Simulating the effect of bidirectional contact tracing on the spread of infectious diseases

Master Thesis by

Jim Vollebregt

Studentnumber: 5683548

Supervisors

Martin Bootsma
Daily supervisor

Paul Zegeling
Second reader

Mathematical Sciences
Graduate School of Natural Sciences
Utrecht University
June 2021

Utrecht University

Abstract

The impact of the COVID-19 pandemic can hardly be overstated. In February 2020, the World Health Organisation recommended to prioritize extensive testing and contact tracing to contain the virus. Through bidirectional contact tracing, we can not only find the people infected by an infectious individual, but the infector of the individual as well. Theoretically, this makes bidirectional contact tracing an effective countermeasure for diseases where a proportion of the population is asymptomatic or a super spreader. Due to its recursive nature, it is hard to incorporate the effects of bidirectional contact tracing into existing analytical models for infectious diseases. Instead, we built an event-by-event simulation model. Using this simulation model, we obtain results regarding the effectiveness of bidirectional contact tracing in various scenarios.

Contents

1 Introduction

On January 15, 2020, the World Economic Forum considered *infectious diseases* a quite unlikely risk [27]. Yet, on January 30, 2020, the World Health Organization (WHO) declared the outbreak of the SARS-CoV-2 virus, also known as COVID-19 virus or Coronavirus, a Public Health Emergency of International Concern [18], and a pandemic on 11 March 2020 [19]. As of May 2021, estimates by the WHO show that more then 150 million people worldwide have been infected by COVID-19 resulting in over 3 million deaths [17]. Aside from that, the COVID-19 pandemic has had a major impact on hospital admissions. In the United States, hospital admissions for non-COVID-19 patients were 16% below baseline on average, while admissions for pneumonia and asthma are estimated to have dropped by 40% or more by September 2020 [20]. In the Netherlands, 23% of treatments in the 12 most frequently provided medical specialties were cancelled during the initial wave of the epidemic [21]. Furthermore, general well-being and mental health are affected as well. The review article by Vindegaard and Benros from October 2020 [23] includes 43 studies investigating the mental health of COVID-19 patients, health care workers and the general public that show lower psychological well-being and higher scores of anxiety and depression. In a survey in China from March 2020, 16.5% of people reported experiencing moderate to severe depressive symptoms, 28.8% reported moderate to severe anxiety symptoms, and 8.1% reported moderate to severe stress levels [24]. Lastly, the COVID-19 pandemic will likely have major economic effects [26]. A proportion of COVID-19 infections happens through super spreading events. A list of over 2000 such events has been complied by the London School of Hygiene and Tropical Medicine [11]. In addition, A study from Italy [25] shows that a proportion of over 40% of infections happens through infectious people who have not developed symptoms at the time of infection. Fortunately, vaccination for COVID-19 has begun. As of June 2021, 2.5 billion vaccines have been administered worldwide [32]. Still, another pandemic may hit in the future.

The impact of the COVID-19 pandemic has been reduced due to countermeasures employed by governments. As early as February 2020, the WHO recommended to prioritize extensive testing and contact tracing in order to contain the spread of the virus [16]. Other countermeasures include lockdowns, curfews and partial closing of borders. The latter countermeasures apply to the whole population, while testing and consecutive contact tracing target people who are already infected specifically: An individual diagnosed with COVID-19 is put in isolation, and their contacts are traced and tested as well. This creates a chain reaction where infectious individuals can be found, positively diagnosed, and isolated even before they develop symptoms or enter their most infectious state.

Through contact tracing one may find the people infected by a diagnosed individual (forward contact tracing) as well as the infector (backward contact tracing). The combination of the two we call bidirectional contact tracing [7]. Usually, it takes several days before infected people become infectious themselves. Moreover, people are usually at their peak infectiousness around the time they develop symptoms. Hence, for forward tracing, one only needs to screen contacts in the past few days. For backward tracing on the other hand, one needs to go back several days to find the infector of a diagnosed individual. Therefore, bidirectional contact tracing requires more effort than forward contact tracing only. Contact tracing can either be applied manually or through an application that tracks contacts.

The goal of this project is twofold. First of all, we want to give an indication of how infectious a disease can be for it to still be controlled by contact tracing. We have the following hypothesis, in line with the findings of Bradshaw and others [7, 8]:

> H1: Bidirectional contact tracing as a countermeasure can control the spread of infectious diseases similar to COVID-19.

We have mentioned asymptomatic infections and super spreading events. Asymptomatic individuals do not develop symptoms for the entirety of their infective period. Super spreaders are individuals that are significantly more infectious than average, for example because they have a higher viral load or because they had relatively many contacts during their most infectious period. We want to measure the effects of bidirectional contact tracing on the development of an epidemic where a proportion of the population is asymptomatic or a super spreader. We hypothesize the following:

> H2: Bidirectional contact tracing is an effective measure to counteract a disease where a proportion of the population is asymptomatic or a super spreader.

In all cases, it is clear that bidirectional tracing will at least be as effective as forward tracing only. However, considering the added effort required, we will compare the performance of bidirectional tracing with that of forward tracing only so that we can state when the former has a significant advantage over the latter.

It is almost impossible to update existing analytical frameworks to allow the incorporation of bidirectional contact tracing. Therefore, an event-by-event based simulation model with parameter updating is necessary to obtain the measurements we are interested in. In this thesis we present such a simulation model. Part of the project is dedicated to the optimization of the efficiency of the simulation model in order to increase practical usability. Thereafter we use simulations to test our hypotheses H1 and H2. The simulation model is applicable for epidemics caused by pathogens with epidemiological characteristics similar to those of the COVID-19 virus.

This thesis is structured as follows. In Chapter 2 we familiarize the reader with the field of epidemiology. We focus on theory, research and results in the branch of contact tracing. This Chapter motivates the use of a simulation rather than an analytical framework for the modelling of the effects of bidirectional contact tracing. In Chapter 3 we present the simulation model. We state the assumptions upon which the model is built and what we can measure with it. We also motivate the use of parameter updating. It turns out that this process is similar to the controlling of a delayed system. In Chapter 4 we present the obtained results. In particular, we conclude that epidemics similar to COVID-19 can be controlled through bidirectional contact tracing. Moreover, we show that it is a more effective countermeasure than just forward contact tracing in case there are asymptomatic carriers of the disease. A step-by-step guide of how the simulation model works is given in Appendix A. Moreover, we suggest and implement some adaptations to make it as efficient as possible. Appendix B is a follow-up where we thoroughly outline the process of parameter updating. We motivate the reader to have a look at these appendices if they are interested in the fine details of the simulation model. However, the contents of this thesis can be understood without them.

2 Epidemiology

The current field of mathematical epidemiology is largely an expansion of the contributions of Kermack and McKendrick, o.a. [1] and later publications. The publication by Diekmann and others, [4], gives an excellent overview of [1] with more modern notation, and expands on it as well. Two primary sources for modern epidemiological studies are Diekmann and others' [2] and Murray's [3]. Throughout this Chapter we will use theory from both works. The starting point is always a population of susceptibles where an initial population of individuals infected with a pathogen is introduced. These individuals start infecting susceptibles. This we call an outbreak. Infected individuals can be removed from the system through recovery or death. In mathematical epidemiology we try to determine three things:

1. Does the outbreak cause an epidemic?
2. If so, at what rate does the number of infected individuals increase during the rise of the epidemic?
3. What proportion of the population will ultimately have been infected?

For reasons that will become apparent in Chapter 3, this project focuses on questions 1 and 2. For the scope of this project we make the assumption that the demographic turnover, i.e. the introduction of new individuals through birth and disappearance of individuals through natural death (death not caused by the epidemic), happens at a time scale much larger than the time scale of the epidemic and therefore does not affect the course of the epidemic.

In Section 2.1 we introduce the *basic reproduction number* and what this parameter can tell us regarding questions 1-3. In Section 2.2 we give an example. In Section 2.3 we explore why it may be convenient to assume a fixed proportion of susceptibles in the total population. We continue this theory in Section 2.4, where we analyze how countermeasures can result in an *effective reproduction number* that is sufficiently low to prevent an epidemic, even if the basic reproduction number is not. Lastly, in Section 2.5, we discuss why this analytical framework may not be strong enough to accurately describe the effect of bidirectional contact tracing in the effective reproduction number and therefore the course of the epidemic. In this case we recommend using a simulation rather than analysis.

In the sequel, when we speak of a pair of infective individuals i and j where i has infected j, we say that i is the "parent" of j and j is a "child" of i. In these terms we can consider the *transmission tree* of an individual. As an example, see Figure 1. An individual has only one parent but can have multiple children. Other types of relatives are mentioned in Section 2.5. We expect that it is clear from the context how they relate to the infective individual in question.

2.1 The basic reproduction number

In this Section and Section 2.2, we assume the population to be finite. Together with the assumption on demographic turnover, this results in our total population being closed. That means we can scale the system to the point where we look at proportions of the population. According to custom we let $S(t)$, $I(t)$ and $R(t)$ respectively denote the proportions of Susceptible, Infective and Removed individuals at time t. In a closed and scaled system, we must have $S + I + R = 1$. The outbreak starts at $t = 0$, at which point we have $S(0) + I(0) = 1$ and $R(0) = 0$. In the rest of this Chapter we assume that all contacts are random and that the proportion of newly introduced infectives is small enough that $S(0) \approx 1$. We may also consider

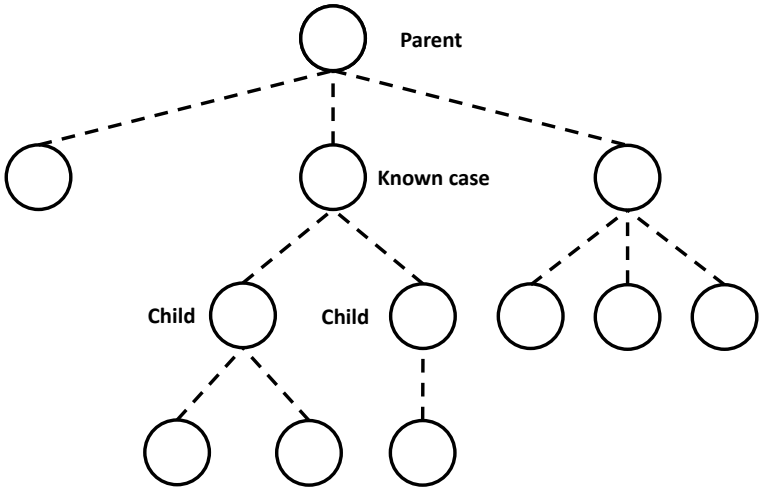


Figure 1: Example of a transmission tree of a known case. Dashed lines represent transmissions.

