Epidemics with Space, Movement, and Asymptomatic Spreading

Blake LeBaron *

Brandeis International Business School

Brandeis University

February 12, 2021

Abstract

This paper describes a simple agent-based model for disease spread and epidemic dynamics. The model is stylized and simple, yet tries to capture some of the core features that are understood to be part of the COVID-19 progression. The model is strongly spatial, but is built in a basic featureless environment to allow for spatial-temporal patterns to emerge endogenously. A key driving force of many of the results is the presence of asymptomatic agents who are contagious and spreading the disease more freely than symptomatic types due to differing levels of mobility. Simulations over several sets of parameters display richer, more realistic dynamics than standard benchmark (SIR) models which can be replicated as a special case. Many of the results show dramatic differences between the fully spatial, and SIR type mechanisms. Finally, the model is used to explore key policy interventions including movement restrictions, testing/contact tracing, and vaccination.

Keywords: Contagion modeling, agent-based, COVID-19, epidemic interventions

*415 South Street, Mailstop 32, Waltham, MA 02453 - 2728, blebaron@brandeis.edu, www.brandeis.edu/~blebaron. The author is grateful to seminar participants at George Mason University for many useful comments and suggestions.

1 Introduction

The COVID-19 global pandemic has driven an explosion in academic research. This event, which at the time of this writing is only one year old, has caused an outpouring of research from many different fields. There are probably many good reasons for this explosion in research activity, including most importantly, the seriousness of the problem, but also the current availability of data, and increased compute power.

This paper presents another model looking at epidemic dynamics in general with a slight lean to COVID-19 related parameters. It is closest in spirit to the original SIR model in trying to represent a platform to visualize a highly stylized outbreak in a uniform spatial environment with relatively similar agents. It emphasizes several key aspects for disease spread in a world where space and agent motion is critical. First, agents will go through a period where they are infected with the disease, but are unaware of this. This is referred to as presymptomatic, and has been commonly referenced as a tricky aspect of COVID-19. Second, the stylized spatial model leads to interesting endogenous patterns, or emergent features that impact the spread of the disease. Outbreaks often take place at spatially distanced pockets, running their course asynchronously. Furthermore, clumping of disease outbreaks in one local space can lead to a local pocket of powerful herd immunity, stopping spread of the disease early in its progression.

Given the availability of massive and detailed data sets, why should we care about a stripped down model that is not particularly fit to any data set? It should be taken first as a useful thought experiment, or a way to get one's head around the dynamics of a model which moves a notch up in realism from the benchmark SIR framework. Another reason is that it gives some indication of which parameters are important for disease spread. This can potentially inform empirical work in terms of where to look in our sea of data, and which values should one work hard to precisely pin down. It also provides a model based monte-carlo platform that can help in understanding methods for estimating key parameters. Finally, it gives a picture of generic aspects in a spatial model. Nothing about it is specific to any real world spatial geography. The benefit of this, is that it can give ideas as to which features of detailed models will travel well when applied and fit to new localities. This is almost a kind of out of sample forecasting feature.

The model generates several features related to the goals just described. First, it can produce relatively realistic time series for epidemic progression. Series where a disease can fade out rather slowly, and occasionally reappear are common in the simulations. It also demonstrates an extreme sensitivity to the level of asymptomatic/presymtomatic activity in the agent population. It displays several critical policy features, some of which are expected, and some are surprising. The most surprising is that contact tracing proves

somewhat ineffective in many cases. For almost all cases it is shown that the fully spatial model generates very different results from the more SIR-like model where space is not a critical feature.

The landscape of infectious disease modeling is enormous. A review of this literature is a daunting task. There are several key surveys and discussions of policy which have appeared recently.[1] Modeling tools seem to split into several relevant bins in terms of connections to this paper. Almost all models trace back some historical roots to the original SIR model of Kermack & McKendrick (1927). Many models used for policy analysis, and forecasting outbreaks, take this model as their core, but augment it in several dimensions. They add many different compartments for the disease to describe individual progression through the disease. Also, the SIR framework becomes a local core for a large scale networked model which includes information on geography, transportation, and contact frequencies between local communities.

The approach used here is agent-based in the sense that it is trying to model individuals and not groups or entire populations. An early example of this can be found in Burke, Epstein, Cummings, Parker, Cline, Singa & Chakravarty (2006) which takes detailed data from a smallpox outbreak, and uses this to build out a realistic small town setting. A recent example of this style of research as applied to COVID modeling is Aleta et al. (2020) which brings modern large data sets to the problem. Agent motion is estimated from geolocation data from a large scale (about 80,000 individuals) anonymized data collection system monitoring locations and contacts.

The model presented here also touches on the large literature placing space and geography as the critical part of disease spreading. An early example of this is Carpenter (1974). Another early example of an explicit spatial approach modeled through transport networks is Bertuzzo, Casagrandi, Gatto, Rodriguez-Iturbe & Rinaldo (2009). This paper also explores some simplified network structures for comparison. A very large scale system handling space is the GLEaM system, Balcan, Goncalves, Hu, Ramasco, Colizza & Vespignani (2010). Several recent papers have presented spatially explicit models for COVID-19 behavior. They include Gatto, Bertuzzo, Mari, Miccoli, Carraro, Casagrandi & Rinaldo (2020), Munchi, Roy & Balasubramanian (2020), and O'Sullivan, Gahegan, Exeter & Adams (2020). The latter is an explicit spatial model applied to New Zealand, and uses a Netlogo framework linked to geographic data. An empirical critic of the SEIR modeling framework is performed in Getz, Salter & Mgbara (2019) which shows how many empirical results differ when one moves to smaller, and more homogeneous subregions in a country. Also, there have

[1] Economists might be most interested in Avery, Bossert, Clark, Ellison & Ellison (2020). Agent-based modelers will find the summary of these tools in Willem, Verelst, Bilcke, Hens & Beutels (2017). For a general survey of mathematical models of infectious diseases there is Siettos & Russo (2013). Recent commentary specifically directed at COVID-19 modeling is given in Vespignani, Tian, Dye, Lloyde-Smith, Eggo, Shrestha, Scarpino, Gutierrez, Kraimer, Wu, Leung & Leung (2020).

been analytic approaches to rethinking standard infection models as in Vrugt, Bickmann & Wittkowski (2020). An early model demonstrating the richness of dynamics which can occur in a theoretical spatial environment is Lloyd & May (1996). Network connections are often a representation for distance in many models. Several articles which are interested in network structure from a theoretical perspective are Lang, Sterck, Kaiser & Miller (2018) and Wolfram (2020).

Asymptomatic spread is a big part of the model considered here. This is driven by growing evidence that this is critical to the spread of COVID-19, (Gandhi, Yokoe & Havlir 2020). Some of the most detailed evidence for this is from a skilled nursing facility, (Arons et al. 2020). The authors found that nearly half of residents testing positive for the disease were asymptomatic.[2] In He et al. (2020) the authors look at viral shedding from throat swabs, and found that the maximum was at or near the onset of symptoms, suggesting a large amount of presymptomatic infection.

Research coming from economists on COVID-19 has been extensive. Avery et al. (2020) give references, and a road map to this work. Papers have concentrated on endogeneity and behavioral feedbacks leading to changes in infection dynamics, agent heterogeneity, and political economy.[3] All of these, except maybe the last area are also part of the epidemiological literature.[4] An important contribution of economists is to try to understand macroeconomic shocks coming from the pandemic, and to build general models of the economy that can measures the costs of the shocks and to guide policy makers for reopening strategies.[5]

This paper proceeds as follows. Section 2 gives details on the primary model used along with comparisons to more standard approaches. Section 3 presents all the model simulations and comparisons. Most of the model output is viewed through a limited set of plot formats from a common monte-carlo comparison across many different parameter sweeps. Section 4 will conclude and present ideas for the future usefulness of this approach.

[2]Most did later become symptomatic.

[3]Another useful primer for economists is John Cochrane's blog post at `https://johnhcochrane.blogspot.com/2020/05/an-sir-model-with-behavior.html`. This gives a simple SIR model with behavioral feedback, and contains some recent references to other work.

[4]For an agent-based approach to endogeneity of behavior see Epstein, Parker, Cummings & Hammond (2008).

[5]Several examples in this area are Acemoglu, Chernozhukov, Werning & Whinston (2020), Scherbina (2021), Bloom, Kuhm & Prettner (2020), and Ludvigson, Ma & Ng (2020).

2 Model structure

2.1 SIR models

Most models of epidemics get some of their foundations from the standard SIR model of Kermack & McKendrick (1927) shown in the following system of differential equations,

$$\frac{dS}{dt} = -\beta(\frac{I(t)}{N})S(t) \tag{1}$$

$$\frac{dI}{dt} = \beta(\frac{I(t)}{N})S(t) - \gamma I(t) \tag{2}$$

$$\frac{R(t)}{dt} = \gamma I(t). \tag{3}$$

It is built from a population which occupies three states, S (susceptible), I (infected), and R (recovered). Agents move from S to I as S agents bump into I agents in the population. Their probability of hitting an infected agent is given by the ratio $I(t)/N$ which implies that all agents are uniformly moving around through space, or the concept of space simply doesn't exist. The other two equations build infected agent changes from flow out of S less the flow into the recovered state. Finally recovered, which can be dead, is the flow out of the infected state. Recovered agents are assumed to be immune. This model is nearly a century old, but it forms the core both for model building and intuition in the world of infectious diseases.

2.2 Spatial model

The primary model considered here is a spatial model based on standard parts from SIR style models. Agents move in a two dimensional cartesian space. For most periods, their behavior is to move one unit in a random direction. Although model calibration is rough, it is assumed that each finite tick of the model is equal to one day. Also, the space is assumed to wrap both from the end on the right, to the start on the left, as well as top to bottom.

There are four states that agents can move through in the model. They begin as susceptible which means they have not been exposed to the disease and might get infected. Their second state is presymtomatic (or asymptomatic). In this state agents have been exposed to the disease, but are not showing any symptoms. Presymtomatic agents follow the same motion rules as do susceptible agents since they are unaware they are sick. Agents next move on to being infected, and in this state they show symptoms and decide to "stay home". For them motion stops in the spatial grid. The final state is recovered where they have passed the

disease and are free to move again. They obtain permanent immunity at this point. One should always keep in mind that some fraction of recovered agents will be dead, so the total deaths in any outbreak are proportional to the recovered.

Agents infect others by landing on the same spot in space. Susceptibles are given the disease by both infected and presymtomatic types with probability p_{infect}. At the start of the run random numbers are drawn for each person that determine the course of the disease. The length of infection is drawn from a uniform distribution $U[10,14]$ days which is a reasonable length for COVID-19 infections. Little is understood about the asymptomatic features of the disease, but they do occur. This asymptomatic period is also crucial to this study, so it will be allowed to vary over different runs. Each agent falls into the presymtomatic bin for a total of $U[0,P_{max}]$ days. Setting $P_{max} = 0$ eliminates asymptomatic behavior from the model, and making it large allows for a lot of infection driven by agents in motion showing no symptoms.

Disease progression using only four boxes (SPIR) is very stylized relative to many other attempts at COVID-19 modeling. For example, asymptomatic types can be modeled as never entering the infected stage of the disease. Also, there can be a latent period in the presymtomatic stage, when agents are infected, but not infectious. The basic idea is to keep the model as simple as possible, enhancing intuition, and minimizing parameters.

The spatial model can be changed to mirror the SIR infection mechanism while maintaining its core structure. This is done by changing agent movement from a unit spatial random walk to a complete random location change. When an agent is asked to move it simply reappears at a random spot in the 2D space. This will be referred to as a "diffuse" style of movement for agents. It is essentially replicating the core assumption in the SIR where space does not matter, and only the various concentrations of different types of agents contribute to infection probabilities.

Two other features will be important for agent motion. First, agents are allowed, with low probability, to jump long distances. This is referred to as “long range travel”. This distance on each jump will be drawn $U[0,d_{maxjump}]$ units. The probability will be controlled by p_{long} for the probability that a long (versus local) trip is taken each period. As with short range travel, long range travel is not allowed when an agent is in the **infected** state, but is allowed for all other states.

2.3 Measuring epidemics

This model is capable of generating very different disease progressions even for the same set of parameters. Many epidemics shut down after only a few days, while others go on for 100's of days with extensive infection through the population. This means it is important to look at cross sections of multiple simulations requiring that runs be summarized with a small set of variables.

