = c\pi d$ .

larger densities:  $\hat{\beta}$  is about half the actual  $\beta$  of the data generating process for cities with 50% larger density than the baseline.

In Figure 15 we illustrate the effects of this mis-specification bias in predicting the dynamics of active cases after a policy intervention (a 25% lockdown at day 20). The solid green lines illustrate the dynamics of the *difference* in active cases between the Spatial-SIR model without and with the policy. The dashed yellow lines report the same dynamics in the SIR model, at the estimated values  $\hat{\beta}$ . We report results for only three density levels, and for the models without (three top panels) and with behavioral responses (bottom panels). The implied errors in the effects of the policy are large and have a complex dynamic structure, to the point of changing in sign in some instances. Note that when density is high (rightmost panels),  $\beta$  is high, therefore both SIR and Spatial-SIR have relatively fast dynamics (not visible in this figure). In this case, the effect of a lockdown is quantitatively similar, at least in the model without behavior. Models with relatively low density (smaller  $\beta$ ), despite being more precisely estimated (see Figure 14), generate relatively fast dynamics in SIR, but not in Spatial-SIR, implying a substantial difference in the effect of the policy, as the leftmost panels illustrate. Therefore, we conclude that it is hard to capture the dynamics implied by the spatial structure and the induced local herd behavior with a SIR model with a general reduction in the accessibility of the susceptible population (a lower  $\beta$ ).

**Figure 15: Infected at  $t$ , difference between models without and with lockdown**



*Note:* Difference between infected at date  $t$  without and with a 25% lockdown policy applied at date  $t = 20$  (fraction of total population). The SIR models are simulated with infection parameter  $\hat{\beta}$  estimated using the data simulated from the Spatial-SIR, as reported in Figure 14. The middle panels are computed with baseline parameters. The left and right-panels Spatial-SIR are computed by varying the citysize to obtain 50% and 150% population density, respectively.

## 7.2 Causal Effects of Policy Interventions

We simulate the policy intervention we just introduced (a 25% lockdown) as a treatment introduced in some cities at day 20, but possibly not in others or introduced at different times in different cities.

**Identifying the time-varying effect of geography in DiD studies of policy interventions.** Let  $\text{Treated}_{i,t}$  take value 1 if city  $i$  is treated by the policy at time  $t$ . Denoting with  $Y_{i,t}$  the variable of interest, the effects of  $\text{Treated}_{i,t}$  can be evaluated by means of a DiD design:

$$Y_{i,t} = \nu + \eta_i + \gamma_t + \delta \text{Treated}_{i,t} + \lambda X_{i,t} \quad (9)$$

where  $\nu, \eta_i, \gamma_t$  are time and location effects and  $X_{i,t}$  are controls. Our analysis suggests that the validity of this research design depends on both the variable of interest,  $Y_{i,t}$ , and the modeling framework. Consider attempting to estimate the effect of policies on the number of contacts,  $Y_{i,t} = c_{i,t}$ , for example. In the standard SIR model, this variable is proportional to  $\beta_i$  and depends on time  $t$  only through the treatment. In this case therefore specification (9) can flexibly capture variation in  $\beta_i$  by location.

Importantly, however, this specification fails to capture the dynamics of contacts in a SIR model with behavioral responses by agents. In this case in fact, contacts  $c_{i,t}$  is proportional to  $\alpha(I_{i,t}; g_i)$ , as in equation (8), and hence it is not separable in  $i$  and  $t$ , as required by the additive form  $\eta_i + \gamma_t$  in (9). Policy and agent behavior have separate effects on the dynamic of the epidemic both because behavioral responses have time-varying effects and because their effects interact with the effects of geography (a point generally disregarded in the few studies that try to account for behavioral responses).<sup>38</sup>

Our analysis shows that including a spatial dimension, as in the Spatial-SIR model, the vector of geographic factors  $g_i$  affects outcomes differently over time, introducing a time-varying heterogeneity that is not fully accounted for by time and location fixed effects as in (9), even after including a vector of geographic factors  $g_i$  among the set of regressors, and even if interacted with time. Furthermore, the direct effects of the treatment also depend on geographic characteristics: a lockdown, for instance, acts as a reduction of density and affects local herd immunity differently depending on the initial density.

Finally, identifying the effects of policy intervention in reduced-form econometric models are even more severe when the variable of interest  $Y_{i,t}$  is the growth rate of infections,  $\Delta I_{i,t}/I_{i,t}$ . In this case, in fact, the structure imposed by SIR, as in (5), generates time-varying heterogeneity not fully captured by time and location fixed effects even without behavioral responses or spatial dimensions and local herd immunity.