the time before the outbreak and say $S(-\infty) = 1$, i.e. before the introduction of the pathogen, the entire population is susceptible. Let $F(t)$ denote the force of infection at time t, i.e. the probability per unit of time that a susceptible becomes infected, at time t. By "incidence" we mean the rate of increase in number of infections at a given time t. An infection happens when a susceptible individual is subjected to the force of infection. Thus, we get

$$\text{incidence} = F(t)S(t).$$

Since we assume our population to be closed, the density of susceptibles only changes due to incidence. Hence,

$$\dot{S}(t) = -\text{incidence} = -F(t)S(t).$$

Let $A(\tau)$ denote the expected contribution to the force of infection by an individual at time τ after they were themselves infected. A is a combination of contact intensity and probability of transmission during contact with a susceptible. To be precise,

$$A : [0, \infty) \longrightarrow [0, \infty)$$

is assumed to be integrable. Since A constitutes an expectation, it is not dependent on the individual itself. This is an assumption that will be further discussed in Section 2.5. For now, we can express the force of infection at time t in terms of past incidence by:

$$F(t) = \int_0^\infty A(\tau)F(t-\tau)S(t-\tau)d\tau \tag{1}$$

and thus

$$\text{incidence} = S(t)\int_0^\infty A(\tau)F(t-\tau)S(t-\tau)d\tau. \tag{2}$$

During the outbreak, it is reasonable to assume that all the contacts of an infective individual are with susceptibles. We define

$$R_0 := \int_0^\infty A(\tau)d\tau. \tag{3}$$

We call R_0 the basic reproduction number. It can be interpreted as the expected number of child cases originating from a newly introduced infective individual, i.e. during the outbreak. It is an intrinsic property of the pathogen. We expect that if $R_0 > 1$, the outbreak will cause an epidemic. Conversely, if $R_0 < 1$, we expect that the infective population will go extinct after the outbreak. We can check this with a standard procedure. Naturally, an epidemic experiences either exponential growth or exponential decline. Therefore, we can describe the incidence during the initial outbreak as

$$\text{incidence}(t) \sim ke^{rt}, \tag{4}$$

where r is the growth rate of the epidemic, with $r > 0$ and $r < 0$ corresponding to exponential growth and decline respectively. If we substitute this into equation (2) and assume that $S(0) = 1$, we get the identity

$$1 = \int_0^\infty A(\tau)e^{-r\tau}d\tau. \tag{5}$$

Now let $\phi(r)$ be the right-hand side of equation (5), then it follows from equation (3) that $\phi(0) = R_0$. Furthermore, $\phi(r)$ is strictly decreasing since $A(\tau)$ is integrable. Since growth rate r is the solution to the equation $\phi(r) = 1$, in the case of epidemic growth with $r > 0$, it follows that $R_0 > 1$, and similarly in the case of epidemic decline with $r < 0$ that $R_0 < 1$, as we expected.

2.2 What we can learn from compartmental models

Let us consider an example. Assume that the force of infection F is directly proportional to the number of infective individuals, with constant of proportionality β, so that we can write $F(t) = \beta I(t)$. Furthermore, assume that infectives have a constant probability per unit of time α to become removed. Then the probability to still be infectious τ units of time after infection is equal to $e^{-\alpha\tau}$. It follows that the expected contribution of an individual to the force of infection τ units of time after being infected is $A(\tau) = \beta e^{-\alpha\tau}$. The system

$$\begin{aligned} \frac{dS}{dt} &= -\beta IS, \\ \frac{dI}{dt} &= \beta IS - \alpha I, \\ \frac{dR}{dt} &= \alpha I, \end{aligned} \tag{6}$$

models an outbreak with the above assumptions. This system is called the SIR-model (Susceptible Infective Removed). From (6) it follows that $d/dt(S+I+R) = 0$ and thus our requirement $S + I + R = 1$ holds. The SIR model lends itself for some direct computations that illustrate what we try to do in the general case. To start off, from equation (3) we get

$$R_0 = \int_0^\infty \beta e^{-\alpha\tau}d\tau = \frac{\beta}{\alpha}.$$

We now know that an outbreak will cause an epidemic when $\beta/\alpha > 1$ and that the growth rate is then positive. Now we may start wondering about question 3. However, the derivation of the final size of the epidemic makes more sense when we first try to determine the maximal proportion of infectives, which we shall denote $I_{\max}$. In order to do this we must know how the

proportion of infectives changes with the proportion of susceptibles. If we devide the second equation from (6) by the first, we obtain

$$\frac{dI}{dS} = -1 + \frac{\alpha}{\beta S}. \tag{7}$$

From equation (7) it follows that $I = I_{\max}$ when $S = \alpha/\beta$. Furthermore, if we integrate equation (7), we see that $I + S - \frac{\alpha}{\beta} \ln S$ is a conserved quantity. If we substitute the situation before the outbreak into this quantity, it follows that

$$I + S - \frac{\alpha}{\beta} \ln S = S(-\infty) - \frac{\alpha}{\beta} \ln S(-\infty). \tag{8}$$

Note that the right-hand side of equation (8) equals 1. Hence, we get

$$I_{\max} + \frac{\alpha}{\beta} - \frac{\alpha}{\beta} \ln\left(\frac{\alpha}{\beta}\right) = 1,$$

and thus

$$\begin{aligned} I_{\max} &= 1 - \frac{\alpha}{\beta}\left(1 - \ln\left(\frac{\alpha}{\beta}\right)\right) \\ &= 1 - \frac{\alpha}{\beta}\left(1 + \ln R_0\right). \end{aligned}$$

Now for the final size, we again consider equation (8) and substitute the situation at the end of the epidemic on the left-hand side. Naturally, we expect there to be no more infectives after the epidemic, i.e. $I(\infty) = 0$. After some rearranging, we get

$$\ln S(\infty) = R_0(S(\infty) - 1). \tag{9}$$

The proportion of the total population that was affected by the disease is then $1 - S(\infty)$. As it turns out, we can derive exactly the same formula for $S(\infty)$ for general A. Consider equation (2), substitute $F(t-\tau)S(t-\tau) = -\dot{S}(t-\tau)$ and divide both sides by $S(t)$ to obtain

$$\frac{\dot{S}(t)}{S(t)} = \int_0^\infty A(\tau)\dot{S}(t-\tau)d\tau.$$

If we then integrate with respect to t from $-\infty$ (before the epidemic) to $+\infty$ (after the epidemic) we immediately get equation (9) again. From the above discussion it becomes apparent that the parameter R_0 tells us a lot about the disease, particularly when we try to answer questions 1-3.

The SIR model is the most basic of a class of compartmental models that all have similar assumptions on $A(\tau)$ and therefore describe special cases of the general discussion in this Chapter. These implicit restrictions on the expected contribution of an individual to the force of infection are too narrow for the scope of this project. However, we can and will make use of a compartmental model as a test model for some aspects of the simulation model that we will build in Chapter 3. This will be the more general SEIR model (Susceptible Exposed Infective Removed). It includes a latency period where infected individuals are not yet infectious themselves. We assume that an exposed individual has a constant probability per unit of time γ to become infectious. Then the probability to still be exposed τ units of time after infection

is equal to $e^{-\gamma\tau}$, while the probability to be infectious τ units of time after infection is equal to $e^{-\alpha\tau} - e^{-\gamma\tau}$. It follows that

$$A(\tau) = \beta\frac{\gamma}{\gamma-\alpha}(e^{-\alpha\tau} - e^{-\gamma\tau}).$$

The system

$$\begin{aligned}\frac{dS}{dt} &= -\beta IS,\\ \frac{dE}{dt} &= \beta IS - \gamma E,\\ \frac{dI}{dt} &= \gamma E - \alpha I,\\ \frac{dR}{dt} &= \alpha I\end{aligned} \tag{10}$$

keeps these assumptions [4]. The corresponding basic reproduction number is $R_0 = \beta/\alpha$.

2.3 A fixed susceptible population

In the remainder of this Chapter, we will assume that the that the number of susceptibles that an infective individual encounters remains unchanged, i.e. we consider $S \equiv 1$. That means that we can remove S from our models, and analyse how the incidence grows or declines in when there is an unlimited supply of susceptibles. This setting translates to the outbreak phase of the epidemic discussed in Section 2.1.

Let us formalize our notion of incidence. Let $Y(t,0)$ denote the rate of people acquiring infection at time t and let this be our definition of incidence. Generally, we let $Y(t,\tau)$ denote the number of people at time t who were infected time τ ago. We specify the boundary conditions $Y(t,\tau) = 0$ if $\tau > t$ and $Y(0,0) = Y_{\text{initial}}$. The first condition makes sense; people cannot be infected before the outbreak. The second condition is just our initial population of infectives. It does not make sense to think about the final size of the epidemic in this setting, as we do not determine $S(\infty)$. That means we will disregard question 3 from now on. However, we can determine the cumulative size of the epidemic at time t as

$$Y(t) = \int_0^\infty Y(t,\tau)d\tau.$$

From the fact that $Y(t+\Delta t,\tau) = Y(t,\tau-\Delta t)$, we arrive at the following relation:

$$\frac{\partial Y(t,\tau)}{\partial t} + \frac{\partial Y(t,\tau)}{\partial \tau} = 0. \tag{11}$$

We can relate the current incidence to past incidence in a fashion similar to equation (1):

$$Y(t,0) = \int_0^t A(\tau)Y(t,\tau)d\tau. \tag{12}$$

Note that $A(\tau)$ should be interpreted slightly differently here compared to Sections 2.1 and 2.2. There, A is implicitly tied to the proportion of susceptibles S, whereas here A is the expected infectiousness of an infective in a population with an unlimited supply of susceptibles. We still have

$$R_0 = \int_0^\infty A(\tau)d\tau.$$

However, with equation (3) we explicitly stated that R_0 is only relevant during the initial outbreak when the proportion of infectives is negligible compared to the proportion of susceptibles. That is exactly the setting we consider for the entire epidemic in this Section and thus R_0 can be interpreted as the the expected number of people infected by an infective individual, regardless of when during the epidemic the individual experiences their infectious period. To show that the same conditions on R_0 hold for epidemic growth or decline as in Section 2.1, we again note that an epidemic experiences either exponential growth or decline, i.e.

$$Y(t,\tau) \sim K(\tau)e^{rt}. \tag{13}$$

Substituting (13) into equation (11), we get

$$rK(\tau)e^{rt} + \frac{dK(\tau)}{d\tau}e^{rt} = 0,$$

and thus $K(\tau) = K(0)e^{-r\tau}$. Now substituting (13) into equation (12), we get

$$\begin{aligned} K(0)e^{rt} &= \int_0^\infty A(\tau)K(\tau)e^{rt}d\tau \\ &= \int_0^\infty A(\tau)K(0)e^{r(t-\tau)}d\tau, \end{aligned}$$

again giving us the identity

$$1 = \int_0^\infty A(\tau)e^{-r\tau}d\tau$$

from equation (5). The same analysis as in Section 2.1 applies here, although we can now conclude that for $R_0 > 1$, not only will an outbreak cause an epidemic, but the epidemic will continue to grow if no countermeasures are taken.