The first set of three variables are related to the severity of each run. "Recovered" estimates the mean fraction of recovered agents. Assuming some fraction of recovered died, this number is the most basic estimate of overall death from the disease. A second estimate is the maximum infection fraction. This is the maximum population fraction of infected people during the course of the disease. It is not necessarily related to total deaths, but it is an indicator of how stressed local health systems might become. One could have a very long lived epidemic, which kills many people, but it would never have a very high infected fraction, so hospitals would not become over extended. The final measure reported is an estimate of $R0$. This value is the estimate of how many others each agent infects, and is a standard summary statistic measured for many diseases. Each agent keeps a record of others infected, and this value is averaged over a recent set of recovered agents. Other agent states are not used, since these are still in progress in terms of infecting others. This is a dynamic value, and it moves toward 0 as the entire population is recovered. The most reasonable estimate to compare with in the field empirical studies is probably the maximum $R0$ observed during any pandemic, and this will be reported.

The time series results of each epidemic spread are relatively complex, and difficult to summarize with a single number. Two numbers give some basic features which try to distinguish between the spatial and SIR types of models. The first is the length of the outbreak from start to the time step when the virus finally dies out, t_E. The second measure tries to capture asymmetry in the time dynamics of the virus. It reports the following value,

$$t_E - 2 * \underset{t}{\operatorname{argmax}}\, I(t),$$

or the final time period less twice the time period when infections hit their maximum value. If the epidemic were perfectly symmetric in terms of infections, then this value would be zero. If there is a long persistent run down of the infection rate, then this value will be positive. The standard SIR model is not symmetric itself, but it is often quite close to zero for this measure.

2.4 Policy interventions

The computational experiments will use three policy interventions. They are all implemented in a stylized, simplified fashion. The first tries to replicate restrictions on mobility and social interactions. The mobility of all agents is uniformly restricted. Each time an agent is ready to move in space, it will draw a random variable, and then only proceeds according to a given probability, P_{move}. It is a kind of probabilistic stay at home order. This slows the spread of the disease by restricting motion of susceptible and presymptomatic agents, and also reduces the amount of contacts between all agents. It is a stand in for almost all motion and interaction restricting policies. Long range travel is also restricted by this probability.

The second policy tool implements a testing and contact tracing program. All agents are subject to testing with a given probability. Tests are assumed to be completely reliable with no false negatives or false positives. Agents testing positive will be quarantined for 15 days. This shuts down the movements of presymptomatic agents, and for most runs the quarantine period is longer then the presymtomatic period, so they step right into the symptomatic phase of the disease which restricts movement too. All agents keep track of a recent set of contacts. When "contact tracing" is activated, a positive test for an agent will also quarantine the entire set of contacts for 15 days as well. In this way, the policy tries to replicate standard testing/contact tracing as implemented in many areas.

The final policy intervention is concerned with vaccination. As with the other experiments this one is a very simplified version designed to explore the impact of an already immune population. The simulation starts in the same way the other runs do with a single infected individual, but some fraction of the population is given immunity. In the model they are moved to the recovered state before the run even starts. This would correspond to an outbreak started in a population with a given fraction already immune.

3 Computational experiments

The model is simulated using the Netlogo platform.[6] Netlogo is well designed for simulations where space and motion are critical. It also allows the creation of a rich user interface that allows for dynamic interactions between experimenters and models.[7] Each run of the model starts with a single infected individual chosen at random. Figure 1 gives a view of what the actual model looks like with the panel for the entire

[6] Wilensky (1999) is the core software, and is available open source for all platforms, `http://ccl.northwestern.edu/netlogo`. Wilensky & Rand (2015) is an excellent introduction to Netlogo and agent-based modeling.

[7] The model used here also gets much early inspiration from code in Brearcliff (2020). This netlogo implementation sets up some of the basic dynamics of a spatial infection model using the SEIR variant of the SIR model (close to the SPIR framework used here), and also puts some basic restrictions on how movement can proceed under a lockdown policy.

space, various sliders and controls, and plots of key variables in real time as the model progresses.

Table 1 gives a summary for the key parameters of the model. Many of these are flexible, and can be changed by the user, but they are held fixed in most of the simulations. The model will be simulated across several parameter sweeps to experiment with their impact. The Netlogo "Behavior space" system is used for this. Runs will report the 5 different summary statistics previously described, and the mean over 25 runs is reported in the figures.

Figure 2 gives an initial picture of the dynamics of the spatial model. Time series are presented for an single run. The presymtomatic range is $U[0,10]$ days which will be standard for many of the future runs. Long range travel probability is set to zero. The upper panel displays the progression of the disease through the two critical states of presymptomatic and infected. The disease shows many features common to pandemic time series, but which are difficult to quantify. There is a well defined peak in infections about about 100 days, but it also displays several other local maximums. This is not quite perfect cyclical behavior, but also not a perfect single peaked progression. Also, the time series displays a long right tail in that the disease takes a while to finally get shutdown.

The lower panel shows the estimates of $R0$ as the model moves through time. Its maximum occurs early in the progression at near 2.5.[8] It then quickly falls to near 1, and will move back and forth around this level for the remainder of the run. It should be noted that the small infection peaks often correspond to local moves of $R0$ above zero, while the troughs correspond to periods when $R0$ falls below zero. Several of our later summary statistics can be seen in this single plot. The maximum infection rate is about 0.04. The total length of the epidemic is about 500 days. The asymmetry would be approximately $500 - 2 * 100 = 300$ days. (100 is the approximate time of the maximum infection.) Finally, $R0$ maximum would be 2.5.[9]

The second primary comparison model is the diffuse model where agents move randomly (as opposed to locally) in space. Figure 3 repeats the previous plots for this model. The dynamics are clearly very different. They now appear much closer to a classic SIR model. The infection and presymtomatic series are strongly single peaked, and are close to symmetric in time. The maximum infection fraction is large at 0.4, and the disease dies out quickly a little after day 100. Interestingly, the maximum $R0$ level is similar to the previous run with a value near 2.5, but in this case the model maintains a steady value for a long period of time which is also consistent with SIR theory. The estimated asymmetry value would be about

[8] There have been many attempts to estimate R0 for COVID-19. This value falls within the wide range that is currently considered reasonable.

[9] An interesting related simulation example is contained in Charbonneau (2017) which also simulates infections moving through a simple spatial grid. These also generate a relative rich time series showing near erratic behavior with many local maximums and minimums.

$120 - 2 * 65 = -10$, or relatively close to zero in comparison to the spatial model. There is no evidence for a long right tail in the time progression. This two model time series contrast couldn't be more stark. It will now echo through most of the following simulations.

3.1 Baseline and asymptomatic spread

The first set of simulations vary the amount of time agents are presymtomatic. In this state they are infected, and are also contagious. Because they are not yet showing any symptoms they move around freely and infect others. This is in contrast to the infected (symptomatic) state when they are still contagious, but stop all movement (ie. stay home).

The presymptomatic length is chosen randomly for each agent and varies over a range of $[0, a_{max}]$ days. Figure 4 shows the three measures representing the severity of the epidemic. It reports outcomes for both the fully spatial model which are labeled "Spatial". It also reports the benchmark comparison, SIR like model which is labeled "Diffuse". As previously described, motion in this model is not local. Agents diffuse across space in a completely random fashion. This aligns with the perfect mixing assumptions that go along with a baseline SIR model with a single spatial compartment. The contrast between these two models will be crucial.

The top panel of figure 4 shows the recovered fraction for different amounts of presymptomatic behavior. The first thing to notice is that for this model structure and parameters, presymptomatic behavior is necessary for the epidemic to start. Recovered represents the agents who have had the disease and are recovered, so when this value is low, or zero, it indicates that the extent of infection was small. Also, it is important to remember that some fraction of recovered will be dead, so this is also related to the total deaths caused by the disease. For both the spatial and diffuse models this value is increasing in the presymtomatic level, but the diffuse model increases much more quickly. Even with only a maximum presymptomatic period of 5 days it already shows an infection fraction of near 75 percent.

The greater spread and severity of the diffuse model also appears in the second panel which reports the maximum infected fraction. This is the maximum fraction of the population in the infected state as the disease progresses. The uniform mixing in the diffuse, SIR type model spreads the disease quickly through the entire population, so things get very bad, very fast. The fully spatial model requires the disease to be spread by the relatively slow motion of the agents, making most of the pockets of infection local. Both increase with the presymtomatic levels, but they are never very close.

The final measure looks at the maximum $R0$ estimate from the runs. This is a rolling estimate of the average number of people infected by each recovered person. The value is the average over 25 runs of the maximum reported for each run. It is increasing over the range, and also shows a higher value for the diffuse case. It is interesting that the actual numerical value for the spatial model is near 2 which is often quoted as a reasonable estimate for COVID-19.

Figure 5 reports the time series features across the same set of simulations. These attempt to summarize the time trajectory of the full epidemic in two numbers. The first panel shows the actual length of time the disease is present in days. Consistent with the previous figures the Diffuse experiments are much shorter in length. The disease moves quickly through the population, and robs itself of susceptible types in a very short time. The fully spatial outbreaks can last much longer with a maximum of a little over 500 days. Both plots show a nonmonotonic relationship. In the first phase, increasing the presymptomatic level increases the fraction of bad outbreaks of the disease, leading to increases in the time length. However, further increases lead to more intensity in the epidemic with a quicker spread, and shorter time frames.

The lower panel displays the asymmetry in the epidemic dynamics. It measures the difference between the start $(t = 0)$ to the maximum infected time period and the finish to the maximum infected period. A disease run with a large "right tail", or very persistent and slow decay, would generate a positive asymmetry measure. The figure shows that the diffuse model is very symmetric with regular increases and decreases in infections. The results are different for the spatial model. For many runs the disease fades out very slowly off its peak, generating large positive values. The results for the spatial model are again nonmonotonic with a maximum of 200 near $a_max = 8$. The reason for this probably parallels that for duration. Initially, epidemics are getting more intense and causing a longer time to run out, but as this intensity increases the asymmetry falls along with the duration.

Figure 6 further pushes the comparisons with an SIR benchmark. In this case the diffuse case is compared with a "Full SIR" model. This model is still operated in space, but agents continue to move randomly. However, this model drops the distinction of motion between symptomatic and presymptomatic types. Symptomatic agents are now also allowed to move freely through space, and are not distinct from presymptomatic types. In this sense it is now very close to a traditional SIR model in space and time. The first two panels in figure 6 display comparisons between the two models. The fraction of presymptomatic types still can have some impact on the Full SIR model since this value does impact the total days in which an agent is infectious. However, the graph shows little impact of this with a severe outbreak showing up for all values of a_{max}. For lower levels of a_{max} there is a large difference between the two models which is

sensible since with the diffuse case most infected agents are locked down (but still infectious). Panel two shows a similar pattern for the severity in the two models.

The lower panel in figure 6 displays the estimated maximum $R0$ for the two models. The plot also shows the theoretical value from the SIR model. In a traditional SIR model R0 is given by,

$$R0 = \frac{\beta}{\lambda} \tag{4}$$

which approximates the others an agent will infect over the average length of their infected period. For the discrete version of the model this equation is approximated in several ways. λ will be $1/(0.5 * a_{max} + 12)$ where the two items in the denominator are the expected length of the presymptomatic and symptomatic periods respectively. β is the expected number of interactions on each day. This is estimated by

$$\beta \approx \sum_{i=0}^{\infty} \Pr(N = i) i \tag{5}$$

where $\Pr(N = i)$ is the probability of i agents sitting on a given patch. It can be determined easily from a binomial distribution. The above sum is only taken out to $i = 10$ since it this probability drops to zero quickly. The plot shows a pretty good agreement between the theoretical level and the estimated benchmark for a_{max} up to 12. Both are well above the diffuse case with values in the range of 3 to 4. Beyond $a_{max} = 12$ the values start to diverge. It is not clear exactly why this is happening, but one possibility is that this simulation points out some of the difficulties in estimating $R0$. As the infection moves through the population more quickly, then the period in which $R0$ is not impacted by the drop off in susceptibles becomes very short, and the estimated value becomes biased down as falling S starts driving the result. Figure 6 connects the simulation model more firmly to the theoretical structure of the SIR model. It also shows that even for the diffuse case, which is taken as a kind of representation for an SIR like model, the marginal impact of shutting down motion on infected types can be very large.

These initial experiments are used to set some baseline parameters for future runs. In particular, the a_{max} value will often be set to 10 days. Given the uniform distribution of this value between 0 and a_max this yields a mean presymptomatic period of 5 days which seems within a reasonable range for what has been discussed for COVID-19. It yields a fairly extensive average epidemic in the spatial case where about 50 percent of the population will eventually be infected and recovered. That fraction is closer to 90 percent for the diffuse case.

