Learning Epidemiology by Doing: The Empirical Implications of a Spatial SIR Model with Behavioral Responses[§]

Alberto Bisin Andrea Moro

June 10, 2020

Abstract

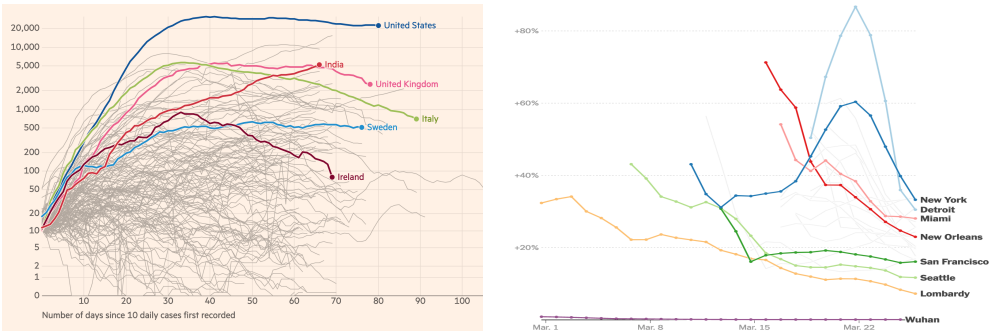
We simulate a spatial behavioral model of the diffusion of an infection to understand the role of geographical characteristics: the number and distribution of outbreaks, population size, density, and agents' movements. We show that several invariance properties of the SIR model with respect to these variables do not hold when agents are placed in a (two dimensional) geographical space. Indeed, local herd immunity plays a fundamental role in changing the dynamics of the infection. We also show that geographical factors affect how behavioral responses affect the epidemics. We derive relevant implications for the estimation of epidemiological models with panel data from several geographical units.

[§]Please check our websites for an updated version of this paper. Bisin: New York University, wp.nyu.edu/albertobisin/, alberto.bisin@nyu.edu. Moro: Vanderbilt University, andreamoro.net, andrea@andreamoro.net. We thank Pedro Sant'Anna and Giorgio Topa for their helpful comments on a preliminary draft of this paper; to Gianluca Violante for help with the calibration.

1 Introduction

Several media outlets report the diffusion of the SARS-CoV-2 epidemic by plotting case statistics over time by location as in Figure 1, sometimes to evaluate the relative effects of policy intervention. 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 epidemics obtained from the city of Vo', a small town near Padua, in Italy, or from the Diamond Princess cruiseship, to inform about the diffusion of the epidemics in New York city?[1] Indeed, SARS-CoV-2 has diffused at very different rates across countries and cities.[2]

Figure 1: Covid-19 disease trends as reported by media outlets



Left panel: Number of new reported cases in selected countries. Right panel: Average daily change in total cases in selected cities. Sources: (Left) Financial Times web site, their analysis of data from the European Centre for Disease Prevention and Control and the Covid Tracking Project. URL: https://ig.ft.com/coronavirus-chart/ (last retrieved: May 22, 2020); (Right) New York Times web site, their analysius of various sources. URL: https://www.nytimes.com/interactive/2020/03/27/upshot/coronavirus-new-york-comparison.html (last retrieved: May 22, 2020).

In this paper we propose a spatial model of epidemic diffusion, the Spatial-SIR model, to study how the dynamics of the epidemics scales in the number and distribution of outbreaks, population size, density, and agents' movements. We show how geography imposes restrictions on outcomes across locations, and how these restrictions cannot be uncovered from the workhorse model or epidemic diffusion, the SIR model.[3] These restrictions are consequential for empirical analysis using

[1] Lavezzo et al. (2020) and Mizumoto et al. (2020) provide a comprehensive review of the outbreak dynamics and steady-state outcomes for the City of Vo', and the Princess Cruise ship, respectively.

[2] See Fernandez-Villaverde and Jones (2020) and the dashboard produced by the authors, available at https://web.stanford.edu/ chadj/Covid/Dashboard.html.

[3] The SIR was developed by Kermack and McKendrick (1927), Kermack and McKendrick (1932).

time-series infection dynamics data.[4]

The spatial extension of the SIR model we study introduces important stylized spatial dimensions of the diffusion process, allowing us to ask a set of interesting questions that the basic SIR model cannot address. Our goal is to provide a better understanding of the core determinants of its dynamic properties and their order-of-magnitude effects.

We begin, in Section 2, by highlighting the conditions in which the standard SIR model is invariant to several factors we focus on in this paper: the number of outbreaks, population size, and density.

In Section 3 we introduce the Spatial-SIR model. In Spatial-SIR, 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 certain probability that we interpret as the strength of the virus. Spatial-SIR determines the diffusion rate of infection depending on epidemiological and geographic factor that are confounded in one single parameter of the standard SIR instead. In Section 3.2 we show how distinguishing these factors is crucial in Spatial-SIR because of what we call "local herd immunities", that are generated by the constrained movement of people in space. In the SIR model, instead, susceptible individuals match with infected individuals randomly. Local herd immunities are responsible for breaking several of the invariance relationships which hold in the SIR model (that we highlight in Section 2).

In Section 4 we calibrate the parameters of the Spatial-SIR model and use simulations to study the roles of the number and distribution of outbreaks, population size, density, and agents' movements on epidemic outcomes. We highlight the quantitatively important effects of these geographic factors in determining infection dynamics. These effects are missed in the standard SIR.

The infection diffusion rate of the SIR model also does not account for behavioral responses of economic agents to the diffusion of the epidemics. Contributions highlighting this important factor attempt to account for these behavioral responses by scaling down the infection diffusion parameter according to a calibrated parameterization that depends on the dynamics of the epidemics.[5] In Section 5, we incorporate

Research in epidemiology has extended this model in many directions, allowing for geographic or network dimensions to the diffusion process; see e.g., Eubank et al. (2004) and the research at https://covid19.gleamproject.org/, https://www.mobs-lab.org/projects.html. The resulting models exploit very detailed descriptions of the demographic characteristics of the population of interest and of the social and geographical environment in which the population lives. These models appear to fundamentally aim at forecasting with accuracy and precision as, say, meteorological models of weather dynamics rather than at identifying the stylized effects of geographical characteristics.

[4]The recent wealth of contributions to the study of the SARS-CoV-2 epidemic in economics has basically restricted its epidemiology component to various extensions of the SIR model that do not account for the geographic characteristics that we focus on in this paper; see e.g., Fernandez-Villaverde and Jones (2020), Atkeson (2020), Eichenbaum et al. (2020), Keppo et al. (2020), Weitz et al. (2020), Brotherhood et al. (2020), Jarosch et al. (2020).

[5]Early contributions in this respect include Goenka and Liu (2012), Geoffard and Philipson

behavioral responses into the model to highlight how their effects depend on geographic factors.

In Section 6 we focus on five implications for empirical analysis we learn from our model and its simulations. In particular, we note that research exploiting geographic variation to study the effect of policy intervention, or to study how epidemic outcomes depend on covariates using longitudinal data, can gain from imposing the cross-location restrictions implied by the epidemiological models and at the same time must deal with time-varying heterogeneity across locations that is hard to control for without imposing specific structure.

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. Let $h_t = [S_t, I_t, R_t]$ denote the distribution of the population across the state-space at time t. The dynamics of h_t is governed by the following transitions: i) a Susceptible agent becomes infected upon contact with an infected, with probability $\beta \frac{I_t}{N}$; ii) an agent infected at t, can recover at any future period with probability ρ; iii) a Recovered agent never leaves this state (this assumes that Recoved agents are immune to infection).

The SIR can be solved analytically.[6] The equations describing its dynamics in discrete time are

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

The parameter β in Equation 1 is to be interpreted as the infection rate in the model. It is related to $\mathcal{R}_0 = \beta/\rho$, which represents the number of agents a single infected agent infects, on average, at an initial condition $R_0 = 0,\ I_0 \to 0$. The infection rate β can be decomposed in terms of the infection rate per-contact between a susceptible and an infected, say π, and the number of random matching contacts per unit of time in the population, say c.[7] A susceptible agent meets with probability c with any other agent; hence with probability $c\frac{I_t}{N}$ with an infected agent.

(1996). Recent work includes Farboodi et al. (2020), Keppo et al. (2020), Greenwood et al. (2019), Bethune and Korinek (2020). In the epidemiology literature, behavioral response seems to have broken somehow into the theoretical literature but much less into the applied literature: see Verelst et al. (2016) for a comprehensive survey.

[6] See e.g., Hethcote (2000) for the analytical solution; see also Atkeson (2020), Moll (2020), Neumeyer (2020).

[7] Distinguishing the role of the number of contacts from the role of the contagion rate is conceptually important because $\mathcal{R}_0$ and β are often interpreted as structural parameters of the model. In our spatial SIR model, they are the product of virological, geographical and, in Section 5, behavioral factors.

We highlight three invariance properties of the dynamics of the SIR model, whose robustness to the introduction of a spatial structure we shall evaluate in the rest of the paper.

Stationary state invariance to initial conditions. Given any initial conditions, $R_0 = 0, I_0 > 0, S_0 = N - I_0$, the dynamical system converges to a unique stationary state. In other words, the inital number of outbreaks of the infection, I_0, has no effect on the stationary state. This stationary state with $R^* > 0$ is characterized uniquely in terms of $\mathcal{R}_0 = \beta/\rho$, as the solution of the following fixed point equation:

$$R^* = -\frac{1}{\mathcal{R}_0} \ln(1 - R^*). \tag{2}$$

Transitional dynamics invariance to population size (in the limit $\frac{I_0}{N} \to 0$). The dynamics of $\frac{1}{N}(S_t, I_t, R_t)$ is invariant to population size, N, as $R_0 = 0$ and the fraction of the population infected at the initial condition converges to zero, $\frac{I_0}{N} \to 0$. The peak of infected cases is

$$\frac{I}{N} = 1 - \frac{1}{\mathcal{R}} (1 + \log \mathcal{R}) . \tag{3}$$

Transitional dynamics invariance to contacts and contagion keeping β constant. The dynamics of $\frac{1}{N}(S_t, I_t, R_t)$ is invariant to changes in the number of contacts c and probability of contagion, π that leave $\beta = \pi c$ constant.

If the epidemics is governed by the SIR model all of these invariances provide restrictions of the model which are testable with cross-city data (see Section 6).