2.4 The effect of countermeasures

It is reasonable to assume that during an epidemic, measures are taken to reduce the force of infection. For example, when an infected individual develops symptoms, they may go into isolation, reducing the amount of people they will infect. Furthermore, contacts of this individual may be traced and tested, which leads to them potentially going into isolation as well. Clearly, in the case that measures are in place, an infective individual is expected to infect less susceptibles than in the no-measure case described in Section 2.3. In other words, we have an effective reproduction number $R_{\text{eff}} \leq R_0$. The effects of countermeasures can be incorporated in compartmental type models, see for example [6]. However, we dispute the implicit restrictions this puts on $A(\tau)$.

Fraser and colleagues [5] consider isolation and contact tracing measures in the general setting. Let $B(\tau)$ denote the proportion of individuals not having developed symptoms τ time after being infected. Note that B should be a decreasing function and $B = 1$ implies that individuals never develop symptoms. After developing symptoms, an infective individual is diagnosed and goes into isolation with efficacy $0 \leq \epsilon_{\text{d}} \leq 1$, i.e. the number of potential transmissions caused by this individual is reduced by the factor $(1 - \epsilon_d)$ after they go into isolation. With this measure in place, the incidence from (2) becomes

$$Y(t,0) = \int_0^\infty A(\tau)(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}}B(\tau))Y(t,\tau)d\tau. \tag{14}$$

From this we deduce that isolation reduces the effective reproduction number to

$$R_{\text{eff}} = \int_0^\infty A(\tau)(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau))d\tau = (1 - \epsilon_{\text{d}} + \epsilon_{\text{d}}\theta)R_0,$$

where θ denotes the proportion of infections occurring before symptoms develop:

$$\theta = \frac{\int_0^\infty A(\tau)B(\tau)d\tau}{\int_0^\infty A(\tau)d\tau}.$$

To control an outbreak, we must have $R_{\text{eff}} < 1$. If for example $\theta = 1/2$, this forces $\epsilon_{\text{d}} > 2 - 2/R_0$. Since ϵ_{d} is a proportion, isolation of diagnosed individuals can only prevent an epidemic if $R_0 \leq 2$. Otherwise, it will only slow down the epidemic.

However, we can further reduce the effective reproduction number by performing contact tracing on the diagnosed individual. For the sake of argument, we start by only considering one-step forward contact tracing. That is, we assume that only people infected earlier by a symptomatic individual can be found. We will summarize Fraser's [5] continued analysis here. We make the unrealistic assumption that isolation and contact tracing are independent. In reality, contact tracing is only performed on symptomatic (and thus isolated) individuals. Let $Y(t, \tau, \tau')$ denote the number of people at time t who were infected time τ ago by people who were themselves infected time τ' ago. Similar to equation (11), we have

$$\frac{\partial Y(t,\tau,\tau')}{\partial t} + \frac{\partial Y(t,\tau,\tau')}{\partial \tau} + \frac{\partial Y(t,\tau,\tau')}{\partial \tau'} = 0. \tag{15}$$

Furthermore, let ϵ_{f} denote the efficacy of isolation after being contact traced. Then the incidence becomes:

$$Y(t,0) = \int_0^\infty A(\tau)(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau)) \int_\tau^\infty \left(1 - \epsilon_{\text{f}} + \epsilon_{\text{f}} \frac{B(\tau')}{B(\tau' - \tau)}\right) Y(t,\tau,\tau')d\tau' d\tau. \tag{16}$$

The factor $B(\tau')/B(\tau' - \tau)$ is the proportion of people whose parent has not yet developed symptoms while they themselves have not yet developed symptoms either. Let us substitute $\rho = \tau' - \tau$, the infection generation time, i.e. the time elapsed between onset of symptoms in the parent case and onset of symptoms of the child case. We get

$$Y(t,0) = \int_0^\infty A(\tau)\left(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau)\right) \int_0^\infty \left(1 - \epsilon_{\text{f}} + \epsilon_{\text{f}} \frac{B(\rho + \tau)}{B(\rho)}\right) Y(t,\tau,\rho+\tau)d\rho d\tau. \tag{17}$$

If we again make an unlikely assumption, namely that the time to onset of symptoms is exponential, i.e. $B(\tau) = e^{-\nu\tau}$, then $B(\rho + \tau)/B(\rho) = B(\tau)$ and equation (17) reduces to

$$\begin{aligned} Y(t,0) &= \int_0^\infty A(\tau)\left(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau)\right) \int_0^\infty \left(1 - \epsilon_{\text{f}} + \epsilon_{\text{f}} B(\tau)\right) Y(t,\tau,\rho+\tau)d\rho d\tau \\ &= \int_0^\infty A(\tau)\left(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau)\right)\left(1 - \epsilon_{\text{f}} + \epsilon_{\text{f}} B(\tau)\right) Y(t,\tau)d\tau, \end{aligned}$$

and thus

$$\begin{aligned} R_{\text{eff}} &= \int_0^\infty A(\tau)\left(1 - \epsilon_{\text{d}} + \epsilon_{\text{d}} B(\tau)\right)\left(1 - \epsilon_{\text{f}} + \epsilon_{\text{f}} B(\tau)\right) d\tau \\ &= \left[(1-\epsilon_{\text{d}})(1-\epsilon_{\text{f}}) + \epsilon_{\text{d}}(1-\epsilon_{\text{f}})\theta + \epsilon_{\text{f}}(1-\epsilon_{\text{d}})\theta + \epsilon_{\text{d}}\epsilon_{\text{f}}\psi\right] R_0, \end{aligned}$$

where ψ denotes the proportion of infections caused by asymptomatic individuals whose parents have not yet developed symptoms either:

$$\psi = \frac{\int_0^\infty A(\tau)B(\tau)^2 d\tau}{\int_0^\infty A(\tau)d\tau}.$$

Let us again consider the example where $\theta = 1/2$. In addition, assume that $\psi = \theta^2 = 1/4$ and that $\epsilon_\mathrm{d} = \epsilon_\mathrm{f} = \epsilon$. Then the outbreak is controlled when

$$1 - \epsilon + \frac{1}{4}\epsilon^2 < \frac{1}{R_0}.$$

This forces $\epsilon > 2 - 2\sqrt{1/R_0}$. Again, ϵ is a proportion, thus isolation of diagnosed and forward traced individuals under these circumstances can be effective when $R_0 < 4$. Note that this is already a lot better than when individuals go into isolation only after symptom onset.

2.5 Discussion

In this Section we will review the assumptions we had to make to build the analytical framework in this Chapter. Although these assumptions were necessary to simplify the analysis to a point where it is workable, we have sacrificed some accuracy. To start off, the models in this Chapter do not take into account differences between individuals. In reality, we can expect that some groups of people are more likely to get infected or are more infectious. Factors that may play a role in this are age, health, social status, etc. We mentioned in the Introduction that we want to distinguish at least two non-standard groups of people. Super spreaders are individuals who are expected to infect an amount of people high above the average. In other words, their expected contribution to the force of infection during their infectious period is higher than normal. Asymptomatic individuals do not develop symptoms for their entire infective period, or do not go into isolation after developing symptoms, for example because they mistake them for another disease.

Secondly, when considering the effect of contact tracing, we made the necessary assumption that there is no correlation between isolation and contact tracing. Our equations (14) and (16) only tell us what happens at time t. We cannot know who infected whom and thus we cannot say exactly who can be found through contact tracing. However, such information can be relevant. For example, if a super spreader is tested and positively diagnosed, we expect to find a greater number of child cases through contact tracing.

Thirdly, we only considered one-step forward contact tracing, i.e. parent $\rightarrow$ child. However, the addition of backward contact tracing has a significant advantage over just forward tracing. When backward contact tracing is performed, the parent of a diagnosed individual may be found as well as their children. Moreover, if forward contact tracing is then performed on the parent, its other children may be found and be put into isolation as well. The added benefit of bidirectional contact tracing over just forward tracing is illustrated in Figures 2 and 3. This becomes especially important if we consider asymptomatic individuals. Through bidirectional contact tracing, we can find asymptomatic individuals that would normally not be found, as shown in Figure 3.

Lastly, we have not considered the recursive aspect of contact tracing. Looking back at equation (16), we implicitly stated that an individual can either go into isolation because they have themselves developed symptoms, or because their parent has developed symptoms and they are found through forward contact tracing. This is inaccurate. After all, an individual can

also be forward traced from their grandparent; when the grandparent is diagnosed and forward contact tracing is applied, the parent can be found, after which forward contact tracing is applied *again*, which reveals the child case.