3.2 Long range travel

The next figures explore another key structural parameter in the model. This is the probability of agents moving longer than one step in space, or "long range travel". This can also be thought of as representing air travel, since agents will simply reappear at the new space without touching anyone in between. Instead of randomly moving 1 unit, they now move randomly $[0, d_{maxjump}]$ with a given probability. The probability of jumping will be swept in the next set of figures, but the $d_{maxjump}$ will be fixed at 25.[10] The presymptomatic maximum, a_{max}, will be set to 5. This is done since this is a region where in the spatial model outbreaks are very small. It will be interesting to see if long range travel can change this situation.

Figure 7 presents the results for increasing probabilities of long range travel for the spatial and diffuse models holding other parameters fixed. It repeats the presentation of the severity summary statistics: recovered, max infected, and $R0$. The upper panel shows a steady increase in the severity of the epidemic as long range travel is increased for the spatial model. More motion leads to more spreading as should be expected. In the diffuse/SIR model there is no change, since this model is already scrambling agents randomly across space. Also, for large amounts of travel the models appear indistinguishable. Similar patterns appear in the next two panels for max infected, and R0, respectively. The severity of the disease outbreaks increases in the spatial model and it steadily approaches the values for the diffuse case.

Figure 8 repeats the experiments for the two time series measures. Results are similar to the previous figure. Duration and asymmetry of outbreaks falls for the spatial model as the long range travel probability increases. It steadily converges to the diffuse model, indicating that frequent long range travel means the SIR framework would provide a reasonable approximation to what is going on.

The results presented in figures 7 and 8 are close to what should be expected. The amount of agent spatial mixing increases as long range travel increases. Are there any more subtle messages in these figures? One is that for compartmentalized SRI models, as long as it is reasonable that agents travel freely at distances on the order of magnitude of the compartment, then these models can give reasonable predictions. A second, more speculative result comes from the fact that the increase in recovered fractions increases very quickly for the spatial model. For only 5 percent long range travel probability, the overall recovered fraction moves to near 50 percent, equivalent to runs with $a_{max} = 10$ without travel. Since long range travel can be part of the overall policy tool kit, this suggests that trying hard to eliminate this travel has been a reasonable thing to do.

[10] The spatial grid is 101 square, so a jump of 25 is significant.

3.3 Motion restrictions

The previous experiments could be thought of as explorations of deep structural parameters of both the disease progression, and agent motion behavior. The remaining experiments are directly connected to policy tools. They explore agent movement, testing, and vaccination policies.

Most policies aimed at slowing the spread of disease are directed at slowing interactions, and infections. In the following experiments agents are slowed down by limiting their movements probabilistically. Instead of moving every day they only move with a given probability less than 1. This can be thought of as a kind of smooth lockdown. Setting the probability to 1 repeats the dynamics of the original model, and setting it to zero gives a model with no motion.

Figure 9 repeats the three severity measures as movement is gradually restricted. In all cases the maximum presymtomatic period a_{max} is fixed at 10, and the long range travel probability is set to 0. This is done to compare with previous results, and to draw the starkest differences between the two types of models. In all cases movement restrictions reduce the outbreak severity. Differences between the spatial and diffuse/SIR models are again apparent. Obviously, the levels of severity are smaller for the spatial model in all cases. Most interesting is the fact that reductions in relative severity are greatest in the spatial model. For a modest reduction in movement to 0.8 the recovered level drops by nearly 50 percent which would translate into a fatality reduction of similar magnitude. The proportionate reduction for the diffuse model is around 10 percent. To gain a 50 percent reduction for that model one would need the movement probability to fall to 0.5. Similar results are seen for the Max infected levels. However, Max $R0$ for the two models tracks very closely. These results ask the interesting question of whether the payoffs to small reductions in activity could be underestimated in a SIR-like framework. The fundamental importance of space and distance in the spatial model must be critical for this result.

In figure 10 the impact of movement restrictions on time series features are presented. The results show that for the spatial model there is an actual increase in the length of the epidemic for small reductions in probability, but quickly the Duration starts to fall as more restrictive movement restrictions are put in place. A similar pattern is observed for the Diffuse/SIR model, but the maximum length is now pushed over to much smaller probabilities and larger movement restrictions. It is also interesting to note that over a large range of movement probabilities the two types of models generate opposite changes in Duration. Increasing restrictions (lowering probabilities) reduces the duration for the spatial model, and increases the duration for the diffuse model. Although somewhat counter intuitive, the results in the diffuse case are

probably the result of smaller infections leading to relatively long, but small, epidemics in space and time.

The lower panel of figure 10 shows no evidence for much asymmetry in the Diffuse/SIR model. The Spatial model is very different displaying its usual asymmetry for movement probabilities near one. Somewhat surprisingly, these drop very quickly as the probability falls to near 0.6. This may indicate that long right tails which are caused by new pockets of the disease starting up in space cannot take hold since agents are not moving enough.

3.4 Testing and contact tracing

The first set of simulations showed the importance of the presymptomatic period for the spread of the disease. Infectious agents moving around as they would in their normal lives are critical in the dynamics for the given parameters. In real epidemics, and in particular COVID-19, this suggests the crucial role that testing can play. This section explores the dual policies of testing and contact tracing. They are done in a highly stylized fashion. First, agents can now be tested with a given probability. Testing is completely accurate (no errors), and agents are assumed to follow quarantine instructions. A positive test will lock them down for a fixed length of 15 days, which will take them well into the infectious (symptomatic) phase when they lock down anyway. This is a policy to find and stop motion of anyone who could spread the disease. As previously described, contact tracing will also be combined with testing. Recent agent contacts, $N_{contact} = 10$, will also be quarantined on discovery of a positive test result.

Figure 11 presents results for the spatial model only. First, it is clear that testing is an effective means of shutting down the disease. With 100 percent daily testing there is no outbreak. This makes sense since in the overall picture of the model this would be nearly equivalent to shutting down the presymptomatic period. All movement ceases once an agent is infected and infectious. Quantitatively the model does require a lot of testing. Reducing test levels to 0.2, or about once every 5 days, moves the recovered fraction back to 0.25 which is a fairly large outbreak.

The most dramatic result in figure 11 is that contact tracing appears to have such a small impact. For all the measures there is a small level of improvement, but for most test probabilities the improvement is small. For example, at a test probability of 0.2 contact tracing reduces the recovered level from 0.25 to 0.20. Not a very dramatic improvement, especially when compared to the big change one would get by increasing testing probability to 0.5. This is a very interesting result, since in the real world policy context contact tracing is viewed to be a key tool. The lower panel shows a reduction in $R0$ as testing increases, and in this

figure contact tracing appears to have almost zero impact.

The time series features are again explored in figure 12. The duration of the disease and its right tail asymmetry all fall as testing is increased. Major reductions in the length of the epidemic require a fairly extensive testing protocol with a test probability greater than 0.5. Once again there is no impact from the addition of contact tracing.

Figures 11 and 12 only displayed features for the pure spatial model. The diffuse model will now be shown in figures 13 and 14. Figure 13 shows that the severity is again decreasing in the amount of testing. As with the previous figures there is some indication that testing must be done very frequently to have a big impact. Even at a testing level of every other day (0.5) the outbreak is still pretty extensive with well over half the population infected (recovered) in the absence of contact tracing. In contrast to the findings from the pure spatial model, contact tracing now shows an impact. It reduces both total recovered and max infected levels significantly for a range of test probabilities. This is especially true for the Max infected values, less so for the total recovered part of the disease. Finally, there is only a small impact of contact tracing on estimated $R0$.

The time series results are reported in figure 14. Again, the important result is that contact tracing now shows a large impact on the dynamic progression of the disease. Contact tracing leads to longer durations (except for very large testing levels), and an interesting asymmetry which is generally not seen in the diffuse model cases. It is important to remember that contact tracing is reducing the overall severity of the outbreaks, but they appear to fade away more slowly. It also should be noted that increasing testing increases the duration of the epidemic for a range of testing probabilities (< 0.4). This is exactly the reverse of the result for the pure spatial model seen in figure 12.

3.5 Vaccination

The final policy explored is vaccinations. Vaccination will be modeled by making some initial fraction of the population recovered. This part of the population is then not sick and cannot carry the disease or infect anyone else. The fraction of the population that is susceptible at the start is smaller since it is just (1-vaccinated).

Figure 15 presents the impact of increasing vaccinations on the severity of the outbreak. For both the spatial and diffuse models vaccinations work to wipe out the epidemic. To better assess the severity of the disease, the fraction recovered is normalized by the fraction initially not vaccinated. In other words

recovered is now measured as,

$$\frac{f_R - f_V}{1 - f_V}. \tag{6}$$

This measures the fraction who moved to recovered from susceptible (were not vaccinated), normalized by the original susceptible fraction. It is just looking at numbers for the unvaccinated as a separate population from the rest. Also, the maximum infected fraction is also normalized by the initial unvaccinated population,

$$\frac{\max I}{1 - f_V}. \tag{7}$$

For both cases a vaccination rate of a little over 0.6 (60 percent) pretty much eliminates spread of the disease. There are again major differences between the two models. The spatial model sees stronger reductions in infection for much smaller vaccination levels. For example, at only a 40 percent vaccination rate the spatial model shows a near elimination of the epidemic. For all cases the value of maximum estimated $R0$ steadily drives toward zero as the level of vaccination increases. For this case, the two models generate relatively similar numbers.

Figure 16 looks at the time series properties as the vaccine levels are changed. For the spatial model there is a strong reduction in duration, and both models converge to a very short time period corresponding to near elimination of the disease. This is repeated in the lower panel for asymmetry. Vaccination levels have no impact on asymmetry for the diffuse case.

4 Conclusions

This paper presented a stylized model of infectious disease spread. Its simple structure puts it much closer to the SIR model in spirit, and it should be viewed more as a computational thought experiment than a quantitative model for predicting disease spread, or making detailed policy recommendations. There are still several interesting aspects to the entire exercise which will be interesting to eventually see how well they hold up empirically as more data is gathered.

First, the spatial model is able to generate much richer time series than a model which mostly ignores the spatial dimension. Time series can generate near cyclic behavior and long persistent tails. It is important that these features may be a deep aspect of how the disease evolves through space, and not necessarily related to policy feedback dynamics which are often conjectured. The second key result is the powerful impact of asymptomatic spreading. This is already starting to get acknowledged by much of the COVID-

19 research, but it is important to see it here. To researchers it is a reminder that to fully understand the nature of this disease we need to work hard to lock down the parameters for asymptomatic periods where individuals are infectious and moving through space.

Several policies are explored, and must be viewed with some care since they are not aligned with actual data. Restricting movements and contacts behaves as it should, but the impact is stronger for the fully spatial model. Given the abstract nature of this policy, it may be the most difficult to take to the real world. Testing also is a powerful tool for slowing down or eradicating the disease. However, the results suggest some possible caution for testing frequencies on the order of once per week. This may not be enough. Probably one of the most interesting results on testing is related to contact tracing. For the fully spatial model its impact was marginal. At the moment it is not clear why this is true, or how its implementation can be improved. For the fully diffuse model contact tracing behaves in a manor more consistent with the common finding that it is very important. Finally, vaccinations behave as they should in eliminating the disease, but the two models again give a different picture. In a fully spatial framework the threshold necessary for a successful vaccination can be as low as 40 percent, where the equivalent for the diffuse model is closer to 60 to 70 percent. This is a pretty optimistic result in terms of vaccination programs.

It is very tempting to make this model more complex and to get it closer to data calibration. In many aspects this should be avoided since it might quickly lose its easy interpretation, and usefulness as a thought experiment. However, there are probably several parts which could be improved. As more detailed information on disease progression becomes available the length of time spent in presymptomatic and symptomatic stages should be updated. This will be an easy change. A second important improvement would be to make movement in space more realistic. This would probably not involve layering an exact geometry on this system, but it it should be possible to better align with some generic information on how people are moving around. These data sets are steadily becoming more readily available through cell phone information. A final major limitation of the model in relation to most other models is the uniformity of contacts. This is where more serious models bring known contact network structure in to better simulate probabilities and interactions. It is not clear if the lack of this structure is a curse or a blessing. To maintain the thought experiment side of this model, and the endogenous nature of space, network structure would need to be brought on board in a very simple fashion.