3 The Spatial SIR model

We now add a spatial dimension to the SIR model. We also expand the state space to better capture several relevant aspects of the SARS-CoV-2 epidemic.[8] Specifically, we split the I state into Asymptomatics and sYmptomatics, A and Y. We also add explicitly the state D, for Dead. Hence, $\mathcal{S} = \{S, A, Y, R, D\}$. We maintain the notation $h_t^i \in \mathcal{S}$ to denote the state of agent i at time t; and $h_t = [S_t, A_t, Y_t, R_t, D_t]$ to denote the distribution the N agents in the population across the state-space.

[8] This expansion of the state space is inconsequential for the study of the effects of geographical characteristics of cities but its adds realism, thereby helping the study of e.g., policy implications; see Bisin and Moro (2020) for an application.

3.1 The Model

Agents are located in space, e.g., a lattice, which we call "the City." Agents are ex-ante identical in terms of demographic characteristics and symmetric in terms of location in space. Two agents come into contact when they are at a geographical distance in space closer than p. Agents move randomly in space: Every day $t = [0, T]$, agents travel distance μ toward a random direction of $d \sim U[0, 2\pi]$ radians.[9]

Spatial-SIR is represented by the following transitions: i) a Susceptible agent in a location within distance p from the location of an Asymptomatic becomes infected with probability π;[10] ii) an Asymptomatic agent infected at t, at any future period, can become sYmptomatic with probability ν, or can Recover with probability ρ; iii) an agent who has become sYmptomatic at t, at any future period, can Recover with probability ρ, or can Die with probability δ; iv) Dead and Recovered agents never leave these states (this assumes Recoved agents are immune to infection).

The resulting dynamical system is difficult to characterize formally.[11] We turn then to simulations. We calibrate transitions away and between the infected states, A, Y, D, R to various SARS-CoV-2 parameters from epidemiological studies, notably e.g. Ferguson et al. (2020). We calibrate $\beta = \pi c$ and the agents' daily travel distance μ from estimates of initial growth rates of the epidemics (in Lombardy, Italy) and data on average contacts in Mossong et al. (2008).[12]

Figure 2 illustrates the dynamics of the epidemic in space at the calibrated parameters. The epidemic spreads exponentially from the location of the outbreak.[13]

3.2 Local Herd Immunity

To understand how Spatial-SIR differs from the standard SIR, we simulate the evolution over time of the growth rates of the infection and the number of active cases (that is, infected agents, I/N) for SIR and Spatial SIR. We argue that the fundamental differences can be rationalized in terms of the effects of *matching frictions* (implicitly defined by geography and people's movements), inducing a form of *local herd immunity* which characterizes Spatial-SIR. More specifically, in Figure 3 we

[9] When they get close to the boundary, the direction is randomly drawn but constrained to point opposite to the boundary.

[10] Susceptible agents are not infected upon contact with a sYmptomatic agent; this is to capture the fact that sYmptomatic agents are either isolated at home or in the hospital

[11] In the Appendix A we show that it can be written as a Markov chain on configurations in space, along the lines of interacting particle system models (Liggett, 2012; Kindermann and Snell, 1980). 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), Glaeser and Scheinkman (2001), Conley and Topa (2007), Özgür et al. (2019).

[12] See Appendix B for the details on the calibration.

[13] All our simulations, for all parameter values and initial conditions, converge to a unique distribution over the state space $[S, A, Y, R, D]$.

Figure 2: Geographic progression of infections and recoveries

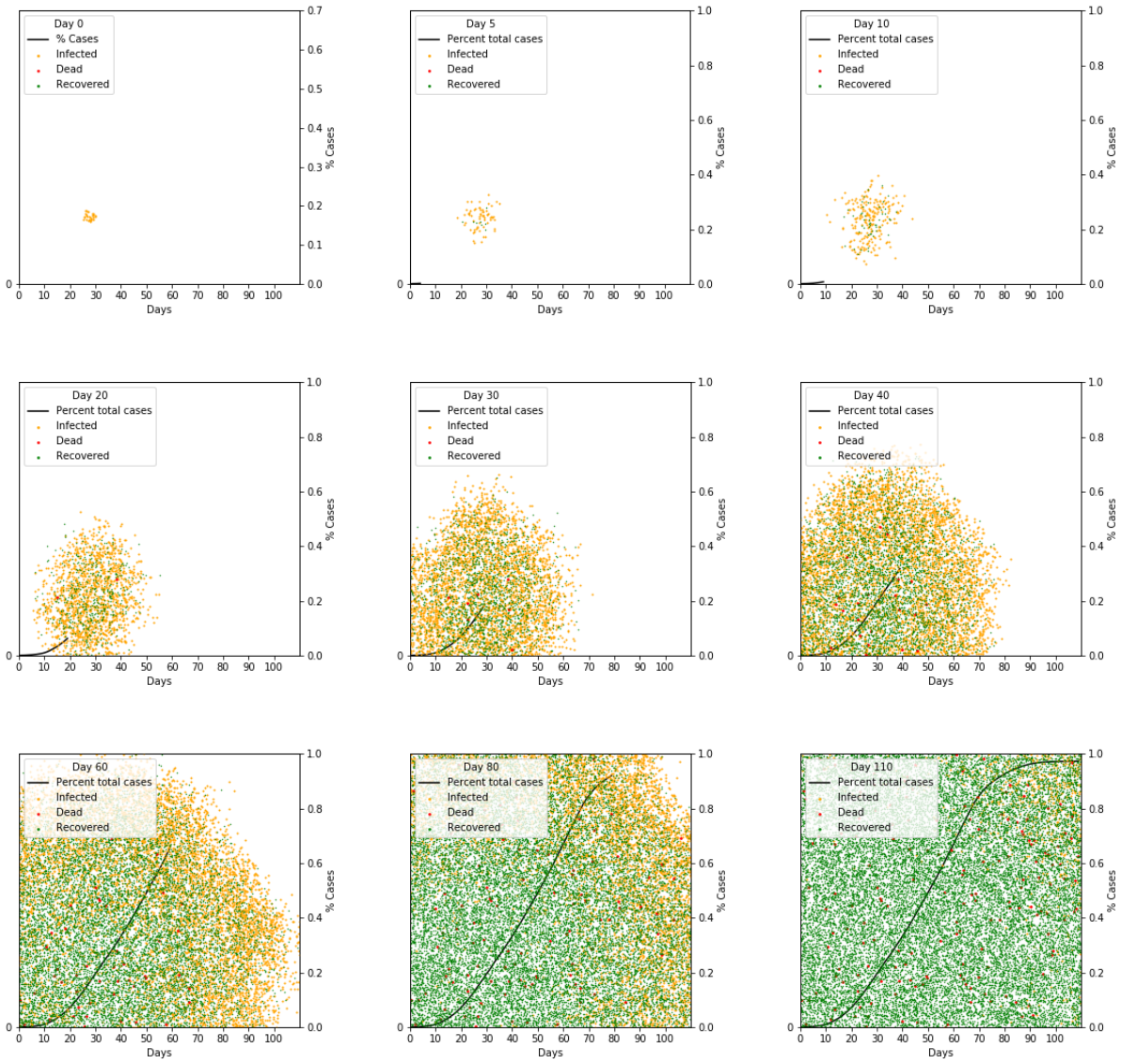
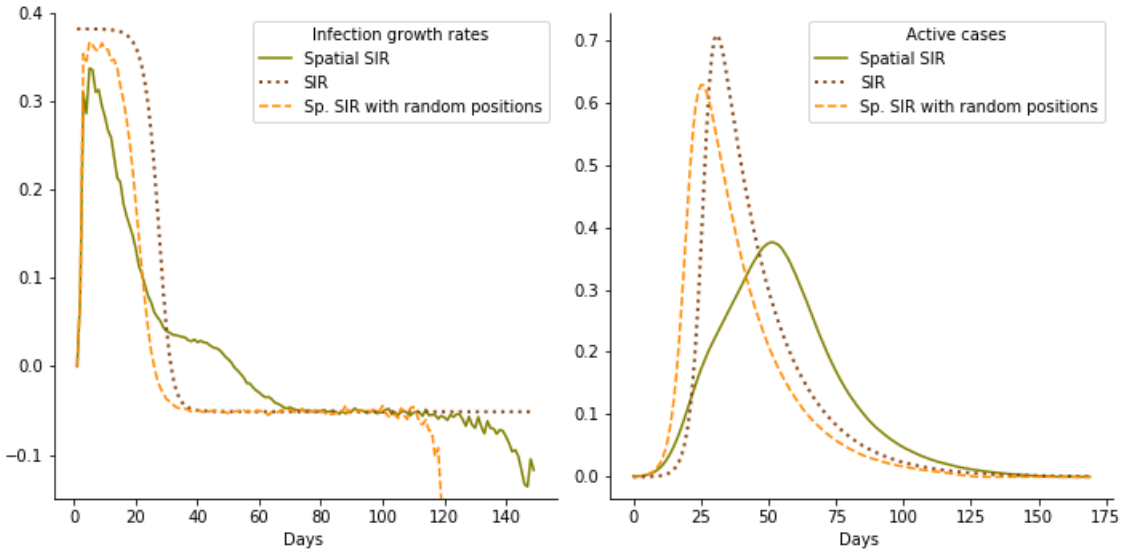


Figure 3: Comparison between SIR and spatial models



compare simulations for three different models:

i. (continuous line) a Spatial SIR (simplified to three states, $(S, Y,$ and $R)$ for comparison), with the same parameters we calibrated for our benchmark model;

ii. (dashed line) a Spatial SIR as in (i), but with agents placed in a random position every day in the city.

iii. (dotted line) a SIR model, with β 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; we set $\rho = 0.05$, as in our benchmark model.

The spatial models (i-ii) all display initially lower growth rates than the SIR model (iii). This is because the agent's movement in space generates "local herd immunities," slowing down the diffusion of infection in the early stages and accelerating it afterwards (as aggregate herd immunity is delayed). Formally, in SIR, random matching implies that the probability that any Susceptible agent is infected at time t is $\beta \frac{I_t}{N}$. In Spatial-SIR this probability is a random variable, say $\beta \lambda_t(\frac{I_t}{N})$, and its expectation across all agents $\beta E\left[\lambda_t(\frac{I_t}{N})\right]$ encodes the effects of local herd immunity over time, as $E\left[\lambda_t(\frac{I_t}{N})\right]$ is initially smaller and then larger than I_t/N.

The effect of local herd immunity is much stronger in model (i) (continuous line) than (ii) (dashed line). In fact, in model (ii) agents are set to a random position every day, mimicking "random matching" as in the SIR model (iii) and therefore minimizing the formation of local herd immunities.

These results indicate that the "reduced form" nature of SIR models is missing a potentially important role of matching frictions and, more generally, of local dynamics. Similar condiderations can be obtained looking at $\mathcal{R}_0$, which is a random variable in Spatial-SIR, as the number of contacts of an individual is random.