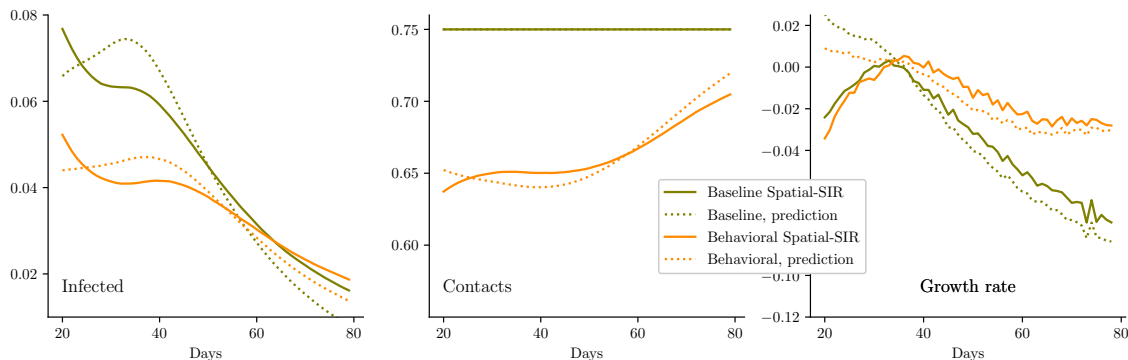
To better examine quantitatively the biases induced by estimating the “two-way” difference-in-differences equation (9) we produce several empirical experiments on simulated data. We estimate (9) using the same data we simulated to generate Figures 14 and 15; that is, several cities with densities from 0.5 to 1.5, with and without a 25% lockdown policy applied at date  $t = 20$ . First of all we study the case in which the data contain, for each density, both cities with and without lockdown; that is, treated and untreated. In this case, therefore, spatial heterogeneity is fully controlled for in Equation 9. Figure 16 reports, for three different outcome variables (active cases, contacts, growth rate of active cases, respectively in the left, center, and right panel) the dynamics after the lockdown is applied, and its predicted values using the estimated coefficients of Equation 9, averaged over time.

In this case, effectively controlling for spatial heterogeneity, the estimated equation i) predicts contacts correctly when the data does not contain a behavioral component; ii) but

---

<sup>38</sup>When the data is treated by policy, special care must be used because  $\alpha(I_t; g_i)$  is also not invariant to policy by a Lucas critique argument, even in the absence of geographical factors; see [Bisin and Moro \(2020b\)](#) for an analysis of this issue.

**Figure 16: True and predicted outcomes after a lockdown: time-heterogeneity only**

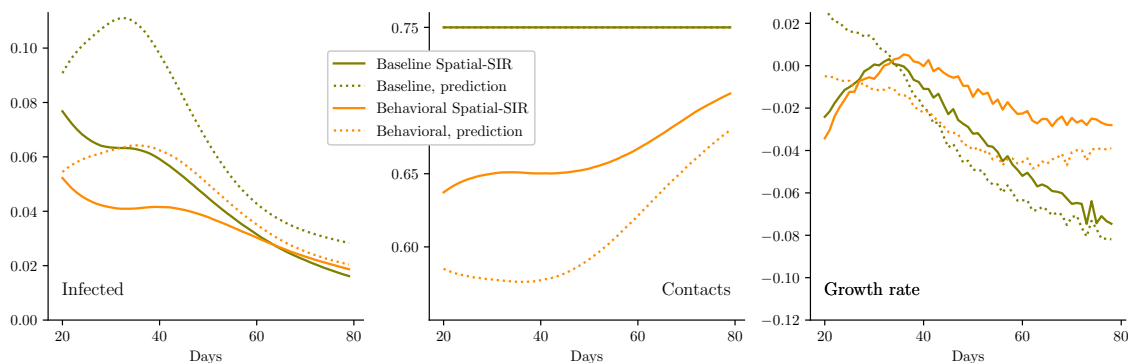


*Note:* Infected (left panel), contacts (middle panel), and growth rate of infected (right panel) at date  $t$  (fraction of population in Spatial-SIR (green lines), and Behavioral Spatial-SIR (orange lines)). The solid lines simulate the outcomes after a 25% lockdown imposed at  $t = 20$ . The dotted lines simulate the average by date  $t$  of the predicted value of  $Y_t$  after estimating (9) using the same data simulated from the Spatial-SIR models of Figure 14. The dataset includes 80 dates for each of 10 cities (5 treated and 5 untreated) per 11 density levels, for a total of 4400 observations.

does not predict contacts with behavior correctly; iii) nor it predicts correctly active cases or their growth rate (with or without behavior). By construction, Equation 9 captures the average treatment effects for all variables correctly (the average values of the solid and dotted lines are the same), but fails to capture the time-heterogeneity of the effects of the lockdown. Figure 17 reports the same simulation experiment, but with estimates computed from data that do not allow to fully control for spatial heterogeneity; that is, data do not contain both treated and untreated cities for each density level. Specifically, we computed an extreme case assuming that the econometricians observe only treated cities if their density is below 1, and only untreated cities if their density is equal or above 1. In this case, the figure illustrates that the estimated value of the treatment parameter fails to identify even the average treatment effect and prediction errors are quantitatively much larger. This is particularly the case when predicting active cases (left panel): after 35 days, the over-prediction of the estimated equation with respect to the simulated data is of the order of 50% (both with and without behavioral responses). But even contacts (central panel) are quite imprecisely predicted when the data are generated by the model with behavioral responses of agents (as noted, without behavior, Equation 9 predicts contacts precisely).