Now let's think about what it would entail to incorporate these aspects into our analysis. First of all, let us say that we consider n different groups of people. The cumulative size of the epidemic within group i can be denoted by Y_i. Of course we must have

$$Y = Y_1 + Y_2 + \cdots + Y_n.$$

The expected contribution to the force of infection by an individual in group i can be denoted A_i. We will need to consider the interaction between the different groups. Let $p_{i,j}$ denote the probability that an individual in group i interacts with an individual in group j. We require

$$\sum_{i=1}^{n} p_{i,j} = 1, \quad \sum_{j=1}^{n} p_{i,j} = 1.$$

Without considering contact tracing, equation (12) turns into the system

$$\begin{aligned}
Y_1(t,0) &= \int_0^t p_{1,1}A_1(\tau)Y_1(t,\tau) + p_{1,2}A_2(\tau)Y_2(t,\tau) + \cdots + p_{1,n}A_n(\tau)Y_n(t,\tau)d\tau, \\
Y_2(t,0) &= \int_0^t p_{2,1}A_1(\tau)Y_1(t,\tau) + p_{2,2}A_2(\tau)Y_2(t,\tau) + \cdots + p_{2,n}A_n(\tau)Y_n(t,\tau)d\tau, \\
&\vdots \\
Y_n(t,0) &= \int_0^t p_{n,1}A_1(\tau)Y_1(t,\tau) + p_{n,2}A_2(\tau)Y_2(t,\tau) + \cdots + p_{n,n}A_n(\tau)Y_n(t,\tau)d\tau,
\end{aligned}$$

We can continue our analysis from here, albeit with a lot more effort. If the reader is interested, we recommend [2] as an excellent source. Instead, we halt the analysis here and go back to equation (16). We will try to expand it to account for multi-step bidirectional contact tracing. In that case, an infective individual can then be found even if neither they nor their parent have developed symptoms, as long as one of their "siblings" has developed symptoms. We now have to consider all possible ways that people can be found through contact tracing, along with their respective proportions, rather than just through one step forward tracing, with proportion $1 - B(\tau')/B(\tau' - \tau)$ as in (16). Furthermore, we have to keep in mind that we only consider the earliest possible moment that an individual can be found through tracing. An individual can be traced multiple times, but can only go into isolation once. For example, the proportion of people at time t who were infected time τ ago by people who were themselves infected time τ' ago that can be found through backward- and then forward tracing, i.e. sibling $\to$ parent $\to$ child is then

$$\int_0^{\tau'} (1 - B(\tau' - \nu)) \frac{B(\tau')}{B(\tau' - \tau)} d\nu.$$

To describe the proportion of the population that can be found through two-step forward contact tracing, i.e. grandparent $\to$ parent $\to$ child, we are required to consider $Y(t, \tau, \tau', \tau'')$, the number of people at time t who were infected time τ ago by people who were infected time τ' ago by people who were infected time τ'' ago. Going on from here will become increasingly more difficult. Moreover, there is no definitive point where we can stop adding terms to our equations due to the recursive nature of the tracing process.

We conclude our discussion here with the remark that incorporating multistep bidirectional contact tracing into the analytical framework built in this Chapter is hard. Instead, we built a simulation model to study its effect on the containment of infectious diseases.

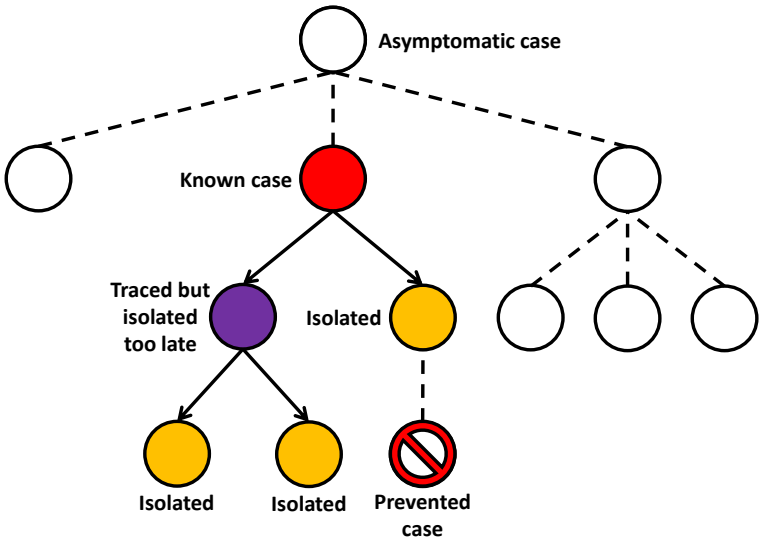


Figure 2: Multi-step forward contact tracing only. White nodes represent individuals that are not found through tracing. Arrows show the direction of tracing. Dashed lines represent (potential) transmissions.

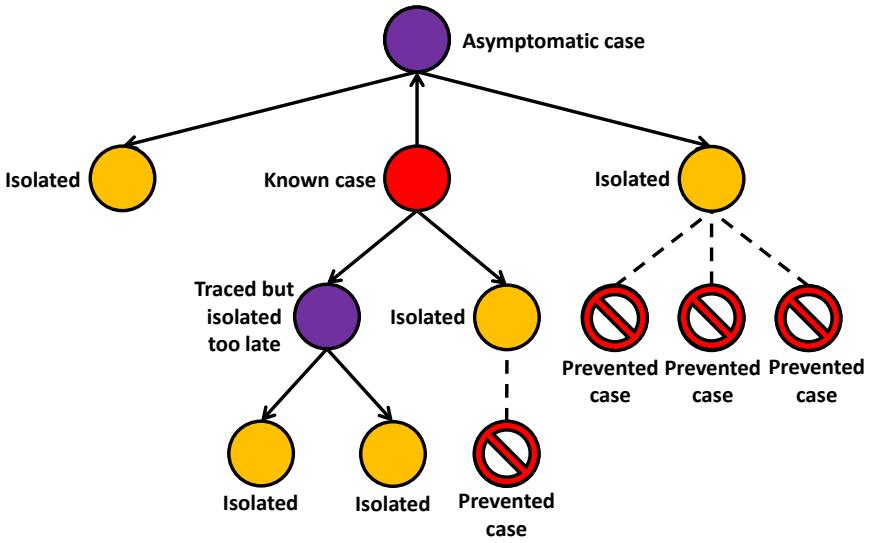


Figure 3: Multi-step bidirectional contact tracing. Arrows show the direction of tracing. Dashed lines represent (potential) transmissions.

3 Simulating bidirectional contact tracing

From the discussion in Section 2.5, we conclude that the existing analytical framework has not sufficiently incorporated the effects of bidirectional contact tracing in order to provide qualitative answers to the questions this project poses. Moreover, updating the analytical framework is unfeasible, especially due to the recursive nature of (bidirectional) contact tracing. Instead, we built a simulation model. This simulation model is described in Appendix A, Sections A.1-A.3. It is based on a simulation model by Bradshaw and others [7, 8]. Primarily, the simulation model we built simulates the spread of an infectious disease with characteristics similar to COVID-19. It can be used to obtain results regarding our hypotheses H1 and H2 in the context of questions 1 and 2. The simulation model has various input parameters that together describe a scenario. It runs an event-by-event simulation that produces a cumulative list of people that were infected, along with their infection times and their infector. The simulation model takes an initial list of infectives with corresponding events. Units of time in the simulation model correspond to days. The simulation model can do the following:

1. For a given scenario, predict the resulting R_{eff} and the growth of the infective population over time.
2. For a given scenario with R_0 free, determine a critical region for R_0 where contact tracing is a sufficient measure for inducing epidemic control.

The simulation model can be used with bidirectional contact tracing or with forward contact tracing only, thus allowing us to compare the two. The simulation model helps us get around some of the restricting assumptions from the analytical framework. However, it is itself not built without any assumptions. We start this Chapter by explicating the assumptions made in the construction of the simulation model. In Section 3.2 we discuss some features of the simulation model as well as some difficulties. Most importantly, it turns out that predictions made by the simulation with regards to 1 are not always reliable. We need the freedom given by 2 to say anything with certainty.

3.1 Assumptions

While building the simulation model we make three significant assumptions. These assumptions are necessary for the simulation model to work in the way we want to while still being able to adequately perform the intended functions.

First and foremost, we assume that the course of the epidemic is not affected by the relative proportions of susceptibles and infectives in the population. In other words, the number of contacts of infective individuals with susceptibles remains unchanged. This assumption is in line with what we discussed in Section 2.3.

Secondly, we assume that contacts are random. In reality, infective individuals naturally infect people that are close to them in a spatial sense, for example housemates, as well as in a relational sense, such as family or friends. Nevertheless, the simulation model does not consider spatial dimensions, so individuals are neither "moving", nor can we pin down their "location". The simulation model also does not keep track of relations between individuals, other than transmissions. We feel that the omission of these effects will not drastically change any conclusions drawn.

Lastly, we assume that contact tracing and resulting isolation is the only countermeasure that affects the workings of the simulation model directly. This assumption allows us to look at the

effect of contact tracing in a vacuum without other influences. In reality, other countermeasures such as curfews and lockdowns will further decrease the effective reproduction number. We will therefore treat the input parameter R_0 in the simulation model as the reproduction number for a disease given any countermeasures besides contact tracing are in place. This interpretation becomes especially advantageous in Section 3.3.

We also make a lot of minor assumptions. We list them below for completeness although we feel that not much is lost if the reader skips over these. How these assumptions appear in the simulation model is explained in Section A.2.

1. Demographic turnover does not affect the course of the epidemic.
2. The initial infective population does not have parents that initiate contact tracing.
3. transmission times are independent of the time of day.
4. There is only one type of super spreader.
5. There is only one type of asymptomatic individual.
6. An individual can be both a super spreader and asymptomatic.
7. There is no correlation between being asymptomatic and being a super spreader.
8. There is no contact correlation between groups.

Let P_s be the proportion of super spreaders. Let C_s and C_r be constants such that

$$C_\mathrm{s}P_\mathrm{s} + C_\mathrm{r}(1 - P_\mathrm{s}) = 1.$$

9. We have $P_\mathrm{s} \leq 0.4$.
10. Regular infective individuals are expected to infect $C_\mathrm{s}R_0$ other individuals.
11. Super spreaders are expected to infect $C_\mathrm{r}R_0$ other individuals.
12. The asymptomatic status does not affect the expected number of infections.
13. The actual number of infections is Poisson distributed around the expected infectiousness.
14. The incubation time of an individual is normally distributed around some average, but no earlier than their infection time and no later than their recovery time.
15. The transmission times of an individual are normally distributed around some average, but no earlier than their infection time and no later than their recovery time.
16. The delay for testing and tracing the parent and children of a diagnosed individual are normally distributed around some average, but no earlier than the time of diagnose of that individual.
17. This delay for testing and tracing can be considered as one delay.
18. Contact tracing is not applied to an individual who is already in isolation.