Replicating simulations of our baseline Spatial SIR, we estimate it as the average number of people infected by the individuals who contracted the infection during the first 5 days. We find that this estimate of the $\mathcal{R}_0$, within the range used to calibrate transition rates in many studies (between 2.5 and 3.5), is highly volatile. In 20 random replications of the model, the average $\mathcal{R}_0$ is 2.66, with a standard deviation of 0.48. However, in Spatial-SIR this volatility does not translate into similarly different aggregate outcomes as predicted by standard SIR. The total fraction of cases in steady state averages to 0.97 in the 20 replications, with a standard deviation of 0.001. This suggest that, in our model, $\mathcal{R}_0$ loses its role as the fundamental driving parameter of the epidemics, since outcomes are also highly sensitive to individual characteristics of initial cluster of infection. While the infection rate in the very first days of the infection is uniquely determined by the structural parameters $\mathcal{R}_0$ and ρ, which are (relatively) independent of the spatial structure of the model, the dynamics of the infection rests on the spatial local interaction structure. In other words, the growth rate of the infection might decline early on in the epidemic following a form of local herd immunity. Indeed, this is what we observe in the data and we set parameters to match.

4 Oubreaks, Spatial SIR: Size, Density, and Movements

In this section we simulate Spatial-SIR to highlight the role of geographical characteristics in the determination of the matching frictions and local herd immunity we have identified in the previous section. We will study the role of outbreaks, population size, density, and agents' movement. More precisely, outbreaks are defined as the number of infected agents at the initial condition, I_0. In Spatial-SIR, the specification of this initial condition includes the distribution of outbreaks over the City. Population size is N. The density of a City is $d = \text{area}/N$. In Spatial-SIR, density is related to the number of contacts c by $c = d\Psi$, where Ψ is the contagion area of any (susceptible) individual.[14] Finally, agents' movement is μ, the average distance travelled each day in Spatial-SIR. We denote $g = [I_0, N, d, \mu]$ the vector of the geographical characteristics we study in Spatial-SIR.

[14]Parameter Ψ will be maintained constant in the whole paper.

We study both properties of the dynamics at the stationary state (the fraction of Recovered and Dead) as well as properties of the transitional dynamics (the time it takes to for an outbreak to reach the peak of active cases, a measure of the speed of the epidemic, and the height of the the peak of active cases, a measure of the intensity of the epidemic).

We will show that these outcomes do not satisfy some of the invariance properties of the SIR dynamics we have delineated in Section 2. Fundamentally, the correction of the SIR dynamics due to local herd immunity is a function of geographic characteristics g. We write this correction then as $E\left[\lambda_t(\frac{I_t}{N};g)\right] - \frac{I_t}{N}$, This analysis of the effects of various geographical on the spread of the epidemics has some clearcut implications regarding how to interpret the scale of the model in simulations. Most importantly, Spatial-SIR is not dimensionless. City size, density, the number and distribution of outbreaks, and movements in the city are variables that empirical cross-city studies of epidemic dynamics should account for.

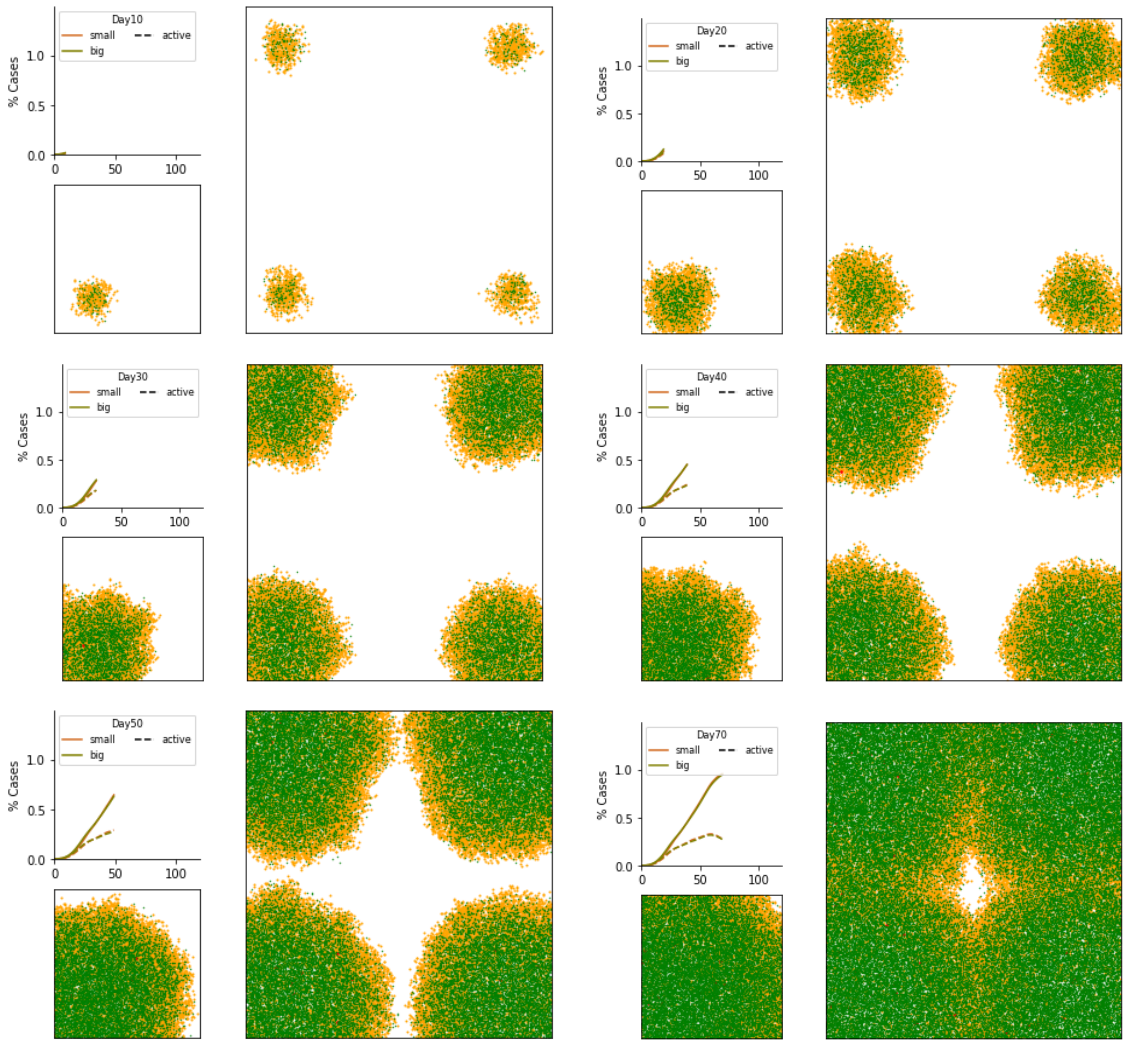
4.1 Outbreaks

The dynamics of the epidemic in Spatial-SIR *is invariant to the initial condition*: scaling in size obtains if we scale initial conditions, that is, if in a $x-$times larger city with $x-$times population, epidemics breaks out from $x-$ times as many locations. This is the case in SIR as well as in Spatial SIR. *In Spatial SIR, however, this is the case only if the initial conditions are appropriately homogeneously spaced.* This point is illustrated in Figure 4 where we can compare the progression of the contagion at days 10, 20, 30, 40, 50, and 70, 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. Each panel reports on the right the geographical location of infections in the bigger city, on the bottom left the geographical location of infections in the baseline (smaller) city, and on the top left the contagion rates.

The progression of the infection is almost entirely symmetric, barring minor effects due to the randomness of people's locations and movement. The top-right chart in each panel shows that both the fraction of active and total cases is nearly identical between the two Cities.

To better understand the role of the distribution of outbreaks in Spatial SIR, in Figure 5 we illustrate the progression of the contagion on days 0, 10, and 25 comparing the baseline City (top 3 panes) with one single initial cluster of infected and an identical City in which however the initial cluster of infected is split and the infected agents are randomly located (bottom 3 panes). While in the baseline model, contagion is relatively concentrated by day 25, contagion is widely spread by the same date if the initial contagions are randomly located.

Figure 4: Rescaling a City



From the top-left panel proceeding right and down: day 10, 20, 30, 50, 70. Yellow dots are the the active cases, green dots are the recovered cases (susceptibles are omitted). Area of small (large) city: 1 (4). Initially infected at $t = 0$: 30 (120).

Figure 5: Progression of contagion: one inital cluster (top 3 panes) vs random initial contagion (bottom 3 panes)