To better understand the source of prediction error in these exercises, let's focus on the prediction of active cases, that is, on the left panels in figures 16 and 17. At the time the policy is applied,  $t = 20$ , the fraction of infected in the population is growing (this is not visible in the figure). But the lockdown policy we simulated is strong enough to cause an

**Figure 17: True and predicted outcomes after a lockdown: both time and spatial heterogeneity**



*Note:* Infected (left panel), contacts (middle panel), and growth rate of infected (right panel) at date  $t$  (fraction of population in Spatial-SIR (green lines), and Behavioral Spatial-SIR (orange lines)). The solid lines simulate the outcomes after a 25% lockdown imposed at  $t = 20$ . The dotted lines simulate the average by date  $t$  of the predicted value of  $Y_t$  after estimating (9) using the data simulated from the Spatial-SIR models of Figure 14, but excluding data from cities treated by the policy if their density is equal or greater than 1, and excluding data from cities not treated by the policy if their density is less than 1. The resulting dataset includes 80 dates for each of 5 cities per 11 density levels, 25 treated and 30 untreated, for a total of 4400 observations.

immediate reduction in the number of infected. The estimated treatment from the regression is however constant over time, by construction, and hence the predicted effects (the dotted lines) continue to increase until the curve reaches its peak in the untreated units. This is the case both when spatial-heterogeneity is controlled for (Figure 16) and when it is not (Figure 17). In this second case, furthermore, the prediction overstates true infection levels. This is because, in our simulations, treatment in cities with low density is used to predict treatment in cities with high density, (for which we don't use treatment data in the estimation). In such cities, the “true” effect of the policy would be larger (in absolute terms), which is why the dotted line in Figure 16 is lower than the dotted line in Figure 17.

Finally, we simulate a more realistic scenario where the data includes cities of various densities and various treatment dates. We report results in Table 2. The table reports the “true” average treatment coefficient, the one we would obtain with data that contain an untreated city for each differently treated one, and the estimated coefficient with partial data.<sup>39</sup> As to be expected, prediction errors are large.

We conclude with some methodological suggestions for the empirical analysis of epidemi-

<sup>39</sup>We report two specifications, one where we did not include density as a regressor, and another in which we included density and density interacted with treatment.

**Table 2: Coefficients of treatment parameter estimates**

Outcome	Treated		Treated#Density	
	True	Estimated	True	Estimated
Without behavioral responses				
Infected	-383.605	-323.088		
	1006.816	214.176	-1390.421	-620.278
Contacts	-0.250	-0.250		
	-0.250	-0.250	-0.000	-0.000
Growth rate	-0.004	-0.009		
	0.249	0.258	-0.253	-0.306
With behavioral responses				
Infected	-324.242	-419.158		
	360.378	-319.468	-684.619	-114.588
Contacts	-0.197	-0.172		
	-0.262	-0.221	0.065	0.056
Growth rate	-0.013	-0.014		
	0.171	0.168	-0.184	-0.208

*Note:* Estimates of the “treatment” parameter  $\delta$  in (9) and of the interaction of treatment with density using data simulated from the Spatial-SIR models of Figure 14, with densities 0.5, 1, and 1.5. The “True” column reports estimates obtained from data with cities that are never treated, treated at date  $t \geq 15$ , and treated at date  $t \geq 40$ , for all density levels. The “Estimated” column reports estimates obtained from data including only cities with density 1 that are never treated, cities with density 0.5 that are treated at  $t \geq 40$ , and cities with density 1.5 that are treated at  $t \geq 15$ . For each outcome we report coefficients obtained from a specification without controls for density, and one with controls for density and density interacted with treatment. We omit standard errors, which can be made arbitrarily small by increasing the number of simulated cities.

ological data across cities and time and for the study of the effects of NPI’s in this context, as they are implied by our study in this paper. Spatial and time heterogeneity render the reduced-form analysis of these data problematic. Short of directly estimating structural models, simulated data can be exploited in an integrated structural/reduced-form empirical analysis.

Simulated data obtained from structural models, accounting for the relevant spatial and behavioral characteristics, are in principle very useful to indicate the sign and the size of prediction errors, as our empirical exercises make clear, e.g., when comparing the prediction errors in Figure 16 and 17.