This paper has presented an agent-based computational model for infectious disease spread. It is highly stylized, and relatively simple, yet still captures enough elements of realism to interest policy makers. It has shown the key importance of combining agent mobility in a rich, but simple space, along with asymp-

tomatic infections. All three parts are important contributors to the dynamics of the disease and its response to policy tools.

References

Acemoglu, D., Chernozhukov, V., Werning, I. & Whinston, M. D. (2020), Optimal targeted lockdowns in a mult-group, SIR model, Technical Report 27102, NBER.

Aleta, A. et al. (2020), 'Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19', *Nature Human Behavior* **4**, 964–971.

Arons et al. (2020), 'Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility', *New England Journal of Medicine* **382**(22), 2081–2090.

Avery, C., Bossert, W., Clark, A., Ellison, G. & Ellison, S. F. (2020), 'An economist's guide to epidemiology models of infectious disease', *Journal of Economic Perspectives* **34**(4), 79–104.

Balcan, D., Goncalves, B., Hu, H., Ramasco, J., Colizza, V. & Vespignani, A. (2010), 'Modeling the spatial spread of infectious diseases: The GLobal Epedimic and Mobility computational model', *Journal of Computational Science* **1**, 132–145.

Bertuzzo, E., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I. & Rinaldo, A. (2009), 'On spatially explicit models of cholera epdemics', *Jounal of the Royal Society Interface* **7**(43), 1–13.

Bloom, D. E., Kuhm, M. & Prettner, K. (2020), Modern infectious diseases: Macroeconomic impacts and policy responses, Technical Report 27757, NBER.

Brearcliff, D. (2020), COVID-19 US model, Technical report, George Mason University.

Burke, D. S., Epstein, J. M., Cummings, D. A. T., Parker, J. I., Cline, K. C., Singa, R. M. & Chakravarty, S. (2006), 'Individual-based computational modeling of smallpox epidemic control strategies', *Academic Emergency Medicine* **13**(11), 1142–1149.

Carpenter, R. G. (1974), 'Goegraphical and temporal development of plagues', *Nature* **250**, 726–729.

Charbonneau, P. (2017), *Natural Complexity*, Primers in Complex Systems, Princeton University Press.

Epstein, J. M., Parker, J., Cummings, D. & Hammond, R. A. (2008), 'Coupled contagion dynamics of fear and disease: Mathematical and computational explorations', *PLOS One* **3**(12), 2.

Gandhi, M., Yokoe, D. S. & Havlir, D. V. (2020), 'Asymptomatic transmission, the achilles' heel of current strategies to control COVID-19', *The New England Journal of Medicine* **382**(22), 2158–2160.

Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R. & Rinaldo, A. (2020), 'Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures', *Proceedings of the National Academic of Science* **117**(19), 10484–10491.

Getz, W. M., Salter, R. & Mgbara, W. (2019), 'Adequacy of SEIR models when epidemics have spatial structure: Ebola in Sierra Leone', *Philisophical Transactions of the Royal Society B* .

He, X. et al. (2020), 'Temporal dynamics in viral shedding and transmissibility of COVID-19', *Nature Medicine* **26**, 672–675.

Kermack, W. O. & McKendrick, A. G. (1927), 'Contributions to the mathematical theory of epidemics', *Proceedings of the Royal Society* **115A**, 700–721.

Lang, J. C., Sterck, H. D., Kaiser, J. L. & Miller, J. C. (2018), 'Analytic models for SIR disease spread on random spatial networks', *Journal of Complex Networks* **6**, 948970.

Lloyd, A. L. & May, R. M. (1996), 'Spatial heterogeneity in epidemic models', *Journal of Theoretical Biology* **179**, 1–11.

Ludvigson, S. C., Ma, S. & Ng, S. (2020), COVID-19 and the macroeconomic effects of costly disasters, Technical Report 26987, NBER, Cambridge, MA.

Munchi, J., Roy, I. & Balasubramanian, G. (2020), Spatiotemporal dynamics in demography-sensitive disease transmission: COVID-19 spread in NY as a case study, Technical report, Lehigh University.

O'Sullivan, D., Gahegan, M., Exeter, D. J. & Adams, B. (2020), 'Spatially explicit models for exploring COVID-19 lockdown strategies', *Transactions in GIS* **24**(4), 967–1000.

Scherbina, A. (2021), Assessing the optimality of a COVID lockdown in the United States, Technical report, Brandeis International Business School.

Siettos, C. I. & Russo, L. (2013), 'Mathematical modeling of infectious disease dynamics', *Virulence* **4**(4), 295–306.

Vespignani, A., Tian, H., Dye, C., Lloyde-Smith, J. O., Eggo, R. M., Shrestha, M., Scarpino, S. V., Gutierrez, B., Kraimer, M. U. G., Wu, J., Leung, K. & Leung, G. M. (2020), 'Modeling COVID-19', *Nature Reviews: Physics* .

Vrugt, M. t., Bickmann, J. & Wittkowski, R. (2020), 'Epidemic spreading modeled via dynamic density functional theory', *Nature Communications* **11**, 5576.

Wilensky, U. (1999), Netlogo, Technical report, Center for Connected Learning, and Computer-based Modeling, Northwestern University.

Wilensky, U. & Rand, W. (2015), *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with Netlogo*, MIT Press.

Willem, L., Verelst, F., Bilcke, J., Hens, N. & Beutels, P. (2017), 'Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006-2015)', *BMC Infectious Diseases* **17**, 612.

Wolfram, C. (2020), 'An agent-based model of covid-19', *Complex Systems* **29**, 87–105.

Table 1: Parameter values

Parameter	Value
a_{max} (asymptomatic)	Often 10, range = $U[0, a_{max}]$
$d_{maxjump}$	25
p_{infect}	1
Infection range (days)	$U[10, 14]$
$R0$ sample size	100
$N_{contact}$ Contact list (agents)	10
Grid	101 square with wrapping

Basic model parameters, and common settings.