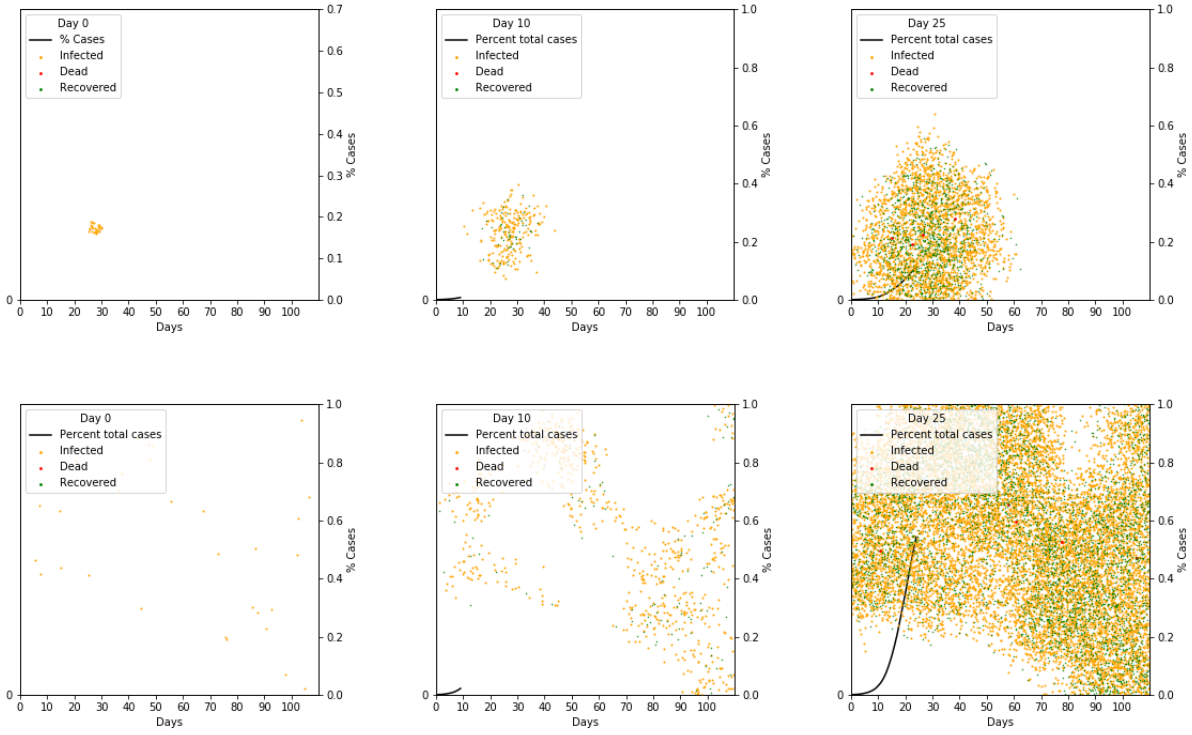
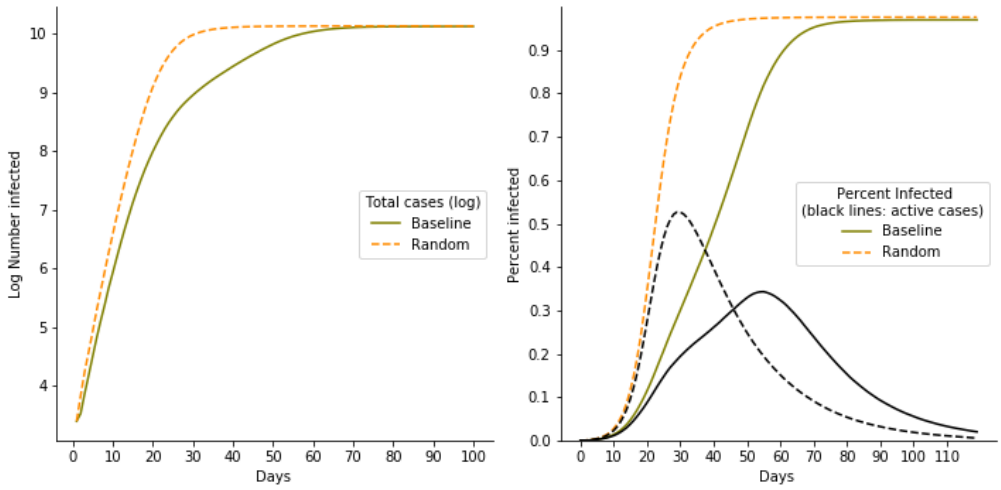


Figure 6 summarizes the infection dynamics in these two simulations: the progression of active cases, $I/N = (A + Y)/N$, is faster when the initial cluster is randomly, reaching a higher peak of active cases (51% rather than 27%) earlier (on day 30 rather than on day 65). However, the fraction of Recovered and Dead at the stationary state, $(R^* + D^*)/N$, is the same (97%)).

Figure 6: One initial clusters vs. random initial locations



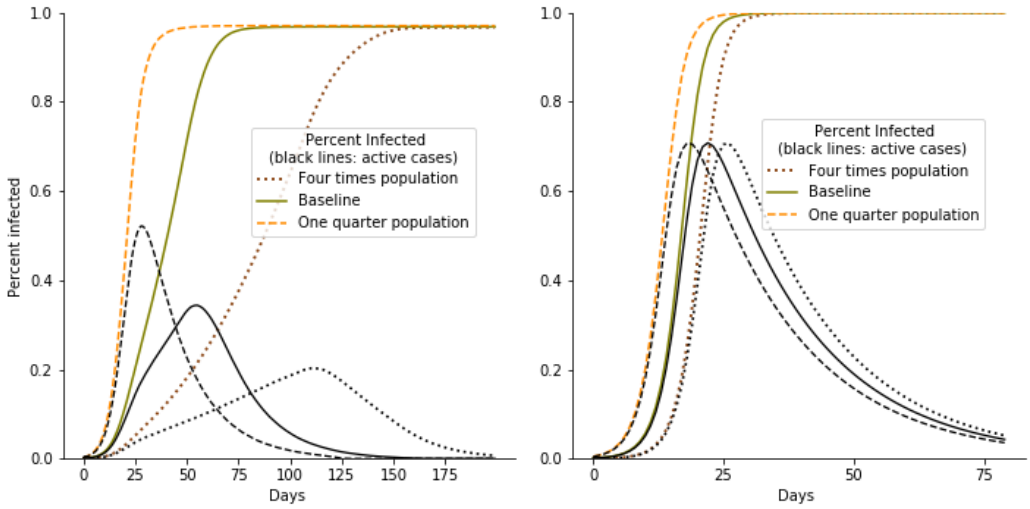
4.2 Population size

In this section we study the effects of changing population size N and city area proportionally so as to keep the City density constant, while fixing the size of the initial outbreak of the infection, I_0. We have shown in Section 2 that in the SIR model these changes have no effect on the stationary state nor on the transitional dynamics (in the limit as I_0/N is converging to zero).

We compare these effects betweeen the SIR and Spatial-SIR in Figure 7 where we report infections as percent of the population; we illustrate for both models three cities: the baseline, a city of size 1/4th of the baseline and one 4 times the baseline. *Changing population size does not change the stationary state fraction of infected*, in both models $(R^* + D^*)/N$ is approximately equal 97 percent of the population. This is consistent with the stationary state invariance property of SIR.

However, the transitional dynamics of the epidemic are not invariant to city size in Spatial-SIR (left pane): the curve displaying the fraction of active cases, I/N, is flatter in larger cities. This differentiates the dynamics between Spatial-SIR and the SIR models. In fact, the transitional dynamics of SIR are not invariant to population size (only in the limit for $I_0/N \to 0$ they are). But their dependence on size is minimal: (it can be shown formally that) an increasing x-time population size increases the peak by $-1/\mathcal{R} \ln x$ percenteage points. With our parameters - the difference is hardly visible (right panel). In Spatial-SIR instead, the same difference in population size reduce the peak in more than half (from .52 to .2 active cases) (left panel). Furthermore, the time to get to the peak is longer in larger cities. But while it goees from 18 to 26 days in SIR, it goes from 28 to 111 in Spatial SIR.

Figure 7: City size comparisons in Spatial SIR and standard SIR

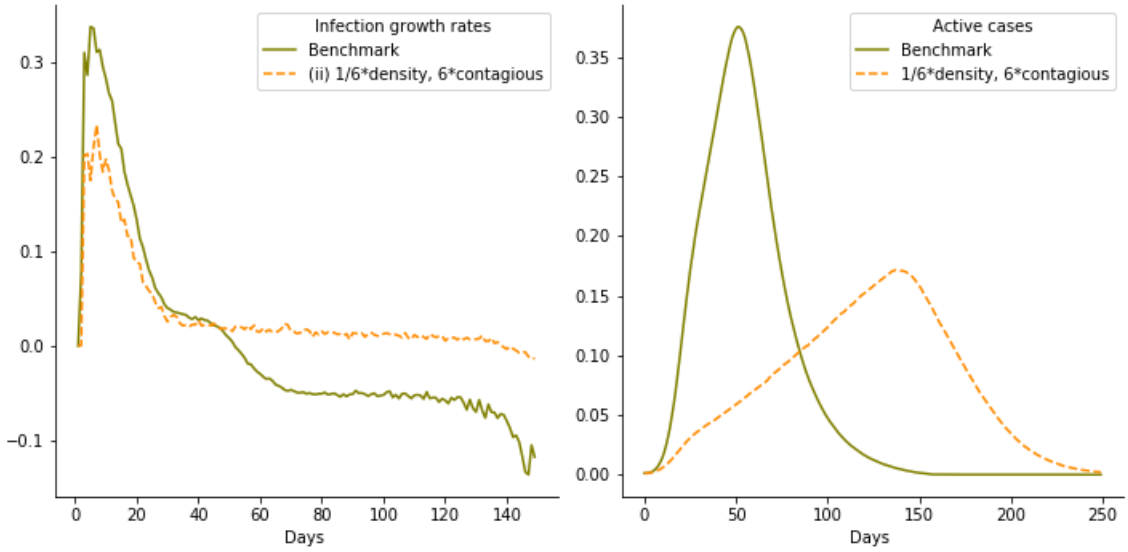


Left pane: Spatial SIR model, Right pane: Standard SIR model. Colored lines: percent ever infected; black lines: current active cases.

4.3 City Density

In this section we study the role of City density on the dynamics of the epidemic. We show that *City density in the Spatial SIR model plays a distinct role from the inverse of the probability of infection, breaking the invariance* we have highlighted in Section 2 *in the SIR model.* This is very clearly shown in Figure 8 where the baseline calibrated Spatial SIR is compared with an environment with 6 times the probability of infection and 1/6th the density: the effect on the infection growth and dynamic is different both qualitatively and quantitatively.

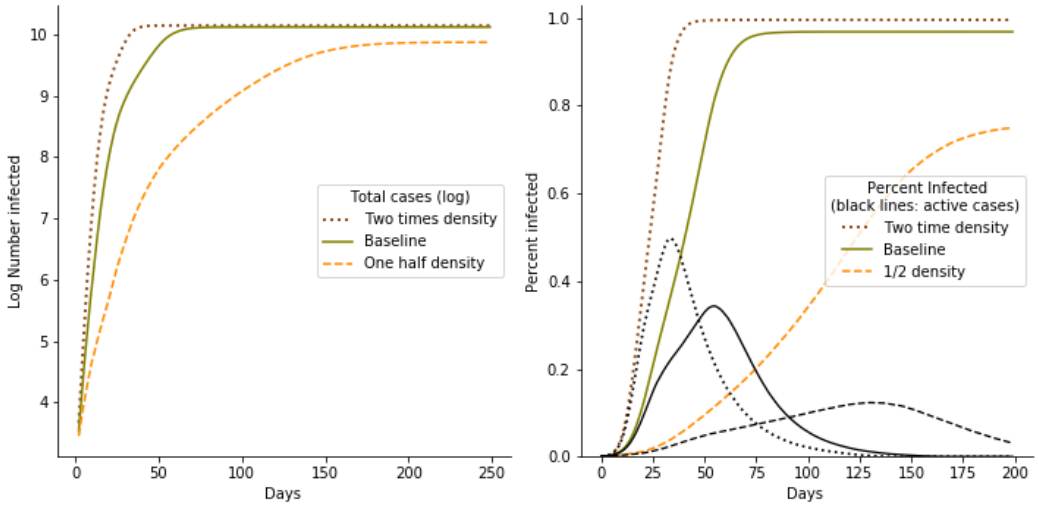
Figure 8: Changing density and contagion rate keeping the probability of infection constant



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 epidemic, as illustrated in Figure 9.