More specifically, simulated data obtained from structural models are useful to capture spatial and time heterogeneity flexibly. For instance, when estimating Equation 9 to predict active cases, breaking the sample to distinguish earlier dates when cases are mostly growing in untreated cities from later ones when they are decreasing appears of first order importance to reduce time-heterogeneity. Similarly, geographic units of analysis should be chosen so that density and other geographic characteristics are relatively homogeneous.<sup>40</sup> But when this is not the case, the sample can be broken by density (and other relevant geographical characteristics) to better control for spatial heterogeneity. In these cases, while identifying the natural breaking points in the data may be difficult, simulations can be used to facilitate this procedure, for example, to identify which geographic variables are most relevant, in order to induce a parsimonious use of fixed effects and controls in Equation 9 and capture more of the structure of the dynamics of the epidemic while avoiding overfitting the data.

## 8 Conclusions

We study the effects of several stylized spatial factors identifying the fundamental role of local interaction and matching frictions as a determinant of the dynamics of epidemic. We highlight important implications for empirical studies on the diffusion of an epidemic, providing a framework for disentangling the effects of local interactions/matching frictions, behavioral responses of risk-averse agents, and policy interventions.

---

<sup>40</sup>For this reason, empirical analyses with data across countries involve additional concerns with respect to data across cities.

## References

- Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar**, “Testing, Voluntary Social Distancing and the Spread of an Infection,” Technical Report, National Bureau of Economic Research 2020. [3](#)
- , **Victor Chernozhukov, Iván Werning, and Michael D Whinston**, “A Multi-Risk SIR Model with Optimally Targeted Lockdown,” National Bureau of Economic Research 2020. [3](#)
- Aguirregabiria, Victor, Jiaying Gu, Yao Luo, and Pedro Mira**, “A Dynamic Structural Model of Virus Diffusion and Network Production: A First Report,” CEPR Discussion Paper No. DP14750 2020. [3](#)
- Alfaro, Laura, Ester Faia, Nora Lamersdorf, and Farzad Saidi**, “Social Interactions in Pandemics: Fear, Altruism, and Reciprocity,” Technical Report, National Bureau of Economic Research 2020. [3](#)
- Allcott, Hunt, Levi Boxell, Jacob C Conway, Billy A Ferguson, Matthew Gentzkow, and Benny Goldman**, “What Explains Temporal and Geographic Variation in the Early US Coronavirus Pandemic?,” Technical Report, National Bureau of Economic Research 2020. [4](#)
- , – , **Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Y Yang**, “Polarization and public health: Partisan differences in social distancing during the Coronavirus pandemic,” *National Bureau of Economic Research Working Paper*, 2020, *w26946*. [4](#)
- Alvarez, Fernando E, David Argente, and Francesco Lippi**, “A simple planning problem for covid-19 lockdown,” National Bureau of Economic Research 2020. [7](#)
- Antràs, Pol, Stephen J Redding, and Esteban Rossi-Hansberg**, “Globalization and Pandemics,” Technical Report, Harvard University Working Paper 2020. [3](#)
- Argente, David O, Chang-Tai Hsieh, and Munseob Lee**, “The Cost of Privacy: Welfare Effect of the Disclosure of COVID-19 Cases,” Technical Report, National Bureau of Economic Research 2020. [3](#)
- Atkeson, Andrew**, “What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios,” National Bureau of Economic Research 2020. [3](#)
- Azzimonti, Marina, Alessandra Fogli, Fabrizio Perri, and Mark Ponder**, “Pandemic Control in ECON-EPI Networks,” Federal Reserve Bank of Minneapolis Staff report 609 August 2020. [3](#)