Figure 1: Netlogo interface

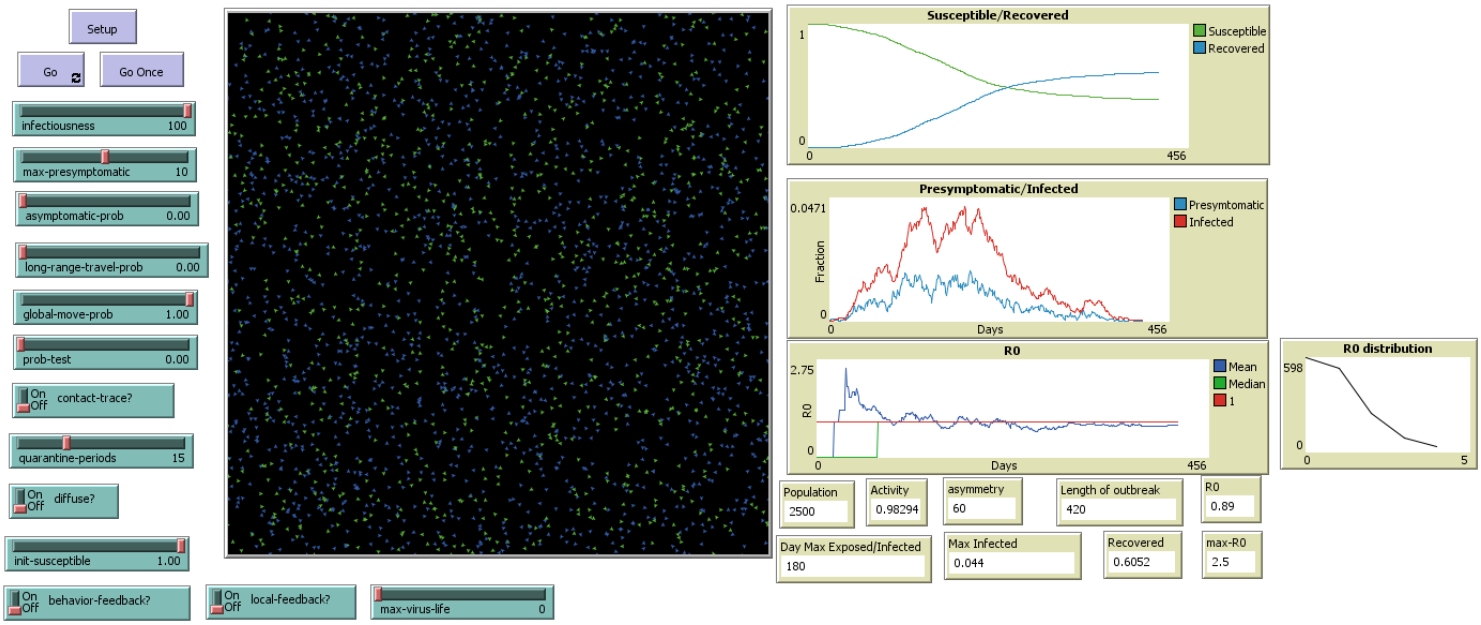

Figure 2: Single run: Spatial model

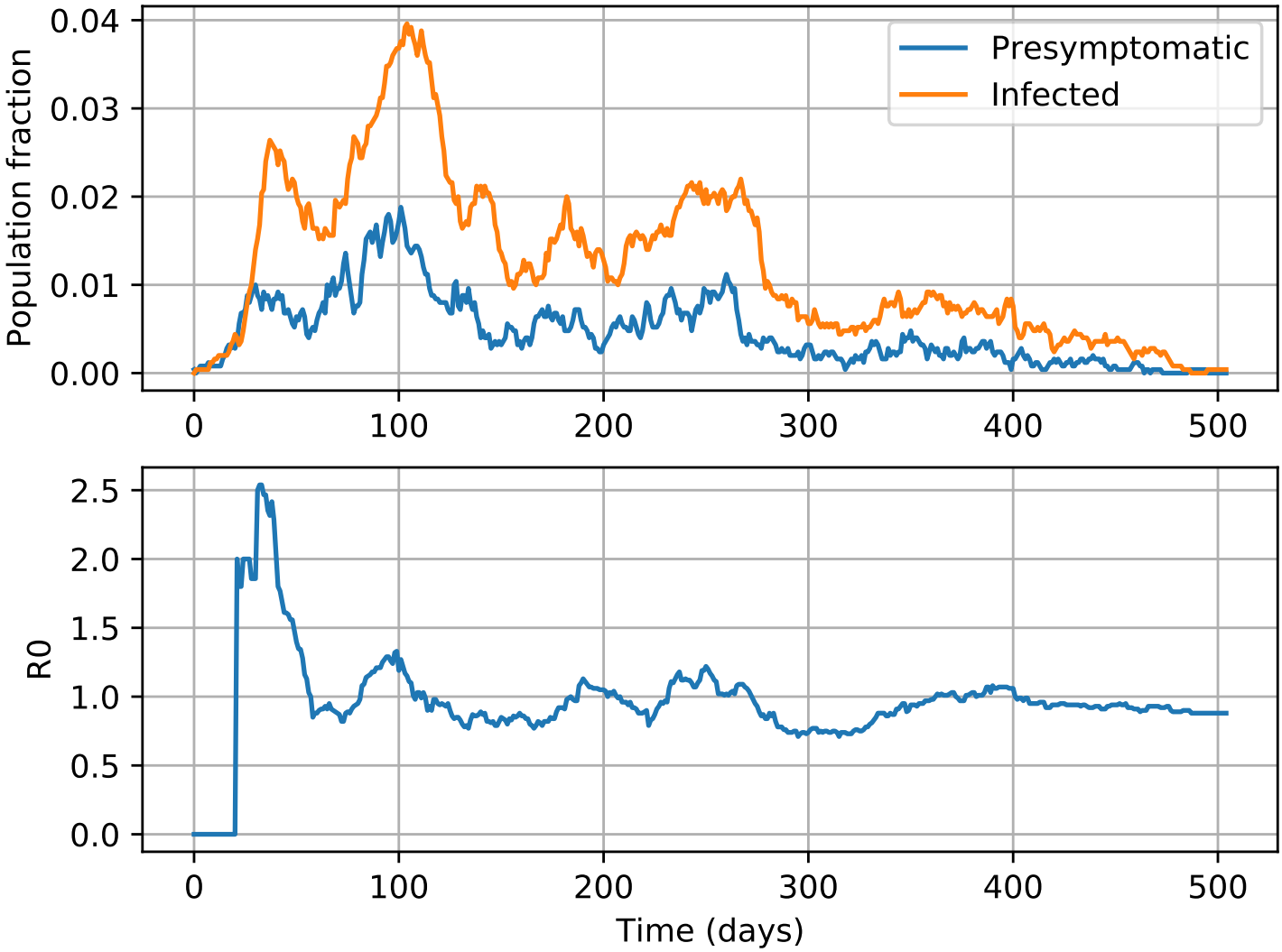

Figure 3: Single run: Diffuse model

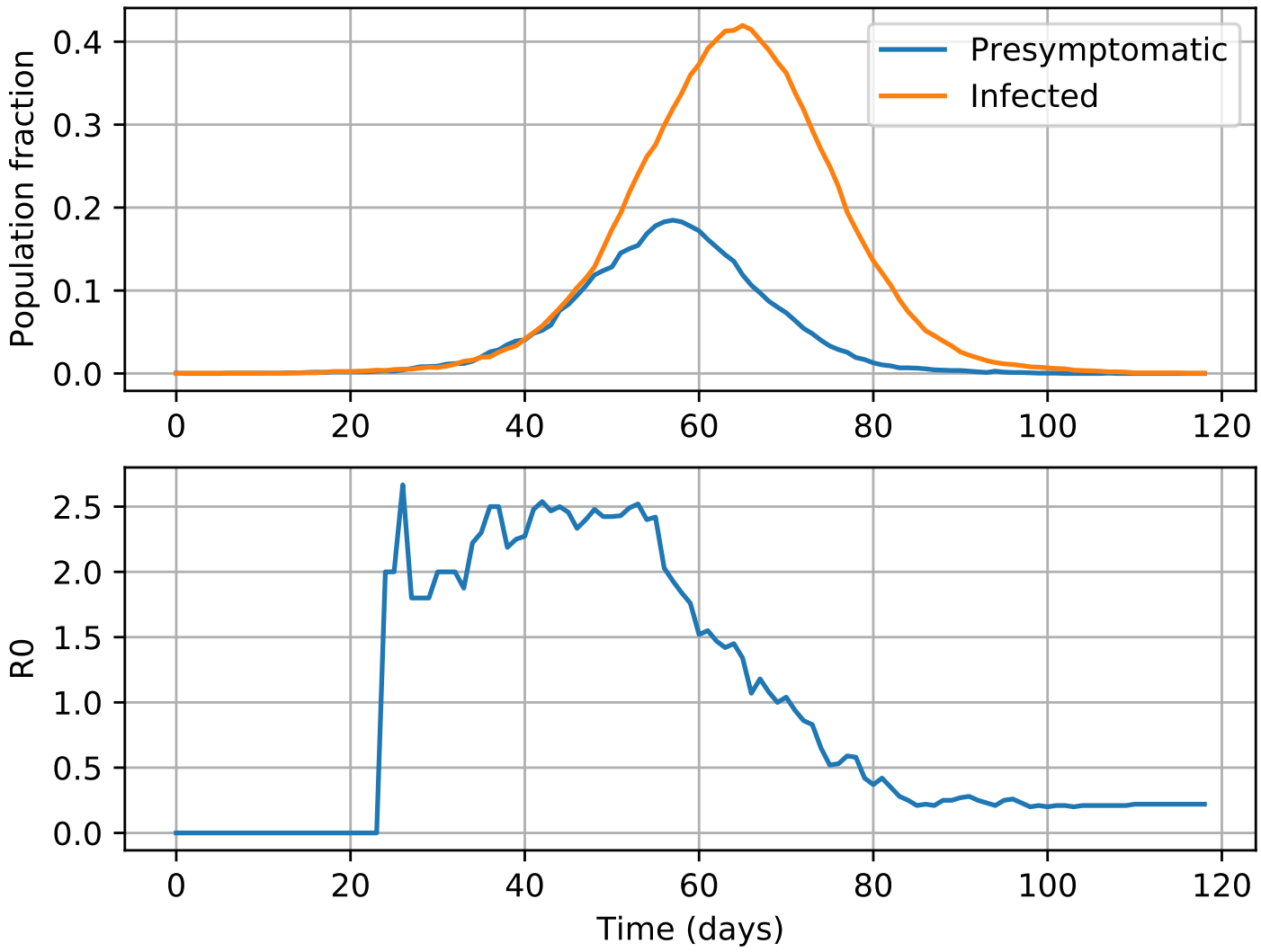

Figure 4: Presymptomatic length: Severity

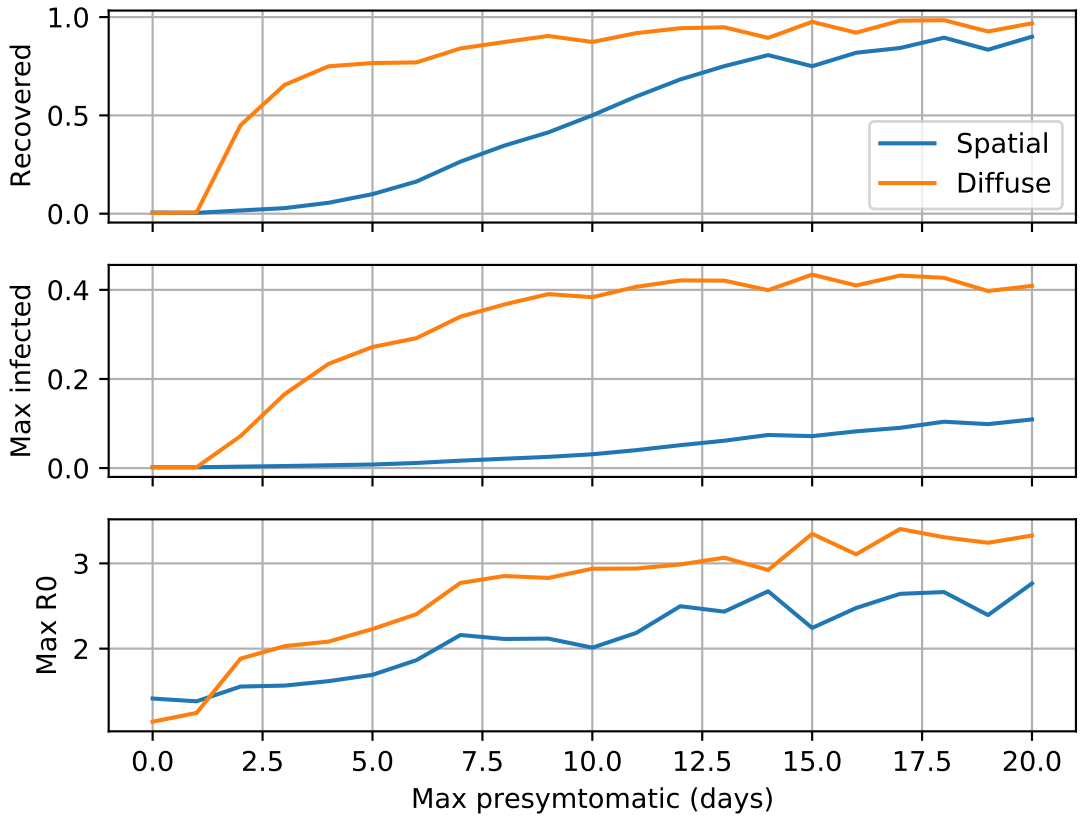