To explore in detail the relationship between density and transitional dynamics of the epidemic, in the right panel of Figure 9 we see that indeed $(R^* + D^*)/N$ is increasing in density. Most importantly, the peak of active cases $\frac{I}{N}$ is very sentitive to density: halving density with respect to the baseline has the effect of dramatically flattening the peak of the infection (more than a half, after more than twice as many days from the outbreak). Density is a crucial determinant of the dynamics of the epidemic because, together with the contagion rates, 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.

Figure 9: The effect of varying city density with constant population



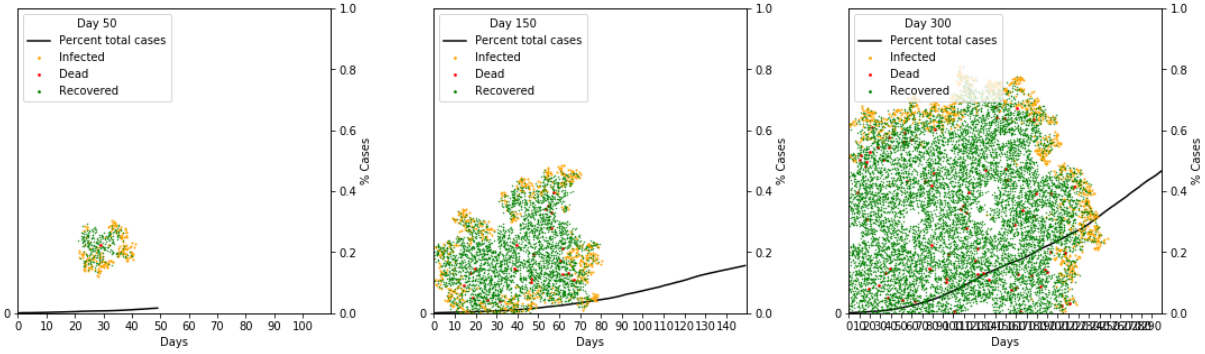
4.4 Movements in the City

In Spatial-SIR several new parameters could contribute to explaining the cross-city heterogeneity in the dynamics of the epidemic. In this section we study variation in the random movement across space. The parameter controlling these movements is the distance traveled every day by each agent, μ.[15] Changing this parameter 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. To provide an intuition of the dependence of the epidemic on the movement speed of agents in the city, Figure 10 reports an extreme case: the progression of contagion over space and the speed and intensity of the spread when agents do not move.

The infection spreads slowly. As 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 affects also the steady state as illustrated in Figure 11.

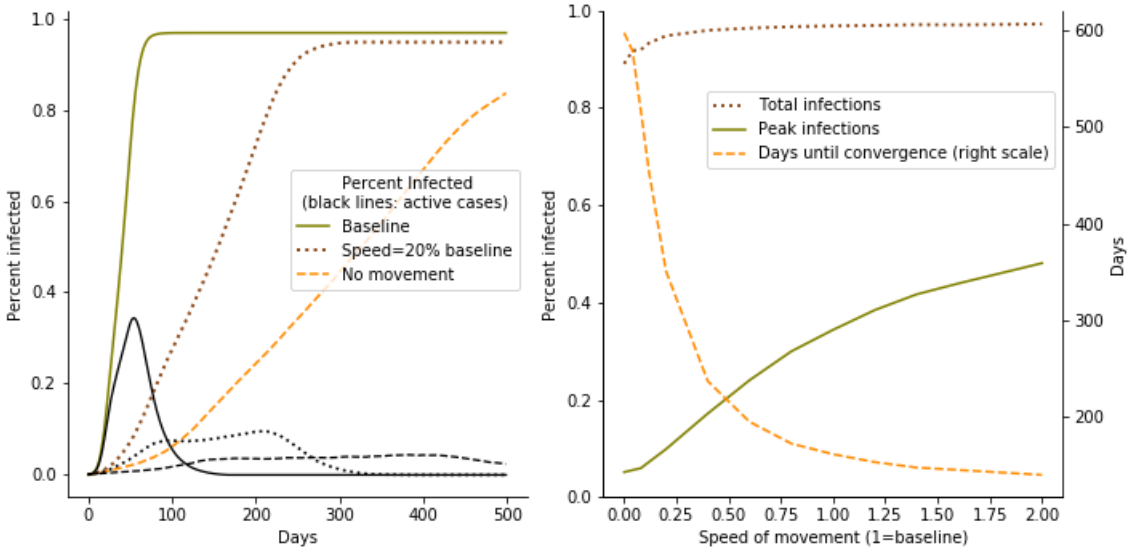
[15] Given our calibration of the spatial structure of the city with respect to the contagion speed (namely, the contagion radius), if all people were placed on an equally spaced grid, contagion would not occur. All infections in the baseline model occur initially because random placement generates clusters of people closer to one another than the infection radious. Contagion expands over time because people randomly move daily around the city.

Figure 10: Progression of contagion when agents do not move: days 50, 150, 300



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 less likely for people in these locations to stay immune until the steady state.

Figure 11: Movement speed



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.

5 Behavioral Spatial SIR

In most predictive epidemiological studies, the analysis disregards behavioral responses to the epidemic; see e.g., Ferguson et al. (2020). In this case, as in the previous sections, the number of daily contacts in the population, c, is a constant.[16] In economic models, on the other hand, agents' choices are generally modeled as the outcome of rational decision making on the part of the agents themselves. Agents react to the dynamics of the epidemic, by choosing to limit their social interactions, their contacts.

In the recent economics literature on SIR, agents' choices are modelled as reduced form behavioral responses, postulating an endogenous dependence of agents' contacts from the number of infected in the population. the behavioral response be represented by a function $0 \leq \alpha(I_t) \leq 1$, acting as a proportional reduction of the agent's contacts:

$$c = \alpha(I_t) d\Psi.^{17}$$

Keppo et al. (2020)e.g., adopt a behavioral response with the functional form:[18]

$$\alpha(I_t) = \begin{cases} 1 & if \;\; I_t \leq \underline{I} \\ \left(\frac{\underline{I}}{I_t}\right)^{1-\phi} & if \;\; I_t > \underline{I} \end{cases} . \tag{4}$$

In this section we extend Spatial-SIR assuming that a fraction of the population self-isolates, thereby reducing their contacts to zero, according to behavioral response (4). We calibrate the dynamics of the epidemics allowing for behavioral response, in both SIR and Spatial SIR.[19] In the calibration, the percent reduction in the number of contacts according to the behavioral response function is reported in Figure 12. As the infection spreads, the number of contacts decreases. As herd immunity begins and the number of infected declines, contacts increase towards the initial (pre-infection) state.

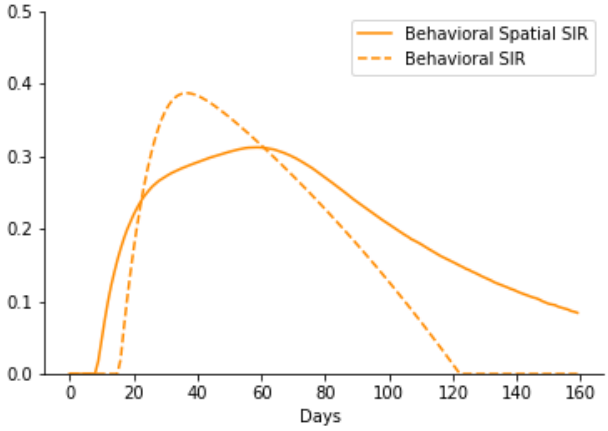
[16] But see Verelst et al. (2016) for a systematic survey of theoretical studies of behavioral respenses in epidemiology.

[17] See Weitz et al. (2020) and Farboodi et al. (2020), for a similar reduced form approach. A fully dynamic rational choice micro-foundation for this approach in SIR is obtained in Appendix B. In Spatial-SIR with state space (S, A, Y, R, D), the behavioral response will depend on A. Since the fraction of asymptomatics is not observable, behavioral response could only depend on the number of symptomatics, as a proxy; with Rational Expectations, however, the agents know (rationally infer) the equilibrium map from Y to A, say $A(Y)$, possibly with noise; see Bisin and Moro (2020).

[18] In Fernandez-Villaverde and Jones (2020) the behavioral response of agents is instead assumed time-dipendent, $\alpha(I_t) = \alpha_0 e^{-\lambda t} + \alpha^*(1 - e^{-\lambda t})$.

[19] Details about the calibration are reported in Appendix C. We calibrated the SIR model as in the simulations in Section 3.2 for this comparison.

Figure 12: Reduction in contacts according to the behavioral response



In Figure 13 we simulate the effects of behavioral responses on the dynamics. 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, but then slowing down the operation of herd immunity. As the number of contacts goes back to normal, the behavioral response has no effects in stationary state. 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 characteristcs g as long and these affect contacts; see 1 in B. We write the behavioral response then as $\alpha(I_t; g)$.

Figure 13: Comparison of SIR with Behavioral model

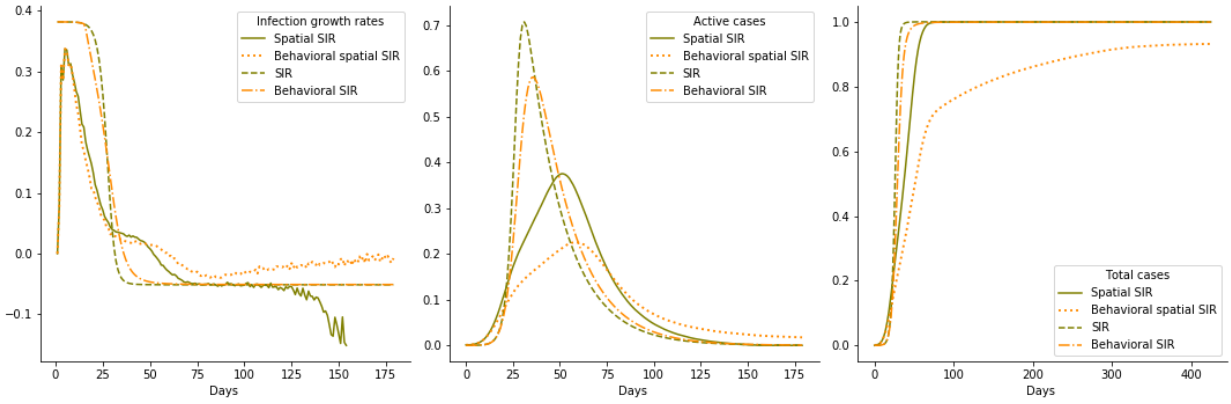


Figure 13 highlights also the differential effects of behavioral responses on SIR and Spatial SIR. The behavioral response is not only much stronger in Spatial-SIR, but qualitatively different when comparing both infection growth rates and the fraction of active cases. The peak of active cases in Spatial SIR is a third, 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 spatial correction on SIR, $E\left[\lambda_t(\frac{I_t}{N}; g)\right] - \frac{I_t}{N}$. The first acts on the number of contacts, while the second acts on the distribution of infected between the contacts.