- Balcan, Duygu, Bruno Gonçalves, Hao Hu, José J Ramasco, Vittoria Colizza, and Alessandro Vespignani**, “Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model,” *Journal of computational science*, 2010, *1* (3), 132–145. [3](#)
- , **Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J Ramasco, and Alessandro Vespignani**, “Multiscale mobility networks and the spatial spreading of infectious diseases,” *Proceedings of the National Academy of Sciences*, 2009, *106* (51), 21484–21489. [3](#)
- Bartik, Alexander W, Marianne Bertrand, Feng Lin, Jesse Rothstein, and Matt Unrath**, “Measuring the labor market at the onset of the COVID-19 crisis,” Technical Report, National Bureau of Economic Research 2020. [4](#)
- Bethune, Zachary A and Anton Korinek**, “Covid-19 infection externalities: Trading off lives vs. livelihoods,” *National Bureau of Economic Research Working Paper*, 2020. [3](#)
- Billari, Francesco C and Alexia Prskawetz**, *Agent-based computational demography: Using simulation to improve our understanding of demographic behaviour*, Springer Science & Business Media, 2012. [2](#)
- Birge, John R, Ozan Candogan, and Yiding Feng**, “Controlling Epidemic Spread: Reducing Economic Losses with Targeted Closures,” University of Chicago, Becker Friedman Institute for Economics Working Paper 2020. [3](#)
- Bisin, Alberto and Andrea Moro**, “Notes on Rational Forward Looking SIR,” mimeo, NYU 2020. [3](#)
- and – , “Spatial-SIR with Network Structure and Behavior: Lockdown Policies and the Lucas Critique,” Working paper, NYU and Vanderbilt University 2020. [4](#), [6](#), [10](#), [20](#), [27](#)
- Blume, Lawrence E, William A Brock, Steven N Durlauf, and Yannis M Ioannides**, “Identification of social interactions,” in “Benhabib, J, A. Bisin, and M. Jackson (eds.), Handbook of social economics,” Vol. 1, Elsevier, 2011, pp. 853–964. [6](#)
- Bognanni, Mark, Doug Hanley, Daniel Kolliner, and Kurt Mitman**, “Economic Activity and COVID-19 Transmission: Evidence from an Estimated Economic-Epidemiological Model,” mimeo, IIES 2020. [3](#)
- Borri, Nicola, Francesco Drago, Chiara Santantonio, Francesco Sobbrío et al.**, “The " Great Lockdown": inactive workers and mortality by Covid-19,” *medRxiv*, 2020. [4](#)
- Borsati, Mattia, Silvio Nocera, and Marco Percoco**, “QUESTIONING THE SPATIAL ASSOCIATION BETWEEN THE SPREAD OF COVID-19 AND TRANSIT USAGE IN ITALY,” Bocconi University Working Paper Series 2020. [4](#)

- Brady, Ryan Robert, Mike Insler, and Jacek Rothert**, “The Fragmented United States of America: The impact of scattered lock-down policies on country-wide infections,” *COVID Economics (43)*, 2020, pp. 42–94. [3](#)
- Breen, Richard and John Ermisch**, “The distributional impact of Covid-19: Geographic variation in mortality in England,” *Demographic Research*, 2021, *44*, 397–414. [4](#)
- Britton, Tom, Frank Ball, and Pieter Trapman**, “A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2,” *Science*, 2020. [16](#)
- Brotherhood, Luiz, Philipp Kircher, Cezar Santos, and Michèle Tertilt**, “An economic model of the Covid-19 epidemic: The importance of testing and age-specific policies,” *CEPR Discussion Paper No. DP14695*, 2020. [3](#)
- Bruch, Elizabeth and Jon Atwell**, “Agent-based models in empirical social research,” *Sociological methods & research*, 2015, *44* (2), 186–221. [2](#)
- Callaway, Brantly and Pedro H.C. Sant’Anna**, “Difference-in-Differences with multiple time periods,” *Journal of Econometrics*, 2020. [4](#)
- **and Tong Li**, “Policy Evaluation during a Pandemic,” arXiv working paper 2021. [4](#)
- Chernozhukov, Victor, Hiroyuki Kasahara, and Paul Schrimpf**, “Causal Impact of Masks, Policies, Behavior on Early Covid-19 Pandemic in the U.S.,” *medRxiv*, 2020. [4](#)
- Chetty, Raj, J Friedman, Nathaniel Hendren, and Michael Stepner**, “The economic impacts of COVID-19: Evidence from a new public database built from private sector data,” Opportunity Insights Working Paper. 2020. [4](#)
- Chinazzi, Matteo, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun et al.**, “The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak,” *Science*, 2020, *368* (6489), 395–400. [3](#)
- Chinviriyasit, Settapat and Wirawan Chinviriyasit**, “Numerical modelling of an SIR epidemic model with diffusion,” *Applied Mathematics and Computation*, 2010, *216* (2), 395–409. [6](#)
- Conley, Timothy G and Giorgio Topa**, “Estimating dynamic local interactions models,” *Journal of Econometrics*, 2007, *140* (1), 282–303. [6](#)
- Courtemanche, Charles J, Joseph Garuccio, Anh Le, Joshua C Pinkston, and Aaron Yelowitz**, “Did Social-Distancing Measures in Kentucky Help to Flatten the COVID-19 Curve?,” Working paper 2020. [4](#)