Figure 5: Presymptomatic length: Shape

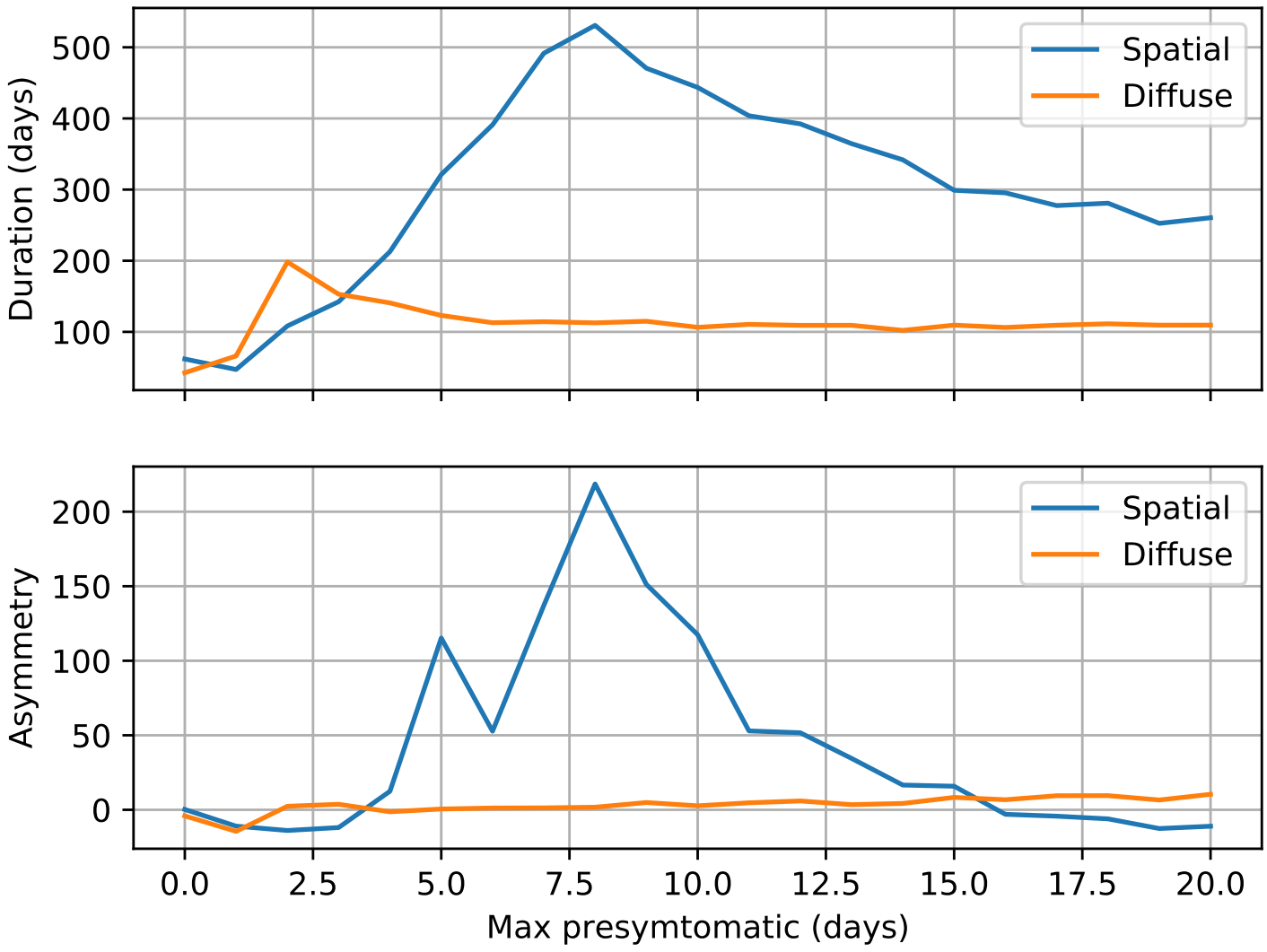

Figure 6: Full SIR model comparison

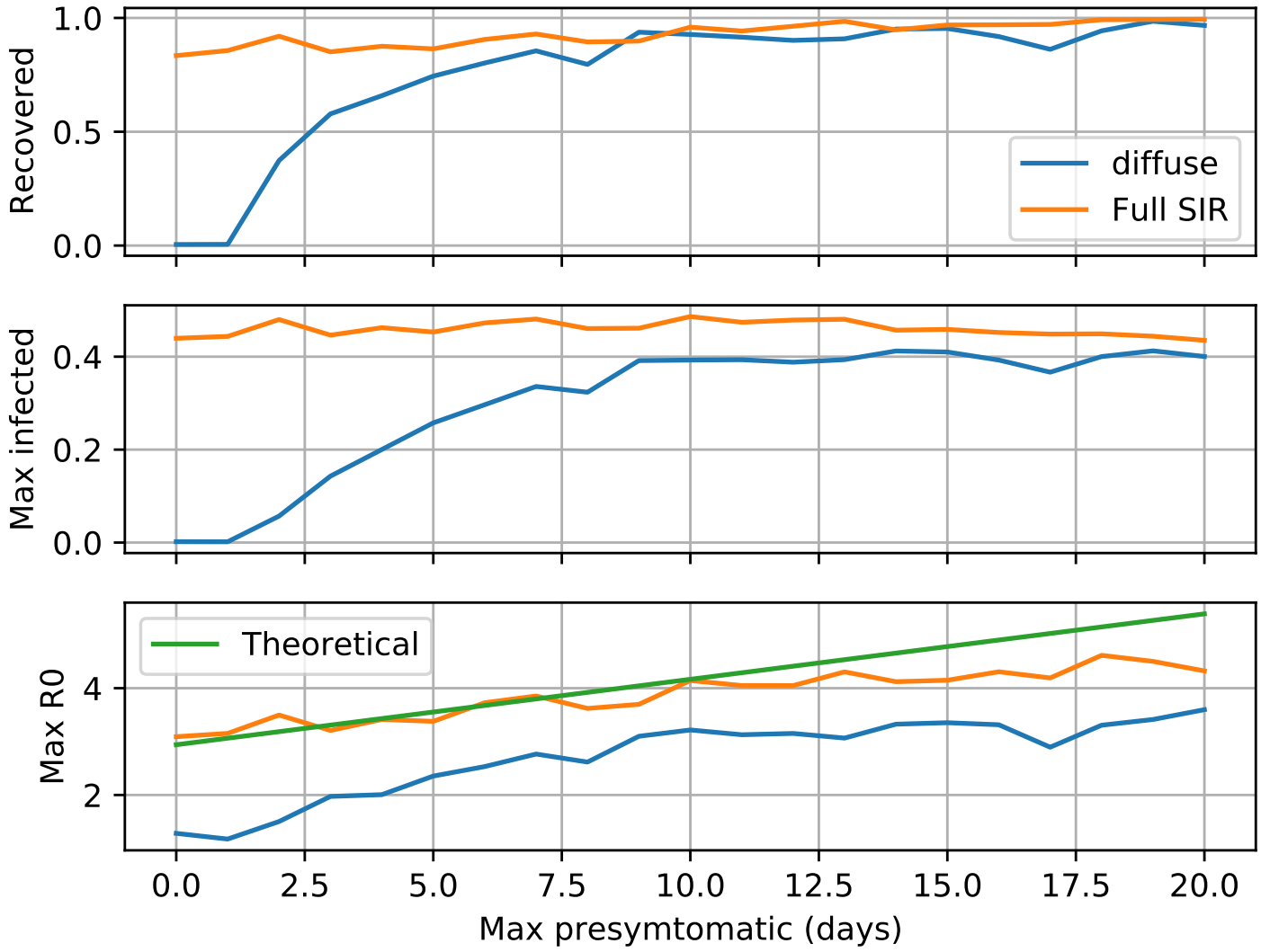

Figure 7: Long range travel probability: Severity

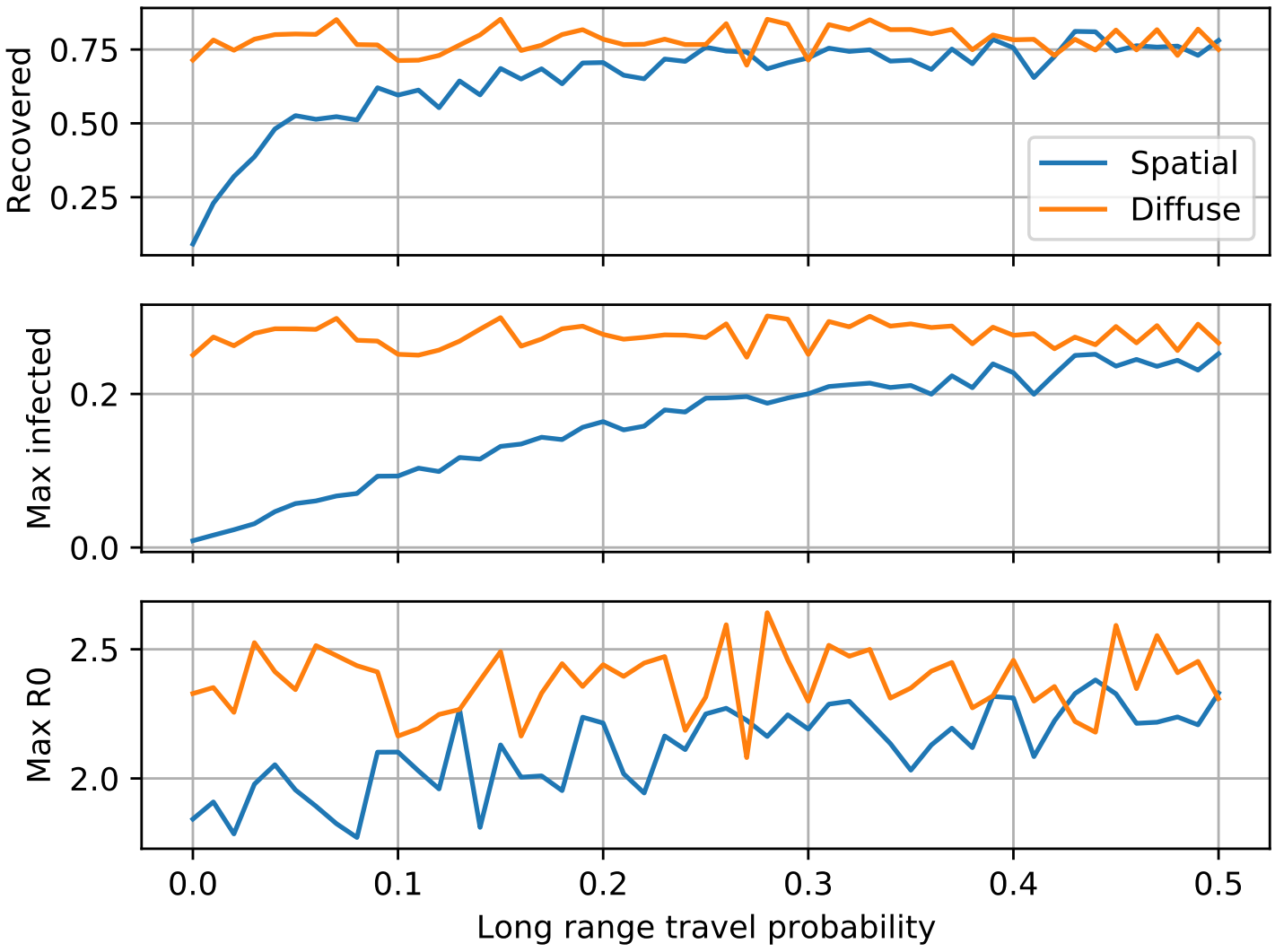

Figure 8: Long range travel probability: Shape

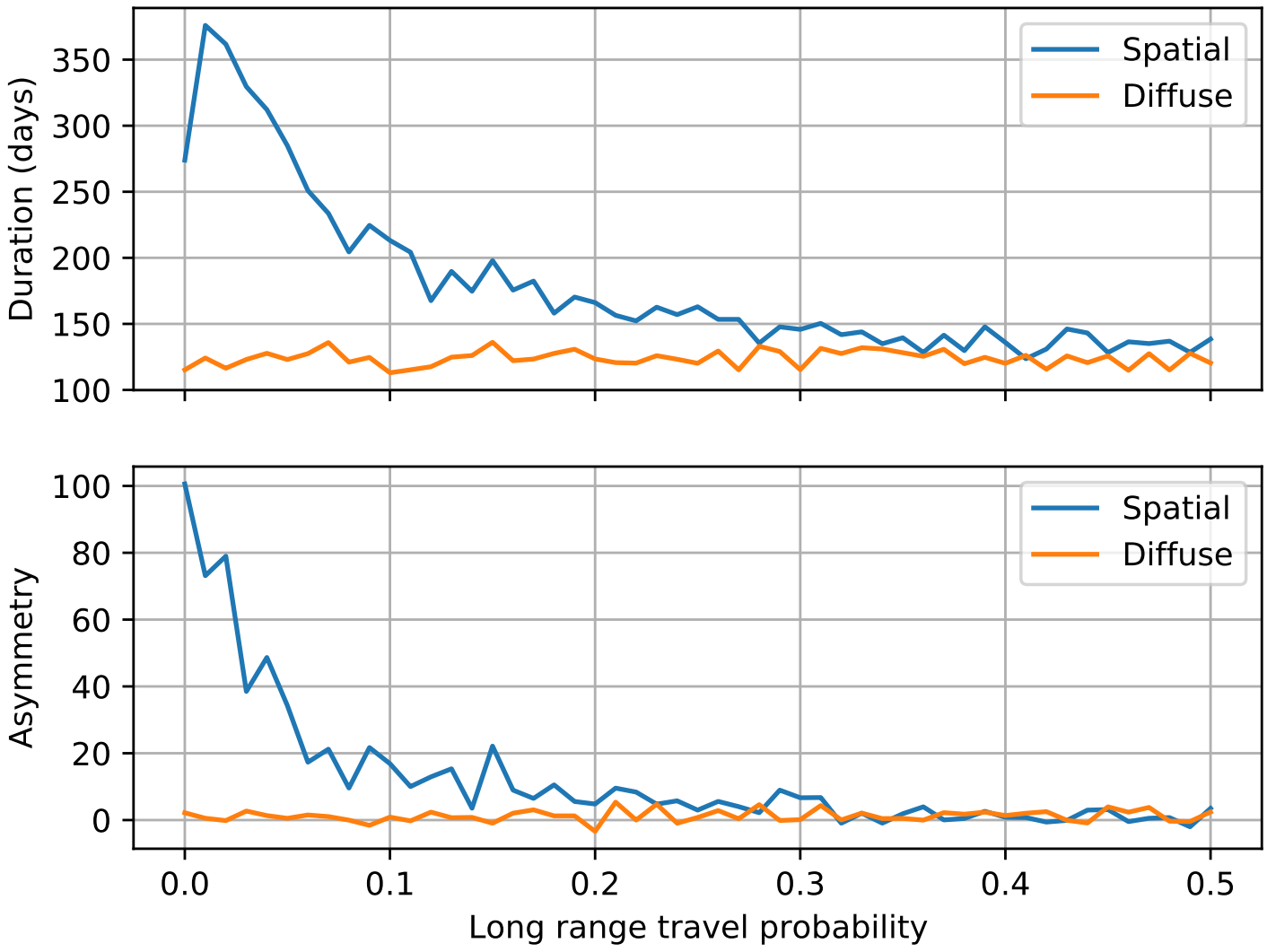