In assumptions 14-15, if the generated normally distributed time stamp is earlier then the infection time, we set it equal to the time of infection. If the generated normally distributed time stamp is later then the recovery time, we set it equal to the recovery time. In assumption 16, if the generated normally distributed time stamp is earlier then the time of diagnose of the initiating individual, we set it equal to the time of diagnose.

3.2 Features of the simulation model

Let us first discuss in what way a disease should be similar to COVID-19 in order for its course to be accurately described by the simulation model. The COVID-19 virus is mainly transmitted on short-range contact between any group of individuals through aerosols in their breath [17]. Therefore, transmissions can effectively be reduced by isolating infective individuals. These characteristics are key to the accuracy of the simulation model. This excludes for example sexually transmittable diseases. Additionally, the average incubation time, i.e. the time between infection and symptom onset of an individual, should be in the order of days or weeks. Likewise, the average serial interval, i.e. the time between successive cases should also be no more than a few weeks. Contact tracing over longer periods of time will be less accurate and therefore not a consistent countermeasure. Furthermore, effective tests for the disease that can also identify presymptomatic or asymptomatic individuals should be available.

Having identified the types of diseases that can be modelled using the simulation model, we move on to the features that we have incorporated. We especially focus on the aspects lacking in the analytical framework. First of all, in order to distinguish between different groups of individuals, a simulation archives the relevant characteristics of each individual in its respective i-state (see Section 6.1 in [2]). The i-state includes an entry that indicates the parent of the individual. Therefore, the simulation can keep track of the transmission tree through the i-states. When incorporating contact tracing, this feature ensures the direct correlation between a diagnosed individual and the resulting tracing process, since it can target the "relatives" of the diagnosed individual directly. Additionally, this feature enables the simulation to perform recursive contact tracing. That is, when the contacts of a diagnosed individual are traced and diagnosed as well, we can apply contact tracing again simply by following the transmission tree. The i-state also includes a status entry that can mark the individual as a super spreader or as being asymptomatic.

Epidemiological Parameters		
Description	Name	Value
Basic reproduction number	R_0	3
Average incubation time	$T_{\text{incubation}}$	5.5 days
Average serial interval	T_{serial}	5.5 days
Average recovery time	T_{recovery}	16.5 days
Proportion of the population that remains asymptomatic	P_{a}	0.3
Proportion of the population that is a super spreader	P_{s}	0.1
Measure Parameters		
Description	Name	Value
Isolation efficacy	ε	0.8
Probability to be contact traced	P_{trace}	0.8
Average tracing delay	T_{delay}	0.5 days

Table 1: Relevant parameters in the simulation model.

The detailed characteristics of the disease under scrutiny can be approximated through the setting of a set of epidemiological parameters. Similarly, the projected effectiveness of bidirectional contact tracing can be varied through the setting of a set of measure parameters. These parameters with chosen values form a scenario. We have some standard values for these parameters that we use, unless stated otherwise, in examples moving forward. See Table 1.

Essentially, one simulation boils down to a stochastic process. Due to coincidence, two sim-

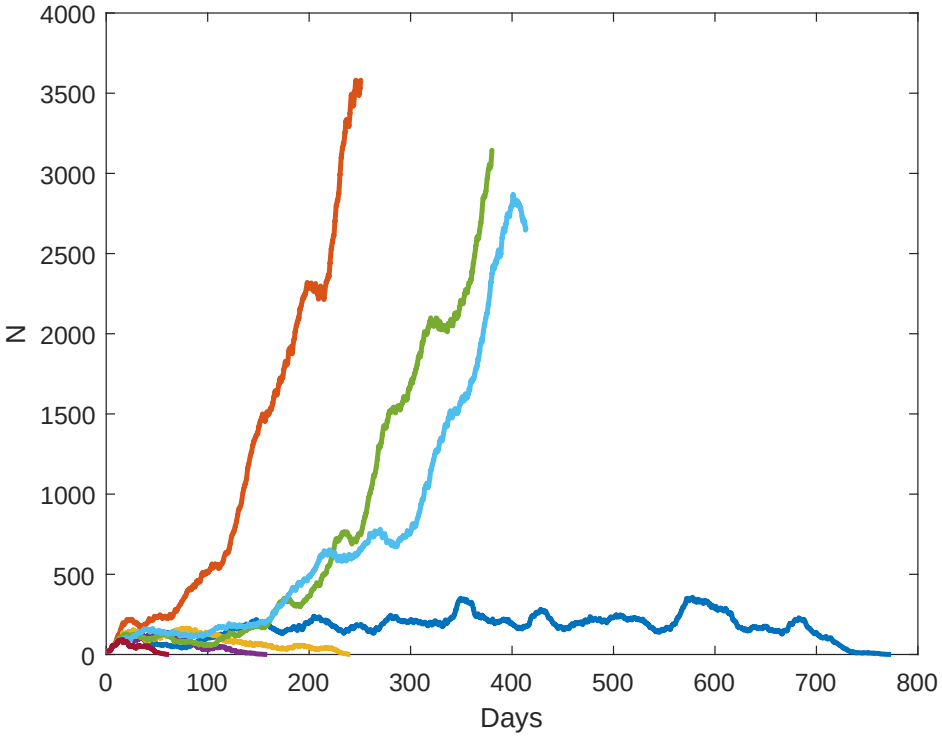


Figure 4: Infective population sizes N over days resulting from 7 simulations with the exact same parameter inputs (regular parameter values except $\varepsilon = 0.93$).

ulations running with the same given set of epidemiological- and measure parameter values can predict different outcomes in terms of epidemic control as described in Section A.5, see Figure 4. If one parameter is free and the others are fixed, then there exists a region for the free parameter where the unpredictability of the outcome becomes more pronounced. There is some value for the free parameter where we the simulation is as likely to predict growth as it is to predict decline. We call this the critical parameter value.

Let us consider an example. Figure 4 suggests that the critical parameter value for ε for the standard values of the epidemiological- and measure parameters is near 0.93. An experiment where we run 100 simulations with the simulation parameters at the values in Figure 4 and $\varepsilon = 0.93$, we get the predictions specified in Table 2 in terms of epidemic control. For a reference guide on how we determine these observations, see Section A.5. We see that under these conditions, 78% of the simulations predicts extinction of the infective population, while 18% results in a stability and 4% results in epidemic growth. This seems like a fine result. However, note that we have considered an initial population of 20 infective individuals for each simulation. In reality, we should translate this result to saying that on an initial infective population of say 2000 individuals, we expect that at least 8 of these individuals starts a growing branch of offspring. Ultimately, contrary to what the results suggest at first glance, epidemic growth is is still likely to occur under these conditions. We must conclude that near critical parameter values, the predictions made by the simulation are unreliable.

We could circumvent this stochasticity by increasing the initial population. However, we then run into another complication with the simulation model. It frequently employs the Quicksort

Prediction	Frequency
Growth of infective population	4
Stable infective population	18
Extinction of infective population	78

Table 2: Frequency of predicted outcomes in 100 simulations. Parameter values as described above.

algorithm for ordering lists. The Quicksort algorithm is of time complexity $\mathcal{O}(n \log(n))$, where n is proportional to the infective population size. Despite its name, this regular sorting makes a simulation increasingly time consuming as the infective population increases. This is especially detrimental when simulating a scenario that induces epidemic growth. In Section A.3 we discuss some adaptations that we made to the basic simulation model in order to reduce the frequency at which this algorithm is invoked, as well as the length of the list that must be ordered. However, these adaptations do not decrease the time complexity by more then a constant factor. Both problems can be solved simultaneously. We discuss our approach in Section 3.3.

3.3 Control of a delayed system

With the observations from the previous Section in mind, it would be useful if we could determine critical parameter values. We choose to try to find the critical value for R_0, which could then be interpreted as the maximal reproduction number that can still be controlled under the given scenario, i.e. under the conditions described by the other parameter values. Ideally, we would only have to run one simulation in order to determine a critical value for R_0. Moreover, considering the time complexity of the simulation model, we would like the infective population to remain within some reasonable bound throughout the simulation. We accomplish this setting by continually updating the parameter R_0. These updates depend on the growth or decline of the infective population. The details can be found in Appendix B. Incidentally, this approach also solves the difficulties observed in the previous Section. Because the simulation model is stochastic, the updated values of R_0 will not converge, but they will remain within some region. We call the interval where the middle 90% of the updated values of R_0 end up the critical region for R_0. We ignore the first 30 update values because the initial population cannot be found through forward contact tracing. Hence, contact tracing will seem less effective at the start of a simulation.

Remarkably, the process of parameter updating comes down to the control of a delayed system, as it takes time for the effects of an updated R_0 to kick in because of the serial interval. There can be some debate as to how and when we choose to update R_0. The PID (proportional-integral-derivative) algorithm is a control loop mechanism that is widely used to control delayed systems, for example in industrial control systems or cruise control. However, for reasons that are further explored in Section B.2, this algorithm is hard to implement in our simulation model. Instead we use a more naive approach that still works. This approach is also discussed in Section B.2. It comes in two variations. The first requires the user to have a broad upper bound for the permissible basic reproduction number, while the second demands that the user has some idea of where the critical region is.

Considering the third major assumption discussed in Section 3.1, we can interpret this updating of R_0 as the effects of other countermeasures. In that sense, the simulation model with parameter updating outputs a critical region for the reproduction number that must be achieved through other countermeasures in order for the contact tracing to induce epidemic control. The lower bound of the critical region is the reproduction number that should be aimed for.

4 Results

In this Chapter, we consider several scenarios for an outbreak. In Chapter 3 we noted that we can use the simulation model to determine a critical region of the reproduction number for a given scenario. We use this feature as a basis for our conclusions. We will divide our results into two Sections. We first consider hypothesis H1, namely that bidirectional contact tracing can control the spread of infectious diseases similar to COVID-19. Furthermore, we explore the added benefit of bidirectional contact tracing over forward contact tracing only. Then we move on to hypothesis H2 and see if bidirectional contact tracing is an effective measure to counteract a disease where a proportion of the population is asymptomatic or a super spreader. We will test bidirectional contact tracing under scenarios with varying proportions of asymptomatic individuals and super spreaders.