- Courtemanche, Charles, Joseph Garuccio, Anh Le, Joshua Pinkston, and Aaron Yelowitz**, “Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States.,” *Health Affairs*, 2020, *39* (7), 1237–1246. 4
- Couture, Victor, Jonathan I Dingel, Allison E Green, Jessie Handbury, and Kevin R Williams**, “Measuring Movement and Social Contact with Smartphone Data: A Real-Time Application to COVID-19,” Working Paper 27560, National Bureau of Economic Research July 2020. 4
- Crucini, Mario J and Oscar O’Flaherty**, “Stay-at-Home Orders in a Fiscal Union,” Technical Report, National Bureau of Economic Research 2020. 4
- Cuñat, Alejandro and Robert Zymek**, “The (Structural) Gravity of Epidemics,” CESifo Working Paper 2020. 3
- Dave, Dhaval, Andrew Friedson, Kyutaro Matsuzawa, Joseph J Sabia, and Samuel Safford**, “JUE Insight: Were urban cowboys enough to control COVID-19? Local shelter-in-place orders and coronavirus case growth,” *Journal of urban economics*, 2020, p. 103294. 4
- Desmet, Klaus and Romain Wacziarg**, “Understanding spatial variation in COVID-19 across the United States,” National Bureau of Economic Research 2020. 2
- Dunham, Jill Bigley**, “An agent-based spatially explicit epidemiological model in MASON,” *Journal of Artificial Societies and Social Simulation*, 2005, *9* (1). 2
- Duranton, Gilles and Diego Puga**, “The Economics of Urban Density,” *Journal of Economic Perspectives*, August 2020, *34* (3), 3–26. 2
- Eichenbaum, Martin S, Sergio Rebelo, and Mathias Trabandt**, “The macroeconomics of epidemics,” National Bureau of Economic Research 2020. 3
- El-Sayed, Abdulrahman M, Peter Scarborough, Lars Seemann, and Sandro Galea**, “Social network analysis and agent-based modeling in social epidemiology,” *Epidemiologic Perspectives & Innovations*, 2012, *9* (1), 1. 2
- Ellison, Glenn**, “Implications of Heterogeneous SIR Models for Analyses of COVID-19,” National Bureau of Economic Research Working Paper 27373 June 2020. 3
- Eubank, Stephen, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang**, “Modelling disease outbreaks in realistic urban social networks,” *Nature*, 2004, *429* (6988), 180–184. 2

- Fajgelbaum, Pablo, Amit Khandelwal, Wookun Kim, Cristiano Mantovani, and Edouard Schaal**, “Optimal lockdown in a commuting network,” Technical Report, National Bureau of Economic Research 2020. [3](#)
- Fang, Hanming, Long Wang, and Yang Yang**, “Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in china,” *National Bureau of Economic Research*, 2020. [4](#)
- Farboodi, Maryam, Gregor Jarosch, and Robert Shimer**, “Internal and External Effects of Social Distancing in a Pandemic,” Working Paper 27059, National Bureau of Economic Research April 2020. [3](#), [24](#)
- Fenichel, Eli P**, “Economic considerations for social distancing and behavioral based policies during an epidemic,” *Journal of health economics*, 2013, *32* (2), 440–451. [3](#)
- Ferguson, Neil, Daniel Laydon, Gemma Nedjati Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, ZULMA Cucunuba Perez, Gina Cuomo-Dannenburg et al.**, “Imperial College COVID-19 Response Team: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand,” Imperial College London 2020. [3](#), [6](#), [7](#)
- Fernandez-Villaverde, Jesus and Charles Jones**, “Estimating and Simulating a SIRD Model of COVID-19,” mimeo 2020. [2](#), [3](#), [24](#)
- Funk, Sebastian, Marcel Salathé, and Vincent AA Jansen**, “Modelling the influence of human behaviour on the spread of infectious diseases: a review,” *Journal of the Royal Society Interface*, 2010, *7* (50), 1247–1256. [3](#)
- Gans, Joshua S**, “The Economic Consequences of  $\hat{R} = 1$ : Towards a Workable Behavioural Epidemiological Model of Pandemics,” Working Paper 27632, National Bureau of Economic Research July 2020. [3](#)
- Gapen, Michael, Jonathan Millar, U Blerina, and Pooja Sriram**, “Assessing the effectiveness of alternative measures to slow the spread of COVID-19 in the United States,” *Covid Economics*, 2020, *40*, 46–75. [4](#)
- Geoffard, Pierre-Yves and Tomas Philipson**, “Rational epidemics and their public control,” *International economic review*, 1996, pp. 603–624. [3](#)
- Giannone, Elisa, Nuno Paixão, and Xinle Pang**, “The Geography of Pandemic Containment,” *Covid Economics: Vetted and Real-Time Papers*, October 2020, *52*, 68–95. [3](#)
- Glaeser, Edward and José Scheinkman**, “Measuring social interactions,” *Social dynamics*, 2001, pp. 83–132. [6](#)