6 Implications for Empirical Analysis

We summarize five implications of our analysis to guide empirical research using panel data about the diffusion of an epidemic. We discuss both structural estimates of a formal epidemic model, and 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 Difference (DiD) design.

Consider panel data on the dynamics of an infection over time t across different geographic units (Cities) i. The econometrician observes the geographic characteristics $g_i = [I_0, N_i, d_i, \mu_i]$ of each City i for several times t, as well as data on the dynamics of the infection, $I_{i,t}, R_{i,t}$ (hence $S_{i,t}$).[20] We make several points which we are expanding in research we are currently pursuing.

1. Cross-City restrictions in the standard SIR. To highlight how model restrictions could be exploited for empirical analysis, consider first estimating a SIR model without behavioral effects, as in standard epidemiological studies (see Ferguson et al. (2020) for example). Consider the following specification:

$$\ln I_{i,t+1} - \ln I_{i,t} = \beta_{i,t} S_{i,t} \frac{I_{i,t}}{N_i} - \rho \tag{5}$$

$$\text{with } \beta_{i,t} = \pi c_i, \quad c_i = d_i \Psi \tag{6}$$

Equation (6) imposes important (falsifiable) cross-City restrictions; e.g., the growth rate of the fraction of infected in a City, other things equal, is proportional to the density of the City. This can be tested.

2. Cross-City restrictions in Spatial-SIR. Accounting for a spatial structure on the SIR model introduces matching frictions through local social interaction, as we have shown in Section 4. The specification of the dynamics of the infection in

[20] Possibly, in fact, distinguishing $A_{i,t}$ and $Y_{i,t}$ as well as $R_{i,t}$ and $D_{i,t}$.

(5) takes the form

$$\ln I_{i,t+1} - \ln I_{i,t} = \beta_{i,t} S_{i,t} E\left[\lambda_t\left(\frac{I_{i,t}}{N_i}; g_i\right)\right] - \rho. \tag{7}$$

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 function λ, making it it difficult to separately identify the effects of geography from infection strength. However, one can use the full structure of the model to match data with model predictions using simulation methods. Alternatively, one could use simulations to estimate $E\left[\lambda_t\left(\frac{I_{i,t}}{N_i}; g_i\right)\right]$ which can be used as a correction to the dynamics of the SIR model (which is much faster to simulate), to estimate (6-7).

3. Identifying behavioral responses

Accounting for agents' choices, the number of contacts is endogenous and (6) takes the form

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

This amplifies the issues we highlighted so far, requiring a new identiquiring a new identification strategy. The standard SIR 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 π from $\alpha(I_t; g_i)$. Parametric identifcation can be achieved by assuming a functional form for $\alpha(I_t; g_i)$ along the lines of (4) from Keppo et al. (2020) In Spatial-SIR the full specification is (7-8). Identification in this case can rely on simulation methods as suggested at the end of empirical implication 2.[21]

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.[22] Policy and agent behavior have separate effect on the dynamics of the epidemic both because behavioral responses have time-varying effects, as we uncovered, and because their effects interact with the effects of geographi (a point generally disregarded in the few studies that try to account for behavioral responses). However, to identify behavioral responses one could focus on pre-treatment data. 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 by exploiting restrictions imposed by Spatial-SIR.

[21] Fernandez-Villaverde and Jones (2020) adopt simulation methods to estimate parameters separately for each location without imposing geographic restrictions

[22] See Bisin and Moro (2020) for more detailed considerations on this issue.

4. Identifying the time-varying effect of geography in DiD studies of NPIs Reduced-form methods can also be exploited to separately identify the effects of policies and agents' behavioral responses. Consider a treatment, like e.g., an NPI, introduced at different times in different cities[23]. Let $\text{Treat}_{i,t}$ take value 1 if city i is treated at time t. Consider the following 2-way fixed-effects DiD specification:

$$\ln I_{i,t+1} - \ln I_{i,t} = \nu + \eta_i + \gamma_t + \delta \text{Treat}_{i,t} + \lambda X_{i,t} \tag{9}$$

where ν, η_i, γ_t are time and location effects and $X_{i,t}$ are additional controls. Our analysis of Spatial-SIR implies that the vector of geographic factors g_i affects outcomes differently over time, therefore the inclusion of location and time fixed effects may not fully account for the bias arising from the time-varying heterogeneity introduced by $\lambda(\cdot)$ and $\alpha(\cdot)$ defined in empirical implications (2) and (3). The inclusion of geographic factors such as density as controls, even interacted with time, may not be sufficient both because their effect are non linear, and because it is often hard to pin down the beginning of the infection in all localities.[24]

A similar specification is used in the literature to study the effects of the treatment to the growth rate in number of contacts $\ln c_{i,t+1} - \ln c_{i,t}$ rather than on the growth rate of cases $\ln I_{i,t+1} - \ln I_{i,t}$.[25] But if c_i depends on I, then the effect of g_i is not captured by the city and time fixed effects η_i, γ_t.

5. Geographic units of analysis and their characteristics. It is important to choose geographic units of analysis so that density and other geographic characteristics g_i are relatively homogeneous. For this reason, empirical analyses with data across countries involve additional concerns with respect to data across cities.

In Section 4 we found that, besides population size and density, the distribution of outbreaks and the speed of movement of the agents have systematic effects on the dynamics of an epidemic. Proxies like the 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 used in both reduced-form and structural estimates.

We also note that in structural estimates, heterogeneous density and various distribution of outbreaks can be easily included in the estimation of a Spatial SIR (but not in an estimation of the SIR).

[23] See e.g., Allcott et al. (2020), Courtemanche et al. (2020), Fang et al. (2020), Hsiang et al. (2020) Maloney and Taskin (2020), Mangrum and Niekamp (2020), Pepe et al. (2020)

[24] See Goodman-Bacon and Marcus (2020) for a comprehensive analysis of potential threats to the validity of DiD design in the analysis of non-pharmaceutical interventions to fight the spread of COVID-19.

[25] Specifically, Allcott et al. (2020), Maloney and Taskin (2020)

A Appendix: Theoretical Structure of SIR and Spatial-SIR

In this appendix we construct the theoretical structure of SIR and Spatial-SIR as Markov chains processes.

A.1 SIR

The society is populated by N individuals. Agents are ex-ante identical in terms of demographic characteristics. Let $\mathcal{S}$ denote the individual state-space. In the SIR model, the state-space is $\mathcal{S} = \{S, I, R\}$, indicating Susceptibles, Infected, and Recovered. Let $h_t^i \in \mathcal{S}$ denote the state of agent i at time t. Let $h_t = \frac{1}{N}[S_t, I_t, R_t] \in \Delta^{\mathcal{S}}$ denote the distribution of the population across the state-space.[26] The SIR model is represented by a Markov Chain:

$$prob(h_{t+1}^i = h' \mid h_t^i = h) = T_{h\,h'}(h_t)$$

where $T_{h\,h'}(h_t)$ is the generic element of a $\mathcal{S} \times \mathcal{S}$ double-stochastic (transition) matrix $T(h_t)$. The dependence of the transition matrix on h_t, the distribution of the population across the state-space (the aggregate state of the economy), is a mean-field property justified in this class of models by random matching in the population.

More specifically, the matrix $T_{h\,h'}(h_t)$ is determined by the following transitions:

S $\longrightarrow$ I. A Susceptible agent becomes infected upon contact with an Infected, with probability πI.

A $\longrightarrow$ R. An agent Infected at t, at any future period, can Recover with probability ρ.

R Recovered is absorbing state of the dynamic process (agents entering this state never leave). This assumes Recoved agents are immune to infection.

The resulting dynamical system for the distribution of the population across the state-space, h_t, is the following,

$$h_{t+1} = T(h_t)h_t.$$

The dynamical system can be solved for in closed form, see e.g., Moll (2020), Neumeyer (2020).

A.2 Spatial-SIR

We now add a spatial dimension to the SIR model. We also expand the state space to better caputure several relevant aspects of the SARS-CoV-2 infection.

[26] Abusing notation, we let we let the capital letters indicating a state also denote the fraction of the population in that state; and we let $\mathcal{S}$ denote both the set and its numerability.

Specifically, we split the I state into Asymptomatics and sYmptomatics, A and Y. We also add explicitly the state D, for Dead. Hence, $\mathcal{S} = \{S, A, Y, R, D\}$. We maintain the notation $h_t^i \in \mathcal{S}$ to denote the state of agent i at time t; and $h_t = \frac{1}{N}[S_t, A_t, Y_t, R_t, D_t] \in \Delta^{\mathcal{S}}$ to denote the distribution the N agents in the population across the state-space.

Agents are located in space, e.g., a lattice, which we call "the City." Agents are ex-ante identical in terms of demographic characteristics and symmetric in terms of location in space. A (Markov) transition process between states governs the dynamics of the system from the initial condition, at day $t = 0$. The spatial dimension maps the stochastic process into a local interaction model, a model in which agents' contacts are not the results of random matching but rather of local matching, with agents close in space (geographical distance as a metaphor for social distance). Let H_t denote the configuration of agent at time t, a vector $\left[h_t^1, h_t^2, \ldots, h_t^I\right]$; the set of all configuration is denoted $\mathcal{H}$. The local interaction model is characterized by

$$prob(h_{t+1}^i = h' \mid h_t^i = h) = T_{h\,h'}(H_t).$$

More specifically, the matrix $T_{h\,h'}(H_t)$ is determined by the following transitions:

S $\longrightarrow$A. Susceptible agents become infected upon contact with an Asymptomatic, with probability π.[27] A contact is defined to occur when agents are at a geographical distance in space $\leq p$.

A $\longrightarrow$ Y, R. An Asymptomatic agent infected at t, at any future period, can become sYmptomatic with probability ν, or can Recover with probability ρ.

Y $\longrightarrow$ R, D. An agent who has become sYmptomatic at t, at any future period, can Recover with probability ρ, or can Die with probability δ.