It should be noted that we run the majority of our simulations were parameter updating is active until a predetermined termination time, which we will set at 3000 or 5000 days. This does not mean that we expect an epidemic to persist for this period of time. It is just for simulation purposes that we continue a simulation for such an extended time. Firstly, it gives the the system time to settle after the intitial overshoot. Secondly, it provides us with more data which means the critical region for R_0 can be determined more precisely.

4.1 Setting the epidemiological parameters

For the sake of relevance, we want to describe COVID-19 as closely as possible. We do this using the epidemiological parameters from Table 1. Note that we leave R_0 free, as we try to determine a critical region for this parameter. It is worth mentioning that pinning down a value for this parameter is very hard. Estimates found in [13, 16, 29, 30] range from 1.71-3.58. Additionally, the introduction of new variants of the disease will vary the reproduction number as well.

There are no sound sources on the proportions of asymptomatic individuals or super spreaders within the population. However, considering that we want to test hypothesis 2, we will vary the parameters P_a and P_s. As a starting point, we will take $P_\mathrm{a} = 0.3$ and $P_\mathrm{s} = 0.1$, where the first is slightly lower than in [8]. We feel that the proportions presented there are an overestimation due to the inclusion of presymptomatic individuals in the consulted sources. This distinction is made clearly in [31]. Their estimated proportion supports our starting value of P_a. Findings from [30] and other references in [8] suggest that asymptomatic carriers have no difference in viral loads compared to their symptomatic counterparts. Therefore, we do not consider them less infectious in our simulation model. The value for P_s is assumed.

In [13, 29, 30] we find estimates for the remaining parameters. We find mean incubation times of 5.7-7.4 days and mean serial intervals of 5.2-6.9 days. There is some significance as to where the onset of symptoms occurs with respect to the mean serial interval. If the mean incubation time is much longer then the mean serial interval, then it reduces the effectiveness of contact tracing, as most infections happen presymptomatically. However, we feel that they are similar enough to consider them equal, in which case about 50% of potential infections happen presymptomatically. This number seems to be supported by [13, 29, 30]. We therefore set $T_\mathrm{incubation} = T_\mathrm{serial} = 5.5$. Less research is focused on the recovery time. In [30] the time to viral clearance, i.e. the time from the earliest positive test for an individual that was tested more than once, to the earliest negative test, is determined to be 9.3 days on average. However, it is likely that such an individual was infected a few days before their first positive test. To be safe, we set $T_\mathrm{recovery} = 16.5$.

4.2 Bidirectional contact tracing vs forward contact tracing

In this Section, we take the epidemiological parameter values determined in Section 4.1 and run simulations with varying measure parameters to determine corresponding critical regions for R_0. We also compare bidirectional contact tracing with forward contact tracing only to demonstrate its added potency.

We consider eight scenarios. First, we look at scenarios where the isolation of diagnosed individuals is slack; only 80% of potential infections will be prevented through isolation. Then we increase the isolation efficacy to 95%. In both cases, we increase the effectiveness of contact tracing in four stages. We increase tracing probability from 50% to 80% and decrease the average tracing delay from 2 days to 1 day. We feel that these numbers are achievable, especially when using tracing applications. The results are presented in Table 3. We give an estimate for the critical R_0 value in the given scenario; the average of the updated values excluding the first 30 values. We also give the critical region for R_0. We reiterate that the critical region is the interval where the 90% of the updated values end up, excluding the top 5% extremes and bottom 5% extremes, as well as the first 30 values. The plots of the updated R_0 values are supplemented in Appendix C. In the table, we reference the figures upon which the critical regions are based. In Table 3, as well as the other tables displaying critical regions in this Chapter, the width of the critical regions increases with the average R_0. This is due to how the parameter updating works (for details, see Section B.2).

From Table 3, we can conclude that infectious diseases similar to COVID-19 but with a low enough basic reproduction number can be controlled by contact tracing alone. Bidirectional contact tracing is more effective than froward contact tracing only. This benefit is amplified as the effectiveness of the countermeasures is increased. For COVID-19 specifically, considering the pessimistic estimate of 3.58 for its basic reproduction number, we see that it may still be controlled under the strictest bidirectional contact tracing. If this effectiveness cannot be achieved, other countermeasures have to be employed as well.

The observed amplified effectiveness of bidirectional contact tracing in case of increased tracing effectiveness is of interest. We do some follow-up simulations to see if further improvements in the tracing process should focus on increasing the tracing probability or on reducing delay. We consider the ideal theoretical scenarios with a tracing probability of 100% or a delay of 0 days. In practice, we deem it unlikely that this level of effectiveness will ever be achieved, but it may give some insight.

We see clearly in Table 4 that increasing the probability of being contact traced is more beneficial than reducing the tracing delay. This result confirms observations from Bradshaw an colleagues [8]. From this we conclude that in the case of an infectious disease outbreak, the focus should be on tracing as many contacts of a diagnosed individual, rather than on speeding up the tracing process. In the case that an application is available for digital tracing, it should be used by as many people as possible. Hybridized tracing, i.e. a combination of digital- and manual tracing is recommended.

4.3 Bidirectional contact tracing vs asymptomatic individuals and super spreaders

In this Section, we have a closer look at how super spreaders and asymptomatic individuals affect the effectiveness of bidirectional contact tracing. To test this, we start by considering scenarios where there are no super spreaders or asymptomatic individuals. Then we increase the proportions of super spreaders and asymptomatics separately. We consider only scenarios

Scenario	Estimate for critical value and critical region for R_0 with bidirectional contact tracing.	Estimate for critical value and critical region for R_0 with forward contact tracing only.
$\varepsilon = 0.8$, $P_{\text{trace}} = 0.5$, $T_{\text{delay}} = 2$ days.	1.6985 [1.6300, 1.7935] (10a)	1.6772 [1.6001, 1.7659] (13a)
$\varepsilon = 0.8$, $P_{\text{trace}} = 0.6$, $T_{\text{delay}} = 1.5$ days.	1.8166 [1.7135, 1.9615] (10b)	1.7616 [1.6656, 1.8772] (13b)
$\varepsilon = 0.8$, $P_{\text{trace}} = 0.7$, $T_{\text{delay}} = 1$ days.	1.9364 [1.8641, 2.0270] (10c)	1.8314 [1.7354, 1.9324] (13c)
$\varepsilon = 0.8$, $P_{\text{trace}} = 0.8$, $T_{\text{delay}} = 1$ days.	2.0938 [1.9621, 2.2531] (10d)	1.9098 [1.8220, 2.0003] (13d)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$, $T_{\text{delay}} = 2$ days.	2.0812 [1.9639, 2.2605] (11a)	2.0146 [1.8720, 2.2333] (12a)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$, $T_{\text{delay}} = 1.5$ days.	2.3339 [2.1666, 2.5918] (11b)	2.1187 [1.9995, 2.2859] (12b)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$, $T_{\text{delay}} = 1$ days.	2.7522 [2.5731, 3.0398] (11c)	2.2679 [2.1367, 2.4700] (12c)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$, $T_{\text{delay}} = 1$ days.	3.2085 [3.0009, 3.4263] (11d)	2.4994 [2.3330, 2.7462] (12d)

Table 3: Performance of bidirectional contact tracing compared to forward contact tracing only in various scenarios.

were the isolation efficacy is at 95%. Since isolation is then almost perfect, this puts more emphasis on the tracing process itself.

We conclude that the effectiveness of contact tracing in general is not noticeably affected by the proportion of super spreaders. In the case of forward tracing, this makes sense, because the average number of people infected by an infectious individual does not change, and therefore we do not expect to find fewer or more cases in total when there are super spreaders. In the case of bidirectional contact tracing this result is a bit surprising. A super spreader is more likely to be detected presymptomatically through one of their children, which would result in a larger branch of the transmission tree being screened. Note that this is unlikely, because it requires the child case to have developed symptoms before the parent. However, in extreme cases, this could result in a slight increase in the effectiveness of bidirectional contact tracing. We do not observe such an increase in effectiveness however.

Secondly, we conclude that the effectiveness of contact tracing in general decreases if there are people who will remain asymptomatic. This is unsurprising; branches of the transmission tree initiated by these asymptomatic individuals can only be found after diagnoses of later generations in those branches. However, if we compare Tables 5 and 6, we can see that in

Scenario	Estimate for critical value and critical region for R_0 with bidirectional contact tracing.	Estimate for critical value and critical region for R_0 with forward contact tracing only.
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$, $T_{\text{delay}} = 0$ days.	3.4807 [3.2634, 3.7721] (14a)	2.4642 [2.2868, 2.6896] (14b)
$\varepsilon = 0.95$, $P_{\text{trace}} = 1$, $T_{\text{delay}} = 1$ days.	4.4462 [4.1589, 4.7456] (14c)	2.8994 [2.7132, 3.1710] (14d)

Table 4: Performance of bidirectional contact tracing compared to forward contact tracing only in scenarios with no tracing delay or perfect tracing.

Scenario	Estimate for critical value and critical region for R_0 with $P_{\text{s}} = 0$, $P_{\text{a}} = 0$.	Estimate for critical value and critical region for R_0 with $P_{\text{s}} = 0.25$, $P_{\text{a}} = 0$.	Estimate for critical value and critical region for R_0 with $P_{\text{s}} = 0$, $P_{\text{a}} = 0.5$.
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$, $T_{\text{delay}} = 2$ days.	3.3160 [3.1279, 3.5745] (15a)	3.2408 [3.0264, 3.6364] (16a)	1.6267 [1.5279, 1.8015] (17a)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$, $T_{\text{delay}} = 1.5$ days.	3.8648 [3.5562, 4.3501] (15b)	3.8205 [3.5510, 4.3376] (16b)	1.7647 [1.6709, 1.8937] (17b)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$, $T_{\text{delay}} = 1$ days.	4.4627 [4.2106, 4.7883] (15c)	4.4099 [4.1433, 4.7997] (16c)	2.0014 [1.8816, 2.1540] (17c)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$, $T_{\text{delay}} = 1$ days.	5.1823 [4.8616, 5.7840] (15d)	5.2603 [4.8646, 5.9033] (16d)	2.3237 [2.1793, 2.4746] (17d)

Table 5: Performance of bidirectional contact tracing in scenarios with no asymptomatics or super spreaders, super spreaders only and asymptomatics only.

scenarios with asymptomatics, bidirectional tracing did noticeably outperform forward tracing. With bidirectional contact tracing, siblings of a diagnosed individual may still be found through their parent, even if the parent was asymptomatic, as demonstrated in Figure 3. This will prevent some additional cases. This explains why bidirectional contact tracing does better in the scenarios with a large proportion of asymptomatics.