- Glaeser, Edward L, Caitlin S Gorbach, and Stephen J Redding**, “How Much does Covid-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities,” mimeo 2020. 3
- , **Ginger Zhe Jin, Benjamin T Leyden, and Michael Luca**, “Learning from deregulation: the asymmetric impact of lockdown and reopening on risky behavior during COVID-19,” Technical Report, National Bureau of Economic Research 2020. 4
- Goenka, Aditya and Lin Liu**, “Infectious diseases and endogenous fluctuations,” *Economic Theory*, 2012, 50 (1), 125–149. 3
- Gomes, M Gabriela M, Ricardo Aguas, Rodrigo M Corder, Jessica G King, Kate E Langwig, Caetano Souto-Maior, Jorge Carneiro, Marcelo U Ferreira, and Carlos Penha-Goncalves**, “Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold,” *medRxiv*, 2020. 16
- Goodman-Bacon, Andrew and Jan Marcus**, “Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies,” *DIW Berlin Discussion Paper*, 2020. 4
- Goolsbee, Austan and Chad Syverson**, “Fear, lockdown, and diversion: comparing drivers of pandemic economic decline 2020,” Technical Report, National Bureau of Economic Research 2020. 4
- Grassberger, Peter**, “On the critical behavior of the general epidemic process and dynamical percolation,” *Mathematical Biosciences*, 1983, 63 (2), 157–172. 6
- Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michèle Tertilt**, “An equilibrium model of the African HIV/AIDS epidemic,” *Econometrica*, 2019, 87 (4), 1081–1113. 3
- Grefenstette, John J, Shawn T Brown, Roni Rosenfeld, Jay DePasse, Nathan TB Stone, Phillip C Cooley, William D Wheaton, Alona Fyshe, David D Galloway, Anuroop Sriram et al.**, “FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations,” *BMC public health*, 2013, 13 (1), 1–14. 2
- Gupta, Sumedha, Laura Montenovo, Thuy D Nguyen, Felipe Lozano Rojas, Ian M Schmutte, Kosali I Simon, Bruce A Weinberg, and Coady Wing**, “Effects of social distancing policy on labor market outcomes,” Technical Report, National Bureau of Economic Research 2020. 4
- Hethcote, Herben W.**, “The mathematics of infectious diseases,” *SIAM Review*, 2000, 2000, 42 (4), 599–653. 5

- Hsiang, Solomon, Daniel Allen, Sebastien Annan-Phan, Kendon Bell, Ian Boliger, Trinetta Chong, Hannah Druckenmiller, Andrew Hultgren, Luna Yue Huang, Emma Krasovich et al.**, “The effect of large-scale anti-contagion policies on the coronavirus (covid-19) pandemic,” *Nature*, 2020. 4
- Huang, Chaolin, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu et al.**, “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China,” *The lancet*, 2020, 395 (10223), 497–506. 7
- Hunter, Elizabeth, Brian Mac Namee, and John D Kelleher**, “A taxonomy for agent-based models in human infectious disease epidemiology,” *Journal of Artificial Societies and Social Simulation*, 2017, 20 (3). 2
- , – , and – , “A Comparison of Agent-Based Models and Equation Based Models for Infectious Disease Epidemiology.,” in “AICS” 2018, pp. 33–44. 2
- Jarosch, Gregor, Maryam Farboodi, and Robert Shimer**, “Internal and External Effects of Social Distancing in a Pandemic,” 2020. 3
- Juranek, Steffen and Floris Zoutman**, “The effect of social distancing measures on the demand for intensive care: Evidence on COVID-19 in Scandinavia,” CESifo Working Paper 2020. 4
- Kaplan, Greg, Ben Moll, and Gianluca Violante**, “Pandemics According to HANK,” *University of Chicago*, 2020. 7
- Keppo, Juusi, Marianna Kudlyak, Elena Quercioli, Lones Smith, and Andrea Wilson**, “The behavioral SIR model, with applications to the Swine Flu and COVID-19 pandemics,” in “Virtual Macro Seminar” 2020. 3, 19
- Kermack, William Ogilvy and Anderson G McKendrick**, “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*, 1927, 115 (772), 700–721. 2
- and – , “Contributions to the mathematical theory of epidemics. II. -The problem of endemicity,” *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 1932, 138 (834), 55–83. 2
- Kindermann, R and JL Snell**, “American Mathematical Society,” *Markov random fields and their applications*, 1980. 6
- Kong, Edward and Daniel Prinz**, “Disentangling policy effects using proxy data: Which shutdown policies affected unemployment during the COVID-19 pandemic?,” *Journal of Public Economics*, 2020, 189, 104257. 4