D, R. Dead and Recovered are absorbing states of the dynamic process. As we noted, this assumes Recoved agents are immune to infection.

Abusing notation, a transition matrix $T(H_t)$ in the space of possible configurations $\mathcal{H}$ can be constructed from $T_{h\,h'}(H_t)$.[28] The resulting dynamical system for configurations H_t is

$$H_{t+1} = T(H_t)H_t.$$

But Spatial-SIR accounts for agents possibly coming into contact after moving randomly in space.[29] Let the operator P_t, mapping $H_t \in \mathcal{H}$ into $P_t \circ H_t \in \mathcal{H}$, represent a configuration after a random permutation of the position of the agents,

[27] Susceptible agents are not infected upon contact with a sYmptomatic agent; this is to capture the fact that sYmptomatic agents are either isolated at home or in the hospital. They movements in the City are vacuous.

[28] This is an ugly looking operation, but formally straighforward, as purely arithemetical.

[29] This is different from most mathematical literature on local interactions; see e.g., Kindermann and Snell (1980) and Liggett (2012).

indexed by i. Before transitioning from the state at t to the state at $t+1$ the agents' locations are permutated randomly. The local interaction model is characterized by

$$prob(h^i_{t+1} = h' \mid h^i_t = h) = T_{h\,h'}(P_t \circ H_t).$$

The resulting dynamical system for configurations H_t is:[30]

$$P_t \circ H_{t+1} = T(P_t \circ H_t)P_t \circ H_t. \tag{10}$$

The dynamical system is difficult to formally characterize, besides (possibly) an ergodicity result, with respect to initial conditions specifying, at day $t = 0$, a random allocation of agents on evenly spaced locations in the City, all of them Susceptible, excepts for $A_0 > 0$ agents who are exogenously infected Asymptomatics. All our simulations, for all parameter values and initial conditions, converge to a unique ergodic distribution over the state space $h_t = \frac{1}{N}[S_t, A_t, Y_t, R_t, D_t] \in \Delta^{\mathcal{S}}$.

B Behavioral SIR

Consider a representative agent in SIR. Assume infected agents have no choice problem. Assume any susceptible agent can instead choose the number of contacts he/she has daily, c. His/her contemporaneous utility is increasing and concave in contacts, say $u(c)$. The present discounted utility he/she obtains if he/she gets infected is a constant V_I. Discounting the future at a daily rate δ, we can write the agent dynamic choice problem recursively as:

$$V(h) = \max_c u(c) + \delta \left[p\left(c, h\right) V_I + \left(1 - p\left(c, h\right)\right) V(h')\right]; \tag{11}$$

where $h = (I, R)$ is the state of the system,[31] $V(h)$ the value function, and $p(c, h)$ the probability a susceptible agents infected each one day, which in SIR is

$$p(c, h) = \pi c \frac{I}{N}. \tag{12}$$

[30] This representation is complicated in that the state space $\mathcal{H}$ is very large, and the permutation does not help. A simpler representation of $prob(h^i_{t+1} = h' \mid h^i_t = h)$ can be obtained as follows. Let I_t map locations $l \in \mathcal{L}$ into agents $i \in \mathcal{I}$. Assume at time $t = 0$ the map I_0 is an identity map so that the index i coincides with l. (This assumes, just for simplicity, that the numerability of agents is the same as that of locations.) Let $I_{t+1} = P \circ I_t$, $t \geq 0$. Fix an agent i and let l be the unique solution to $I_t(l) = i$. (As we constructed it, I_{t+1} is a byjection.) Let $NBHD_t(i) = \{i \in \mathcal{I} \mid i = I_t(l'),\ l - d \leq l' \leq l + d\}$. Then

$$prob(h^i_{t+1} = h' \mid h^i_t = h) = T_{h\,h'}([h^{i'}_t]_{i' \in NBHD(i)}).$$

[31] We are abusing notation. We have defined $h = (S, I, R)$, but S is residual.

The agent takes the state h and h' as given. The first order conditions (necessary and sufficient under concavity) are:

$$u'(c) = \delta\pi\frac{I}{N}\left[V(h') - V_I\right] \tag{13}$$

where $[V(h') - V_I] > 0$ (to avoid trivialities) and $V(h)$ decreases with I and increases with R (S is residual) [this needs to be shown]

At the equilibrium dynamics of the system, $h' = (I', R')$ satisfies

$$I' = I + \pi c(1 - R - I)\frac{I}{N} \tag{14}$$

$$R' = R + \rho I \tag{15}$$

While the existence (and possibly uniqueness) of an equilibrium needs to be formulated as fixed point in the (infinite dimensional) space of sequences $(I, R)_t$, these sequences are observed in empirical work and hence this equilibrium condition need not be imposed in empirical analysis.

Consider now a perturbed problem in which $c = \alpha d\Psi$ and the agent chooses α, at a convex cost $C(\alpha)$ such that $C(1) = 0$. It is straighforward to show then that

Proposition 1. *Given (h, h') if d is higher, α is lower, but $c = \alpha d\Psi$ is higher.*

C Appendix: Calibration

We calibrate the parameters of Spatial-SIR to the dynamics of the SARS-CoV-2 epidemic.[32] The parameters we choose in the calibration of the aggregate model are summarily reported in Table 1. We discuss in turn the parameters governing the transitions away and between the infected states, A, Y, D, R, and then the infection and contact rates governing how Susceptibles are infected. Finally, we set initial conditions.

Geography We place people in initially random location in a square or area equal to 1. At all $t > 0$, individuals are relocated at distance μ from their location at $t - 1$, in a direction drawn randomly from a uniform distribution on $[0, 2\pi]$

Initial conditions At time $t = 0$ we set 30 individuals in Asymptomatic state; all others are Susceptibles. In all specification where we do not test for the effect of cluster location, the asymptomatics at $t = 0$ are those initially located in a position

[32] 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: Parameter values in the calibration

Parameter	Notation	Value
number or people	N	26,600
initially infected	A_0	30
prob. of recovery	ρ	0.05
prob. of becoming symptomatic	υ	0.09
prob. of dying	δ	0.00013
contagion probability	π	0.038
mean distance traveled	μ	0.034
contagion radius	p	0.013

closest to location $[x = 0.25, y = 0.25]$.

Transitions away and between the infected states, A, Y, D, R. The probability any agents transitions away from being Asymptomatic, state A, is $\rho + \nu$ in our model. We assumed agents are infective only in state A (we assumed that all sYmptomatic agents reduce to zero social contacts). The average time an agents stays in state A is then $T_{inf} = \frac{1}{\rho+\nu}$. We set $\rho+\nu$ to match a theoretical moment which holds exactly at the initial condition in the basic SIR model. Recall 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,

$$\frac{(\mathcal{R}_0 - 1)}{T_{inf}} = g_0. \tag{16}$$

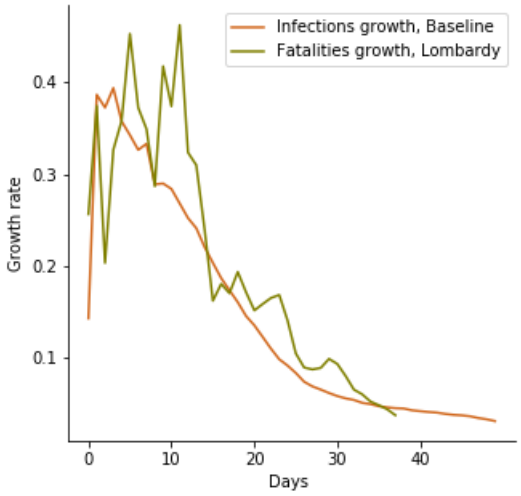
For the current SARS-CoV-2 epidemics, R_0 is reasonably estimated between 2.5 and 3.5.[33] The daily rate of growth of the infection g is estimated = .35 by Kaplan et al. (2020).[34] This implies, from Equation (16), that T_{inf} is between 4 and 7 days (respectively for $\mathcal{R}_0$ between 2.5 and 3.5). Ferguson et al. (2020) use 6.5 days. We set $\rho+\nu = .14$, so that $T_{inf} = \frac{1}{\rho+\nu} = 7$.[35] Furthermore, the average time from infection to death or recovery is reasonably estimated to be 20 days; see e.g., Ferguson et al. (2020). Therefore we set $\rho = .05$ so that $1\rho = 20$. This implies $\nu = .09$.

[33] More preciseely, Wang et al. (2020) estimates $\mathcal{R}_0 = 3.1$ for Wuhan, China; Remuzzi and Remuzzi (2020) estimate it between 2.76 and 3.25 for Italy; Zhang et al. (2020) has 2.5 from the Princess Cruise ship; Fauci et al. (2020) estimates $\mathcal{R}_0 = 2.2$ in the U.S.; the European Centre for Disease Prevention and Control, at https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases, estimates $\mathcal{R}_0$ between 2 and 3 and the last Imperial College report, Ferguson et al. (2020) uses 3.5. Note that the range of Recovered agents in stationary state is R^* implied by R_0 between 2.5 and 3.5, is between .89 and .97; from equation 2.

[34] Alvarez et al. (2020) have .2; Ferguson et al. (2020) have .15.

[35] We thank Gianluca Violante for suggesting this calibration strategy.

Figure C.14: Growth rate of infections



The case fatality rate, the probabilty of death if infected, is estimated between .005 and .01; see e.g., Ferguson et al. (2020). Since agents remain sYmptomatic in the model, before Recovering, on average $\frac{1}{\nu} = 11$ days, we set the probability of Death for a sYmptomatic, δ, to be 0.001.[36]

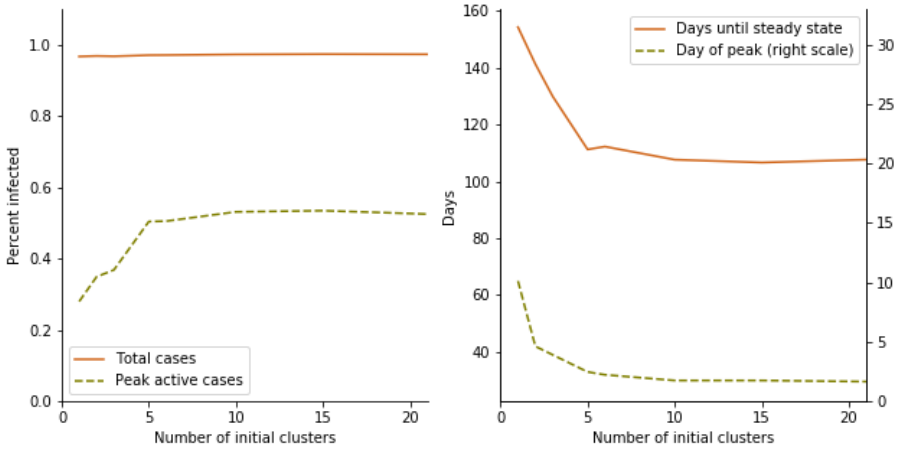
Infection and contact rates. We calibrate the infection rate π, the contagion radius p, and the mean distance μ to match i) the daily growth rates of the dynamics of infections, g_t, observed in the first 30 days of epidemics; and ii) the average number of contacts observed in demographic surveys. For g_t, we use data for Lombardy, Italy; see Figure C.14.[37] For contacts, data in Mossong et al. (2008) suggests an average of 12.5 contacts every day.