These conclusions put our findings in the previous Section in a new light. The superiority of bidirectional tracing over forward tracing in the case of COVID-19 is due to the fact that a proportion of the population consists of asymptomatic individuals.

4.4 Discussion

The biggest discussion point to address here is the fact that in order for a simulation to run smoothly, we must have a somewhat accurate upper bound for the critical R_0 region in the given scenario to use as initial guess. Otherwise, the simulation might result in extinction of the infective population before an equilibrium is attained. Ideally, such a result would be prevented by the simulation itself.

Scenario	Estimate for critical value and critical region for R_0 with $P_s = 0$, $P_a = 0$.	Estimate for critical value and critical region for R_0 with $P_s = 0.25$, $P_a = 0$.	Estimate for critical value and critical region for R_0 with $P_s = 0$, $P_a = 0.5$.
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$, $T_{\text{delay}} = 2$ days.	3.2806 [3.1516, 3.4499] (18a)	3.2857 [3.0545, 3.6045] (19a)	1.5685 [1.4946, 1.6880] (20a)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$, $T_{\text{delay}} = 1.5$ days.	3.8091 [3.5911, 4.0782] (18b)	3.7822 [3.5528, 4.1474] (19b)	1.5892 [1.5234, 1.6964] (20b)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$, $T_{\text{delay}} = 1$ days.	4.4252 [4.2035, 4.7835] (18c)	4.4640 [4.0518, 5.2208] (19c)	1.6801 [1.6152, 2.7708] (20c)
$\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$, $T_{\text{delay}} = 1$ days.	5.2796 [4.9316, 5.7307] (18d)	5.1969 [4.7616, 5.7717] (19d)	1.7722 [1.6938, 1.8695] (20d)

Table 6: Performance of forward contact tracing only in scenarios with no asymptomatics or super spreaders, super spreaders only and asymptomatics only.

Secondly, we have noted that the width of the critical R_0 regions increases with their average. It can be argued that they are too wide regardless. Although complete convergence cannot be achieved due to the stochasticity of the simulation model, it might be feasible to narrow them until we observe some predetermined deviation from the average. Improvements on this part might for example result in better indications of how much, if at all, bidirectional contact tracing increases in effectiveness because of super spreaders.

Thirdly, some of our assumptions for the simulation model can be disputed. We use normal distributions to determine our time stamps, although other distributions have been suggested in [8]. This could be an issue mainly for the amount of presymptomatic infections we expect to occur. However, we feel that the choice for normal distributions does not have a major effect as long as we can somewhat accurately determine the incubation time and serial interval. Also, it is conceivable that super spreaders are more likely to be in contact with other super spreaders. We expect that this correlation would increase the effectiveness of contact tracing even more, since we will intercept more super spreaders after diagnosing a super spreader.

Lastly, we want to emphasize that when the basic reproduction number for an infectious disease is within the critical region for a certain set of measure parameter values, we still recommended to have other countermeasures in place as well. After all, the critical region indicates that the infective population will remain stable. If the infective population is already relatively large when contact tracing is initiated, it will remain that large.

4.5 Synopsis

We have built a simulation model that simulates the effect of contact tracing on the spread of infectious diseases similar to COVID-19. This simulation model circumvents some of the difficulties that we encounter if we approach this analysis analytically. Through simulations we have been able to conclude that bidirectional contact tracing has a greater effect than just forward contact tracing if there are asymptomatic individuals in the population. This advantage becomes more pronounced when the probability to be traced increases.

Acknowledgements

First of all, I would like to thank my daily supervisor Martin Bootsma. Without him I would not have as much of an appreciation for the field of epidemiology as I currently do. Secondly, I would like to express my gratitude towards my tutor and second reader, Paul Zegeling. He helped me find the correct daily supervisor and was willing to give suggestions to help improve the project. I thank Erik te Rietmole for our weekly get-togheters; this was exactly the support I needed to stay motivated during the COVID-19 pandemic. Also thanks to all the participants of the seminar Mathematical Epidemiology at Utrecht University for presenting to me the basics of epidemiology. Lastly, I would like to thank Utrecht University and the department of Mathematics specifically for allowing me to write this thesis.

Jim Vollebregt

References

[1] W.O. Kermack, e A.G. McKendrick, *A contribution to the Mathematical Theory of Epidemics*; Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 115 (1927) pp. 700-721.

[2] Diekmann, Odo; Heesterbeek, Hans; Britton, Tom, *Mathematical tools for understanding infectious disease dynamics*, Princeton University Press, Princeton, NJ, 2013. xiv+502 pp. ISBN: 978-0-691-15539-5 92-01 (62P10 92D30).

[3] Murray, J. D.; *Mathematical biology. II. Spatial models and biomedical applications. Third edition.*; Interdisciplinary Applied Mathematics, 18. Springer-Verlag, New York, 2003. xxvi+811 pp. ISBN: 0-387-95228-4 92-02 (92B05 92C05 92D30).

[4] D. Breda, O. Diekmann, W.F. de Graaf, A. Pugliese, R. Vermiglio, *On the formulation of epidemic models (an appraisal of Kermack and McKendrick*, Journal of Biological Dynamics, Vol. 6, Suppl. 2, September 2012, 103-117.

[5] Christophe Fraser, Steven Riley, Roy M. Anderson, Neil M. Ferguson, *Factors that make an infectious disease outbreak controllable*, PNAS April 20, 2004 101 (16) 6146-6151. https://doi.org/10.1073/pnas.0307506101.

[6] Traoré, A., Konané, F.V. *Modeling the effects of contact tracing on COVID-19 transmission.* Adv Differ Equ 2020, 509 (2020). https://doi.org/10.1186/s13662-020-02972-8

[7] William J. Bradshaw, Ethan C. Alley, Jonathan H. Huggins, Alum L. Lloyd, Kevin M. Esvelt, *Bidirectional contact tracing is required for reliable COVID-19 control*, medRxiv, May 10, 2020. https://www.medrxiv.org/content/10.1101/2020.05.06.20093369v2.

[8] William J. Bradshaw, Ethan C. Alley, Jonathan H. Huggins, Alum L. Lloyd, Kevin M. Esvelt, *Bidirectional contact tracing dramatically improves COVID-19 control*, medRxiv, July 14, 2020. https://www.medrxiv.org/content/10.1101/2020.05.06.20093369v2.

[9] Jacob E. Lemieux et. al. *Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events*, Science 10 Dec 2020. DOI:10.1126/science.abe3261

[10] Yang Liu, Rosalind M Eggo, and Adam J Kucharskia, *Secondary attack rate and superspreading events for SARS-CoV-2*, Lancet. 2020 14-20 March; 395(10227): e47. Published online 2020 Feb 27. doi:10.1016/S0140-6736(20)30462-1

[11] London School of Hygiene and Tropical Medicine, *SARS-CoV-2 Superspreading Events from Around the World*, Google Docs, retrieved December 2020. https://docs.google.com/spreadsheets/d/1c9jwMyT1lw2P0d6SDTno6nHLGMtpheO9xJyGHgdBoco/edit#gid=1812932356

[12] Ramanan Laxminarayan, Brian Wahl, Shankar Reddy Dudala, K. Gopal, Chandra Mohan B, S. Neelima, K. S. Jawahar Reddy, J. Radhakrishnan, Joseph A. Lewnard, *Epidemiology and transmission dynamics of COVID-19 in two Indian states*, Science 06 Nov 2020: Vol. 370, Issue 6517, pp. 691-697. https://science.sciencemag.org/content/370/6517/691.

[13] Shujuan Ma, Jiayue Zhang, Minyan Zeng, Qingping Yun, Wei Guo, Yixiang Zheng, Shi Zhao, Maggie H. Wang, Zuyao Yang, *Epidemiological parameters of coronavirus disease 2019: a pooled analysis of publicly reported individual data of 1155 cases from seven countries*, medRxiv, March 24, 2020. https://www.medrxiv.org/content/10.1101/2020.03.21.20040329v1.

[14] Mizumoto, Kenji and Kagaya, Katsushi and Zarebski, Alexander and Chowell, Gerardo, *Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020*, Eurosurveillance, 25, 2000180 (2020), https://doi.org/10.2807/1560-7917.ES.2020.25.10.2000180

[15] Großmann G., Bortolussi L., Wolf V. (2020) *Rejection-Based Simulation of Non-Markovian Agents on Complex Networks*. In: Cherifi H., Gaito S., Mendes J., Moro E., Rocha L. (eds) Complex Networks and Their Applications VIII. COMPLEX NETWORKS 2019. Studies in Computational Intelligence, vol 881. Springer, Cham. https://doi.org/10.1007/978-3-030-36687-2_29

[16] World Health Organization, *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. February 2020.

[17] World Health Organisation, *WHO Coronavirus (COVID-19) Dashboard*, retrieved May 13, 2021. https://covid19.who.int/table

[18] World Health Organization, *Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV)*, 30 January 2020.

[19] World Health Organization, *WHO Director-General's opening remarks at the media briefing on COVID-19—11 March 2020*, 11 March 2020.

[20] John D. Birkmeyer, Amber Barnato, Nancy Birkmeyer, Robert Bessler, Jonathan Skinner, *The Impact Of The COVID-19 Pandemic On Hospital Admissions In The United States*, HealthAffairs, VOL.39, NO.11, September 24, 2020.

[21] RIVM, *Impact van de eerste COVID-19 golf op de reguliere zorg en gezondheid: Inventarisatie van de omvang van het probleem en eerste schatting van gezondheidseffecten*, published December 15, 2020.

[22] Worldometer, active cases, retrieved March 17, 2021. https://www.worldometers.info/coronavirus/

[23] Nina Vindegaard, Michael Eriksen Benros, *COVID-19 pandemic and mental health consequences: Systematic review of the current evidence*, Brain, Behaviour and Immunity, Vol.89, pp. 531-542, October 2020.

[24] Cuiyan Wang, Riyu Pan, Xiaoyang Wan, Yilin Tan, Linkang Xu, Cyrus S. Ho, Roger C. Ho, *Immediate Psychological Responses and Associated Factors during the Initial Stage of the 2019 Coronavirus Disease (COVID-19) Epidemic among the General Population in China*, EIJERPH, Vol.17, Issue 5, February 28, 2020.