- Lavezzo, Enrico, Elisa Franchin, Constanze Ciavarella, Gina Cuomo-Dannenburg, Luisa Barzon, Claudia Del Vecchio, Lucia Rossi, Riccardo Manganelli, Arianna Loregian, Nicolò Navarin, Davide Abate, Manuela Sciro, Stefano Merigliano, Ettore Decanale, Maria Cristina Vanuzzo, Francesca Saluzzo, Francesco Onelia, Monia Pacenti, Saverio Parisi, Giovanni Carretta, Daniele Donato, Luciano Flor, Silvia Cocchio, Giulia Masi, Alessandro Sperduti, Lorenzo Cattarino, Renato Salvador, Katy A.M. Gaythorpe, Alessandra R Brazzale, Stefano Toppo, Marta Trevisan, Vincenzo Baldo, Christl A. Donnelly, Neil M. Ferguson, Iliaria Dorigatti, and Andrea Crisanti**, “Suppression of COVID-19 outbreak in the municipality of Vo, Italy,” *medRxiv*, 2020. [2](#)
- Liggett, Thomas Milton**, *Interacting particle systems*, Vol. 276, Springer Science & Business Media, 2012. [6](#)
- Maloney, William F and Temel Taskin**, “Determinants of Social Distancing and Economic Activity during COVID-19: A Global View,” *Covid Economics*, May 2020, *Issue 13*. [4](#)
- Mangrum, Daniel and Paul Niekamp**, “College Student Contribution to Local COVID-19 Spread: Evidence from University Spring Break Timing,” *Available at SSRN 3606811*, 2020. [4](#)
- Mizumoto, Kenji, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell**, “Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020,” *Eurosurveillance*, Mar 2020, *25* (10). [2](#)
- Moll, Ben**, “Lockdowns in SIR models,” *Author’s website, Princeton*, 2020. [5](#)
- Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, and et al.**, “Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases,” *PLoS Medicine*, Mar 2008, *5* (3), e74. [6](#), [7](#)
- Neumeyer, Pablo Andres**, “Clase especial de epidemiologia,” *Author’s website, Class notes, Universidad Di Tella*, 2020. [5](#)
- Özgür, Onur, Alberto Bisin, and Yann Bramoullé**, “Dynamic linear economies with social interactions,” Melbourne Business School working paper 2019. [6](#)
- Paules, Catharine I, Hilary D Marston, and Anthony S Fauci**, “Coronavirus infections?more than just the common cold,” *Jama*, 2020, *323* (8), 707–708. [7](#)

- Pepe, Emanuele, Paolo Bajardi, Laetitia Gauvin, Filippo Privitera, Brennan Lake, Ciro Cattuto, and Michele Tizzoni**, “COVID-19 outbreak response: a first assessment of mobility changes in Italy following national lockdown,” *medRxiv*, 2020. 4
- Remuzzi, Andrea and Giuseppe Remuzzi**, “COVID-19 and Italy: what next?,” *The Lancet*, 2020. 7
- Sears, James, J Miguel Villas-Boas, Vasco Villas-Boas, and Sofia Berto Villas-Boas**, “Are we# stayinghome to Flatten the Curve?,” *Department of Agricultural and Resource Economics, first version April*, 2020, 5. 4
- Tomé, Tânia and Robert M Ziff**, “Critical behavior of the susceptible-infected-recovered model on a square lattice,” *Physical Review E*, 2010, 82 (5), 051921. 6
- Toxvaerd, FMO**, “Equilibrium social distancing,” Faculty of Economics, University of Cambridge 2020. 3
- Verelst, Frederik, Lander Willem, and Philippe Beutels**, “Behavioural change models for infectious disease transmission: a systematic review (2010–2015),” *Journal of The Royal Society Interface*, 2016, 13 (125), 20160820. 3
- Weitz, Joshua S, Sang Woo Park, Ceyhun Eksin, and Jonathan Dushoff**, “Moving Beyond a Peak Mentality: Plateaus, Shoulders, Oscillations and Other’Anomalous’ Behavior-Driven Shapes in COVID-19 Outbreaks,” *medRxiv*, 2020. 3
- Wilson, Daniel J**, “Weather, Mobility, and COVID-19: A Panel Local Projections Estimator for Understanding and Forecasting Infectious Disease Spread,” Federal Reserve Bank of San Francisco 2021. 4
- Wright, Austin L, Konstantin Sonin, Jesse Driscoll, and Jarnickae Wilson**, “Poverty and economic dislocation reduce compliance with covid-19 shelter-in-place protocols,” *Journal of Economic Behavior & Organization*, 2020, 180, 544–554. 4
- Wu, Chufen, Yong Yang, Qianyi Zhao, Yanling Tian, and Zhiting Xu**, “Epidemic waves of a spatial SIR model in combination with random dispersal and non-local dispersal,” *Applied Mathematics and Computation*, 2017, 313, 122–143. 6
- Yilmazkuday, Hakan**, “COVID-19 and unequal social distancing across demographic groups,” *Regional Science Policy & Practice*, 2020, 12 (6), 1235–1248. 4
- Zhang, Sheng, MengYuan Diao, Wenbo Yu, Lei Pei, Zhaofen Lin, and Dechang Chen**, “Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis,” *International Journal of Infectious Diseases*, 2020, 93, 201–204. 7



**Ziedan, Engy, Kosali I Simon, and Coady Wing**, “Effects of state COVID-19 closure policy on non-COVID-19 health care utilization,” Technical Report, National Bureau of Economic Research 2020. [4](#)