Behavioral models. In the simulation of the model (4) we set $\phi = 0.88$ 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.

[36] We also set, for simplicity in the simulations, that fatalities cannot occur to an agent less than 3 days before she becomes sYmptomatic and that every infected individual recovers with certainty after 100 days.

[37] Since the number of infections is not observed, we match the growth rates of infection in the model with the growth rate of deaths in the data. This is justified when, as we assumed, the case fatality rate is a constant and Death follows infection after a constant lag on average.

Figure D.15: The effect of the number of clusters

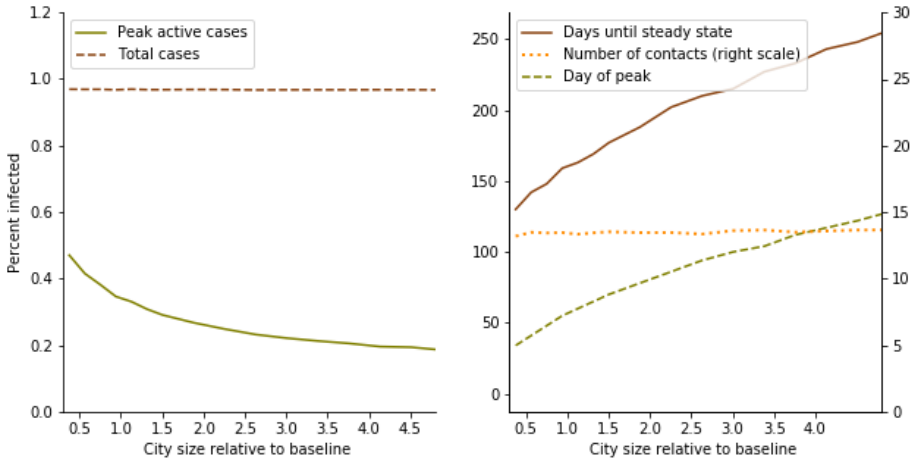


D Appendix: Additional Figures

In Figure D.15 we report the results of simulations varying the number of initial clusters from 0 to 20, in our otherwise baseline city, while keeping the number of initially infected agents constant. We observe that, with our calibrated parameters, the effect of increasing the number of initial clusters converges quite fast: when there are five or more initial clusters, increasing the number of initial clusters while keeping the number of initially infected the same, has no effect on the dynamics of the epidemics.

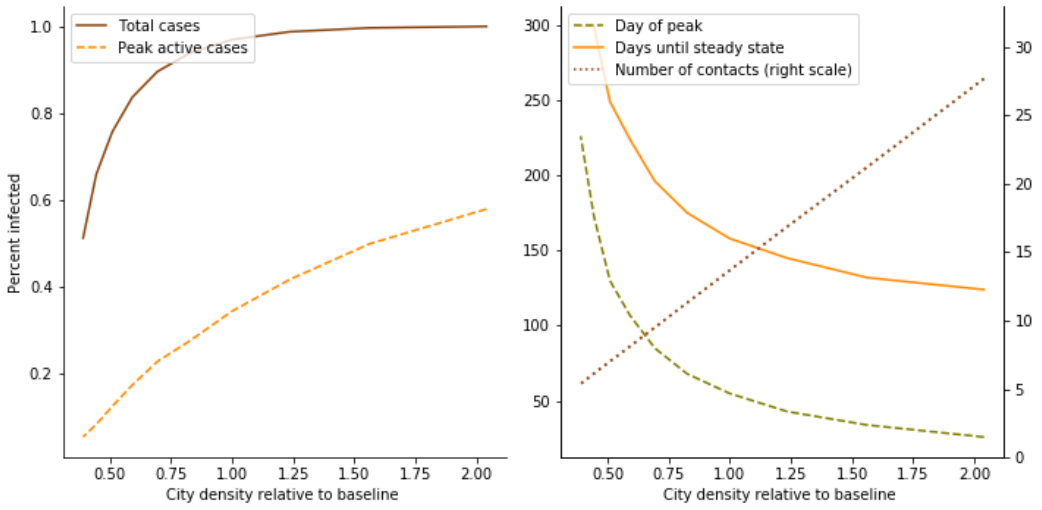
In Figure D.16 we report the results of simulations varying the size of the City, in our otherwise baseline City, while keeping the size of the initial outbreak of the infection and City density constant. We observe that, with our calibrated parameters, the effect of increasing City size: the peak of active cases declines with size in a convex manner (less so the larger the city); the number of days it takes to reach the peak and the number of days to the stationary state (the end of the epidemic) both increases with size and do so with a slight concavity (less so the larger the city).

Figure D.16: City size comparisons



In Figure D.17 we report the results of simulations varying City density, in our otherwise baseline City. We observe that, with our calibrated parameters: i) $(R^* + D^*)/N$ is increasing and concave in density; and the peak of $(A + Y)/N$ is also increasing and concave in density. In the right panel of Figure D.17 we see that the days it takes to reach the peak and the stationary state are decreasing and convex in density.

Figure D.17: The effect of varying city density with constant population



References

Allcott, Hunt, Levi Boxell, 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*. (Cited on page 22)

Alvarez, Fernando E, David Argente, and Francesco Lippi, "A simple planning problem for covid-19 lockdown," National Bureau of Economic Research 2020. (Cited on page 27)

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. (Cited on pages 3 and 4)

Bethune, Zachary A and Anton Korinek, "Covid-19 infection externalities: Trading off lives vs. livelihoods," *National Bureau of Economic Research Working Paper*, 2020. (Cited on page 4)

Bisin, Alberto and Andrea Moro, "Learning Epidemiology by Doing: Lockdown Policies and the Lucas Critique," *Working paper, NYU and Vanderbilt*, 2020. (Cited on pages 5, 18, and 21)

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. (Cited on page 6)

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. (Cited on page 3)

Conley, Timothy G and Giorgio Topa, "Estimating dynamic local interactions models," *Journal of Econometrics*, 2007, *140* (1), 282–303. (Cited on page 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. (Cited on page 22)

Eichenbaum, Martin S, Sergio Rebelo, and Mathias Trabandt, "The macroeconomics of epidemics," National Bureau of Economic Research 2020. (Cited on page 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. (Cited on page 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. (Cited on page 22)

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer, "Internal and external effects of social distancing in a pandemic," National Bureau of Economic Research 2020. (Cited on pages 4 and 18)

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 reduceCOVID-19 mortality and healthcare demand," Imperial College London 2020. (Cited on pages 6, 18, 20, 27, and 28)

Fernandez-Villaverde, Jesus and Charles Jones, "Estimating and Simulating a SIRD Model of COVID-19," mimeo 2020. (Cited on pages 2, 3, 18, and 21)

Geoffard, Pierre-Yves and Tomas Philipson, "Rational epidemics and their public control," *International economic review*, 1996, pp. 603–624. (Cited on page 3)

Glaeser, Edward and José Scheinkman, "Measuring social interactions," *Social dynamics*, 2001, pp. 83–132. (Cited on page 6)

Goenka, Aditya and Lin Liu, "Infectious diseases and endogenous fluctuations," *Economic Theory*, 2012, *50* (1), 125–149. (Cited on page 3)

Goodman-Bacon, Andrew and Jan Marcus, "Using Difference-in-Differences to Identify Causal Effects of COVID-19 Policies," *DIW Berlin Discussion Paper*, 2020. (Cited on page 22)

Grassberger, Peter, "On the critical behavior of the general epidemic process and dynamical percolation," *Mathematical Biosciences*, 1983, *63* (2), 157–172. (Cited on page 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. (Cited on page 4)

Hethcote, Herben W., "The mathematics of infectious diseases," *SIAM Review, 2O00*, 2000, *42* (4), 599–653. (Cited on page 4)

Hsiang, Solomon, Daniel Allen, Sebastien Annan-Phan, Kendon Bell, Ian Bolliger, 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. (Cited on page 22)

Jarosch, Gregor, Maryam Farboodi, and Robert Shimer, "Internal and External Effects of Social Distancing in a Pandemic," 2020. (Cited on page 3)

Kaplan, Greg, Ben Moll, and Gianluca Violante, "Pandemics According to HANK," *University of Chicago*, 2020. (Cited on page 27)

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. (Cited on pages 3, 4, 18, 21, and 28)

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. (Cited on page 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. (Cited on page 2)

Kindermann, R and JL Snell, "American Mathematical Society," *Markov random fields and their applications*, 1980. (Cited on pages 6 and 24)

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, Ilaria Dorigatti, and Andrea Crisanti, "Suppression of COVID-19 outbreak in the municipality of Vo, Italy," *medRxiv*, 2020. (Cited on page 2)

Liggett, Thomas Milton, *Interacting particle systems*, Vol. 276, Springer Science & Business Media, 2012. (Cited on pages 6 and 24)

Maloney, William F and Temel Taskin, "Determinants of Social Distancing and Economic Activity during COVID-19: A Global View," *World Bank Policy Research Working Paper*, 2020, *9242*. (Cited on page 22)

Mangrum, Daniel and Paul Niekamp, "College Student Contribution to Local COVID-19 Spread: Evidence from University Spring Break Timing," *Available at SSRN 3606811*, 2020. (Cited on page 22)

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). (Cited on page 2)

Moll, Ben, "Lockdowns in SIR models," *Author's website, Princeton*, 2020. (Cited on pages 4 and 23)

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. (Cited on pages 6 and 28)

Neumeyer, Pablo Andres, "Clase especial de epidemilogia," *Author's website, Class notes, Universidad Di Tella*, 2020. (Cited on pages 4 and 23)

Özgür, Onur, Alberto Bisin, and Yann Bramoullé, "Dynamic linear economies with social interactions," Melbourne Business School working paper 2019. (Cited on page 6)

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. (Cited on page 22)

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. (Cited on page 6)

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. (Cited on pages 4 and 18)

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. (Cited on pages 3 and 18)