Figure 9: Movement probability: Severity

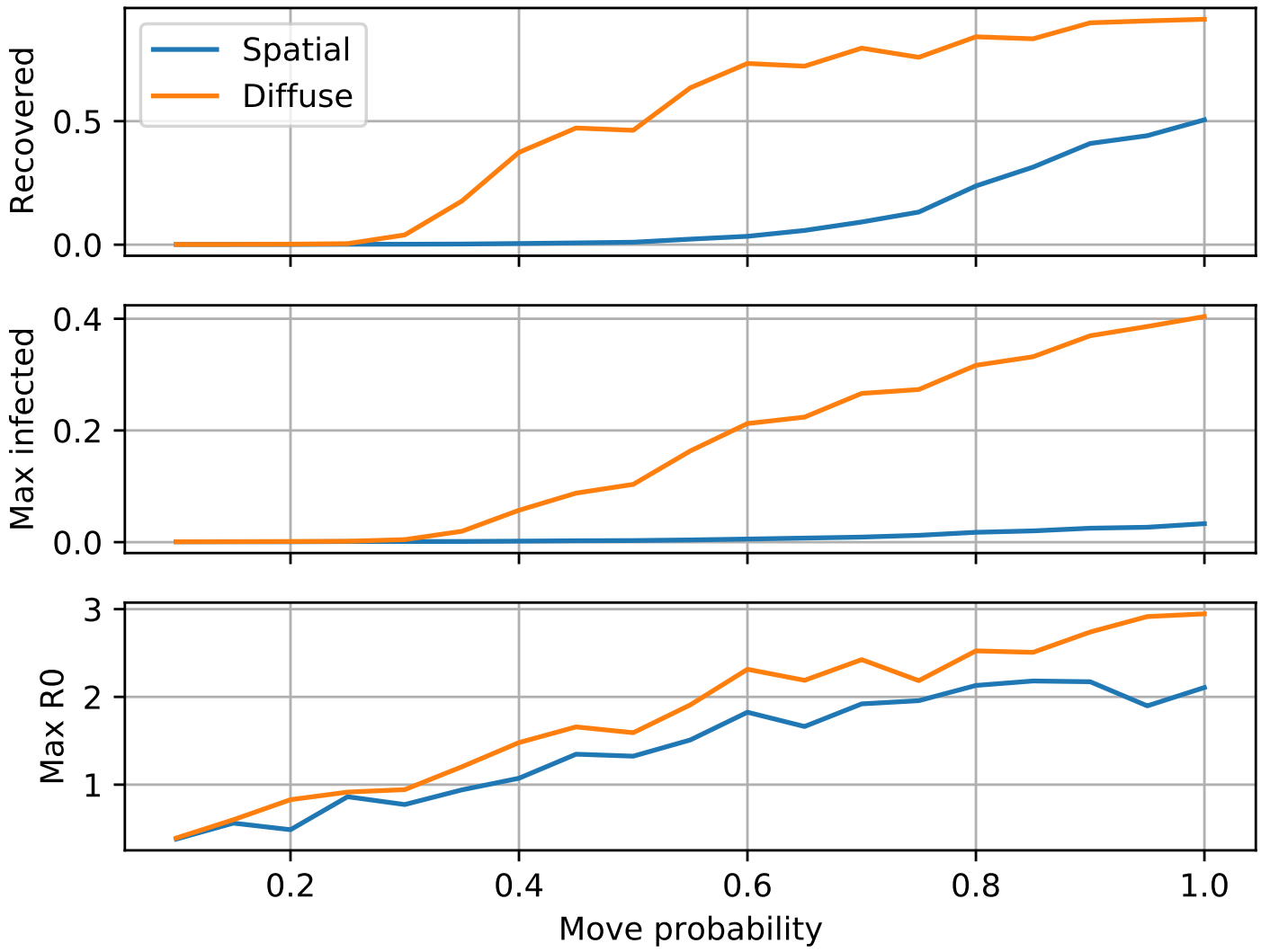

Figure 10: Movement probability: Shape

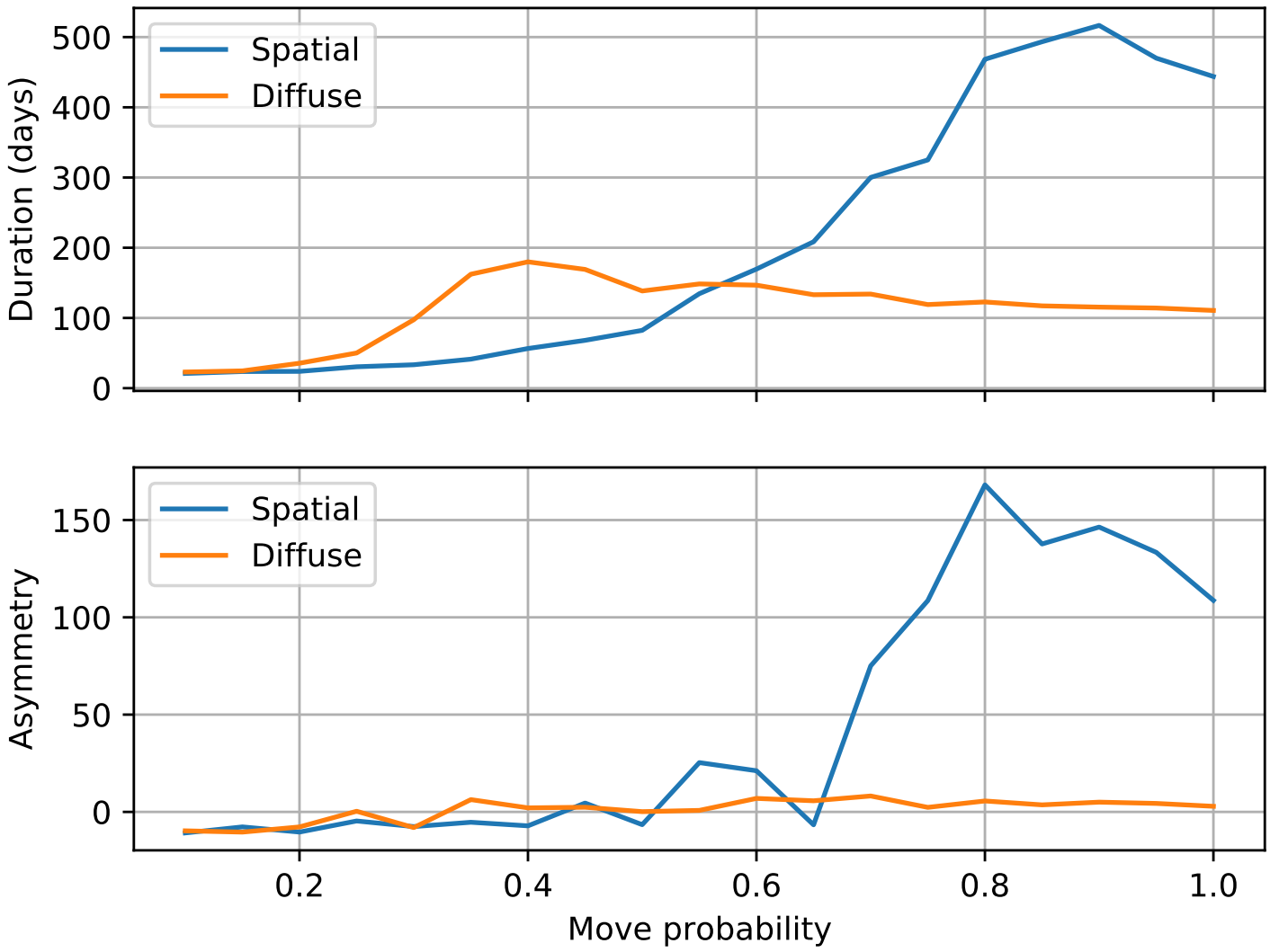

Figure 11: Testing probability (spatial model): Severity

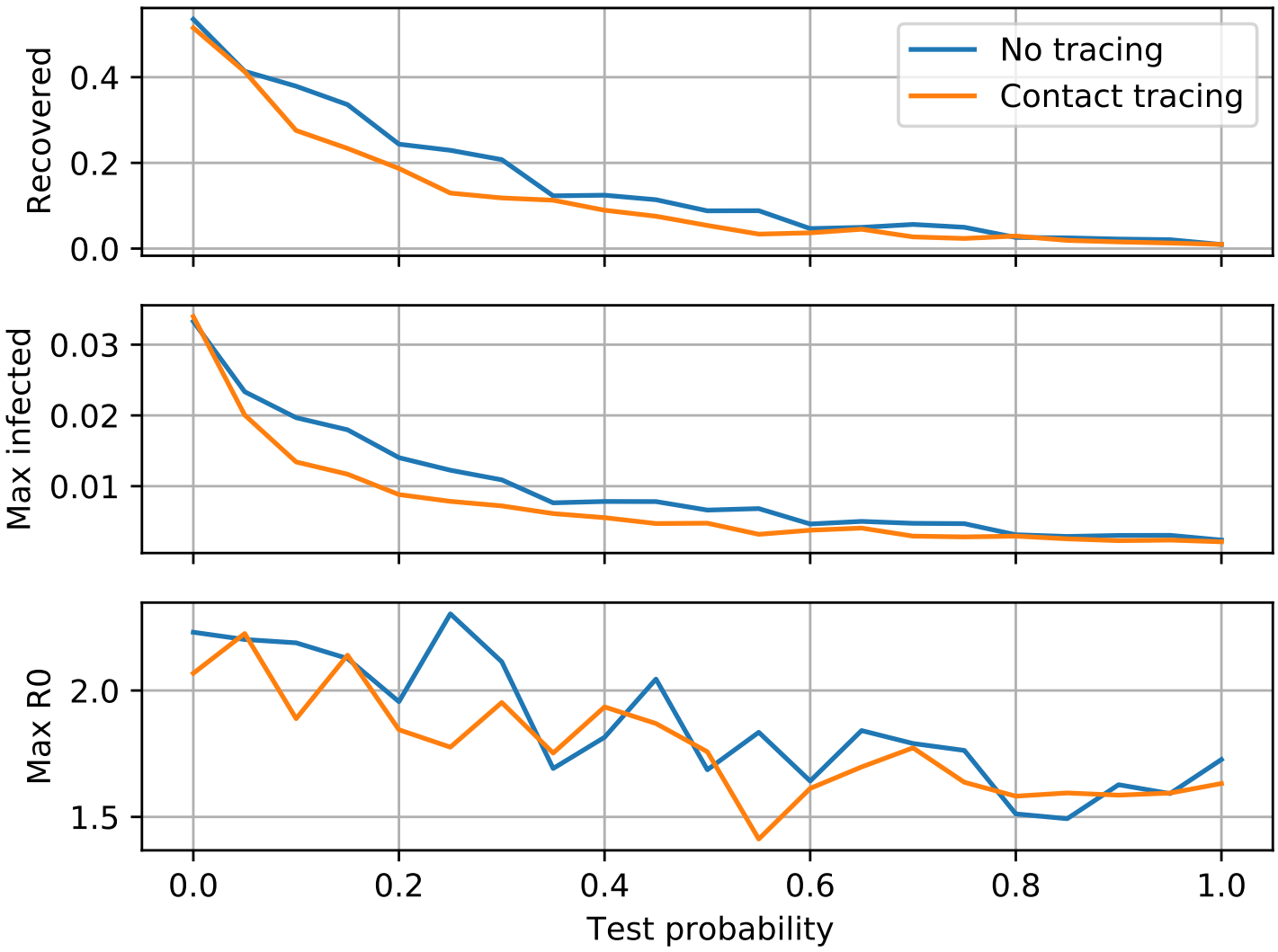

Figure 12: Testing probability (spatial model): Shape

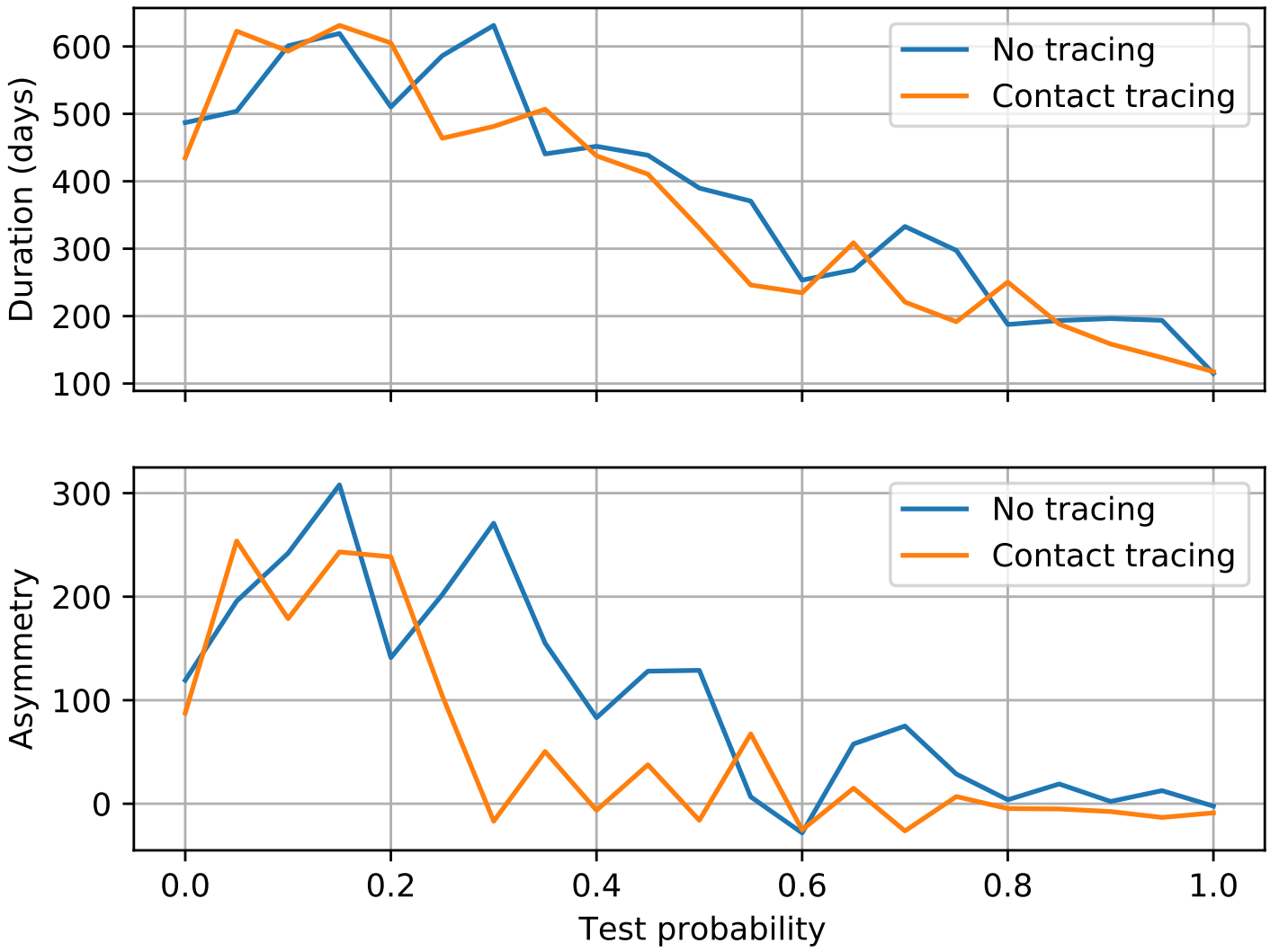

Figure 13: Testing probability (diffuse model): Severity

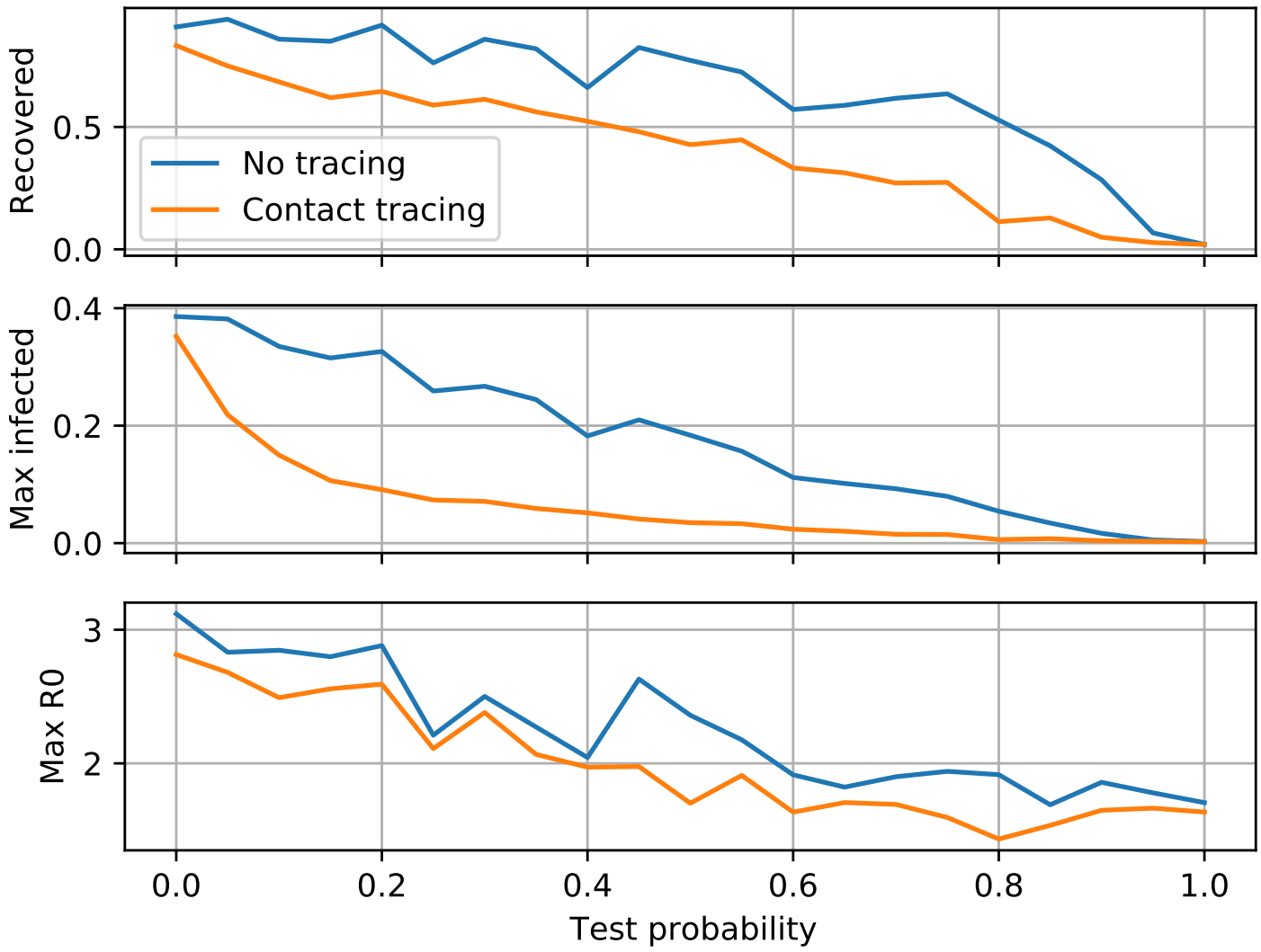

Figure 14: Testing probability (diffuse model): Shape

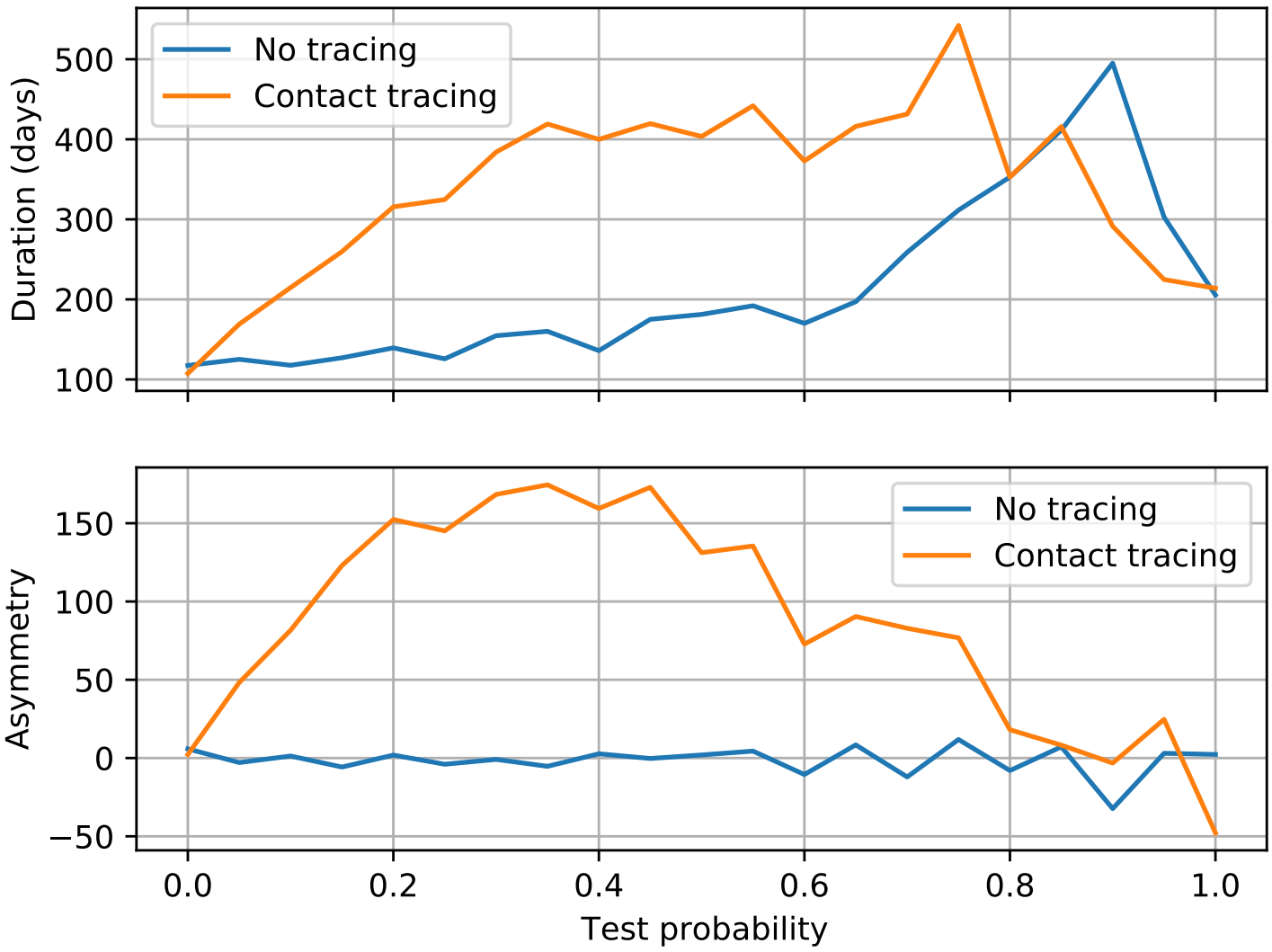

Figure 15: Vaccinations : Severity

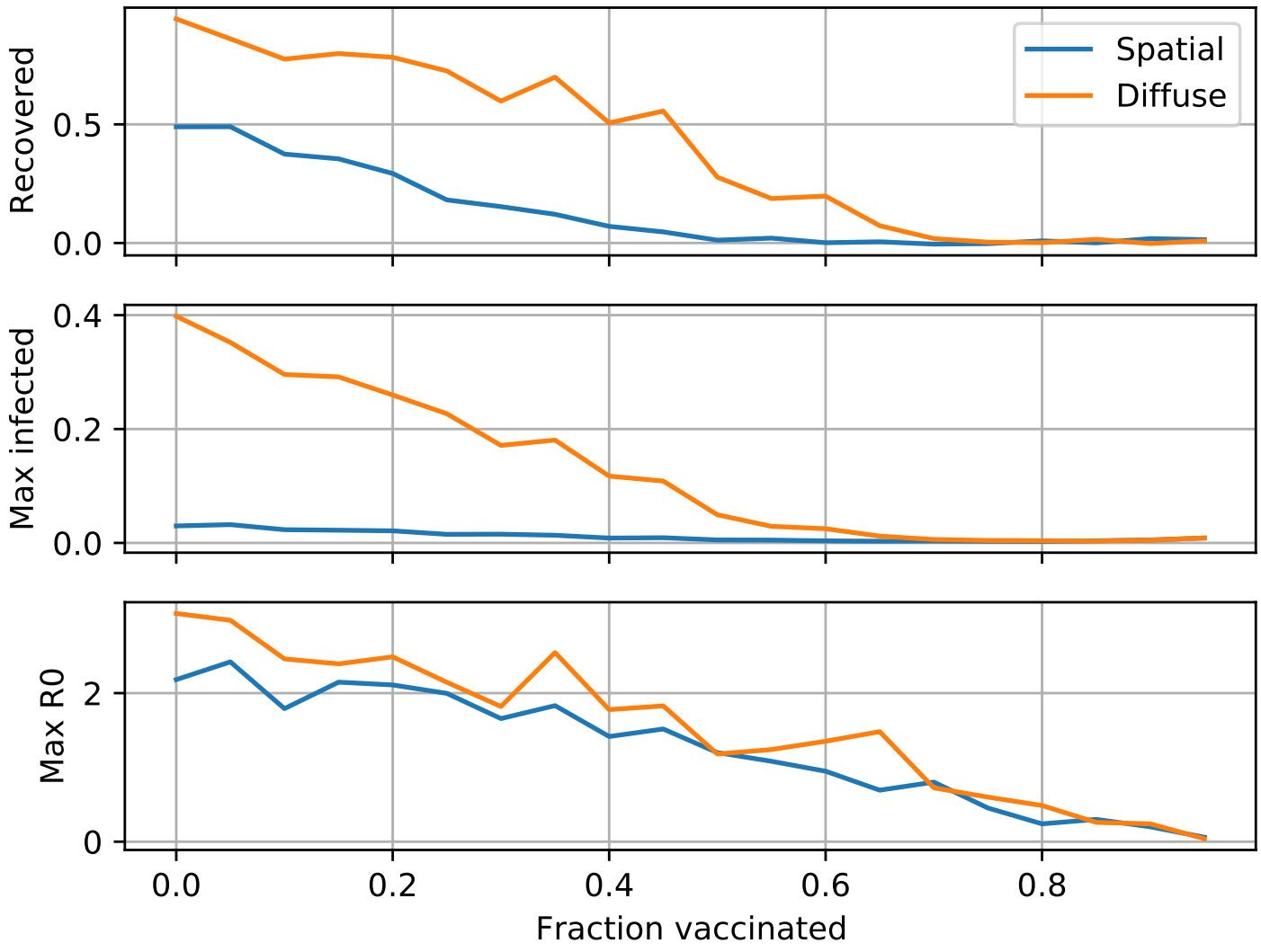

Figure 16: Vaccinations: Shape