[25] Lavezzo E, Franchin E, Ciavarella C, et al. *Suppression of COVID-19 outbreak in the municipality of Vo, Italy*. medRxiv. 2020; published online April 18. DOI:10.1101/2020.04.17.20053157.

[26] Ramelli, Stefano, and Alexander Wagner. *What the stock market tells us about the consequences of COVID-19*. Mitigating the COVID Economic Crisis: Act Fast and Do Whatever 63 (2020).

[27] World Economic Forum, *The Global Risks Report 2020*, January 15, 2020. http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf

[28] Kiam Heong Ang; Chong, G.; Yun Li (2005). *PID control system analysis, design, and technology*. IEEE Transactions on Control Systems Technology. 13 (4): 559–576. doi:10.1109/TCST.2005.847331.

[29] Balram Rai, Anandi Shukla, Laxmi Kant Dwivedi, *Estimates of serial interval for COVID-19: A systematic review and meta-analysis*, Clin Epidemiol Glob Health. 2021 January-March; 9: 157–161. Published online 2020 Aug 26. doi: 10.1016/j.cegh.2020.08.007.

[30] Lavezzo, E., Franchin, E., Ciavarella, C. et al. *Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'*. Nature 584, 425–429 (2020). https://doi.org/10.1038/s41586-020-2488-1.

[31] Diana Buitrago-Garcia and others, *Nicola LowOccurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis*, PLOS Medicine, Published: September 22, 2020, https://doi.org/10.1371/journal.pmed.1003346.

[32] Our World in Data, retrieved June 18, 2021.

A The simulation model

We simulate using MatLab. The simulation model is based on the one created by Bradshaw and others [8], with one key difference: the simulation model used in this project runs an event-by-event simulation instead of a generation-by-generation simulation. This type of simulation model was suggested by Großman and others [15]. A simulation is real-time in the sense that it runs events in chronological order. The simulation model assumes an infinite population of individuals that can potentially be infected. As noted in Section 2.3, this assumption only affects the accuracy of the simulation model if in reality, a significant proportion of the population is infected at any point in time, which we do not expect to happen for diseases similar to COVID-19. However, this approach does imply that we will not be able to achieve any results regarding question 3 from Chapter 2.

In this Appendix we discuss a total of 3 versions of the simulation model. We start off by stating all the important parameters in the simulation model in Section A.1. Then in Section A.2 we give a detailed overview of the basic idea of the simulation model and we present version 1. Lastly, in Section A.3 we adapt the simulation model to make it more time-efficient. We present the alternative versions 2 and 3 and we compare their performance to the original version. The rest of the Appendix is dedicated to the details of how we obtain results from a simulation. In Section A.4 we describe how we deduce the effective reproduction number for a given set of parameters from the data provided by a simulation. In Section A.5 we explicate what other conclusions can be drawn from a simulation and how. With that, we have accomplished function 1 of the simulation model. We describe how the model works for bidirectional contact tracing. When we consider forward contact tracing only, we remove backward contact tracing from the simulation model.

A.1 Parameters

The simulation model depends on three types of parameters; epidemiological parameters, measure parameters and simulation parameters. The epidemiological parameters are intrinsic to the pathogen and must be approximated through research. On the other hand, measure parameters depend on the effectiveness of counteractive measures employed by society. These parameters can be varied to determine how strict counteractive measures should be, depending on the characteristics of the pathogen, in order to control the epidemic. By a scenario, we mean a set of values for the epidemiological- and measure parameters. Lastly, the simulation parameters are necessary to determine a starting point and a termination point. Unless stated otherwise, we will use some standard values for the epidemiological- and measure parameters. We specify the simulation parameters when we feel that they are of note. The relevant parameters for the simulation model with their standard values can be found in Table 7 below. Units of time in the simulation model correspond to days in reality. Thus, time dependent parameter values are given in days as well.

We have discussed the basic reproduction number in Section 2.1. The incubation time is the time between infection and symptom onset. The recovery time is the time between infection and complete recovery. The asymptomatic and super spreader groups are not mutually exclusive; an individual can belong to both.

Since we assume that the only counteractive measure in place is that people go into isolation after having symptoms or after being contact traced (see Chapter A), we have a limited number of measure parameters that are relevant. To start off, ε will be used to update the isolation efficacy of a detected infective individual. Potential infections by this individual will afterwards

Epidemiological Parameters		
Description	Name	Value
Basic reproduction number	R_0	3
Average incubation time	$T_{\text{incubation}}$	5.5 days
Average serial interval	T_{serial}	5.5 days
Average recovery time	T_{recovery}	16.5 days
Proportion of the population that remains asymptomatic	P_{a}	0.3
Proportion of the population that is a super spreader	P_{s}	0.1
Measure Parameters		
Description	Name	Value
Isolation efficacy	ε	0.8
Probability to be contact traced	P_{trace}	0.8
Average tracing delay	T_{delay}	0.5 days
Simulation Parameters		
Description	Name	Value
Initial population size	N_{initial}	–
Maximal cumulative size of the epidemic	C_{max}	–
Termination time	T_{max}	–

Table 7: Relevant parameters in the simulation model.

be rejected with probability ε. Next, P_{trace} is the probability that an individual can be found through tracing. Lastly, T_{delay} is the average of a combination of the delay for being traced and the delay for obtaining test results. It is therefore the average time between the diagnose of an individual and the isolation of their parent and children.

A.2 Basics of the simulation model

A simulation starts with an initial population of infected individuals of size N_{initial} with known i-states stored in a list called *People*. For each individual we determine beforehand their time of symptom onset, their time of recovery, the potential number of people they will infect, and the potential infection times. These time stamps and corresponding events are stored in an ordered *Events* list (the Event queue in [15]). The simulation continually takes the first event in the list and processes its outcome. Throughout the simulation, new individuals with i-states can be added to the *People* list and new events can be added to the *Events* list. Below, (18) depicts the i-state of an individual.

[Number; Number of parent; Infectiousness; Isolation efficacy; Status; Infection time]. (18)

To each individual we assign a number that helps us keep track of the branching process. The personal numbers are assigned chronologically, i.e. a higher personal number means a later infection time. This excludes the initial population, which we assume to have been infected simultaneously at time 0. In (18), infectiousness is a first estimator of the number of other individuals this individual will infect. The infectiousness depends on whether or not the individual is a super spreader. The isolation efficacy in (18) is is a number between 0 and 1 representing the probability that a potential infection initiated by the concerned individual is rejected. It is subject to change under the conditions of an "symptom onset" event (depending on some measure parameters, see Table 7). The status is an indicator of the type of individual. It can take four values:

- Status 0: Individual has no special properties.
- Status 1: Individual is a super spreader.
- Status 2: Individual remains asymptomatic.
- Status 3: Individual is a super spreader and remains asymptomatic.

The infection time is the time stamp at which the individual was infected. An example of the i-state of an individual is:

$$[2; 1; 3; 0.75; 2; 1.0345]. \tag{19}$$

This individual has personal number 2. Hence, we know that it was the second individual to be infected in this simulation. This individual was infected by individual 1. Individual 2 will potentially infect approximately 3 other individuals. A potential infection by individual 2 will be rejected with probability 0.75. Individual 2 will remain asymptomatic while contagious. Individual 2 was infected at time 1.0345. Note that the initial population does not have parents. Thus, for all individuals in the initial population, we set the number of parent entry to -1. Since the individual with number -1 does not actually exist, they cannot be found by backward contact tracing.

Next, (20) depicts an event.

$$\text{[Time stamp; Number of individual initiating the event; Event type]}. \tag{20}$$

The event type in (20) is a number corresponding to a type of event. We distinguish four types of events:

- Type 1: Potential infection.
- Type 2: Symptom onset.
- Type 3: End of infectivity.
- Type 4: Contact traced.

An example of an event is:

$$[2.1543; 2; 1]. \tag{21}$$

This event indicates that at time 2.1543, individual 2 will potentially infect another individual. For each individual, we generate precisely one event of Type 3. We only generate a Type 2 event for individuals with status 0 or 1, since by definition the individuals who remain asymptomatic do not develop symptoms.

Version 1 of the simulation model has the following schematic:

1. Take the first event in the *Events* list.
2. Determine the type of event and process its outcome.
   - 2.1. If the event is Type 2 or 4, the initiating individual goes into isolation. Its isolation efficacy of the initiating individual is updated. We also apply bidirectional contact tracing: new Type 4 events initiated by the parent and children of the initiating individual are be added to the *Events* list with probability p_{trace}.
   - 2.2. If the event is Type 1, the isolation efficacy ε of the initiating individual is consulted. A new individual is infected with probability $1 - \varepsilon$. The new individual is generated with its i-state and added to the *People* list, and new Type 1, 2 and 3 events are generated for this individual and added to the *Events* list.

2.3. If the event is Type 3 the simulation knows the initiating individual is no longer infective.

3. Remove the processed event from the *Events* list.
4. If new events were added to the *Events* list, re-sort the *Events* list with respect to the time stamps.
5. Go back to 1.

When a new individual is added to the simulation as part of the initial population or in step 2.2, say at time T, we immediately determine their i-state and events initiated by them (apart from tracing events). Their time of symptom onset is $T+T_1$, where T_1 is a normally distributed number centered around $T+T_{\text{incubation}}$. If the generated time stamp is earlier then the infection time of the initiating individual, we set it equal to the time of infection. If the generated time stamp is later then the recovery time of the initiating individual, we set it equal to the recovery time. An individual is marked as asymptomatic with probability P_{a}. The time of symptom onset of an individual is always determined. However, we do not generate a corresponding Type 2 event in the case that the individual has the asymptomatic status. The end of infectivity time is exactly $T + T_{\text{recovery}}$.

A newly generated individual is marked as a super spreader with probability $P_{\text{s}} \leq 0.4$. In this case the infectiousness R_{s} of the individual is set to $2.5R_0$. Otherwise, their infectiousness is set to

$$R = \frac{1 - 2.5P_{\text{s}}}{1 - P_{\text{s}}} R_0,$$

so that

$$(1 - P_{\text{s}})R + P_{\text{s}}R_{\text{s}} = R_0.$$

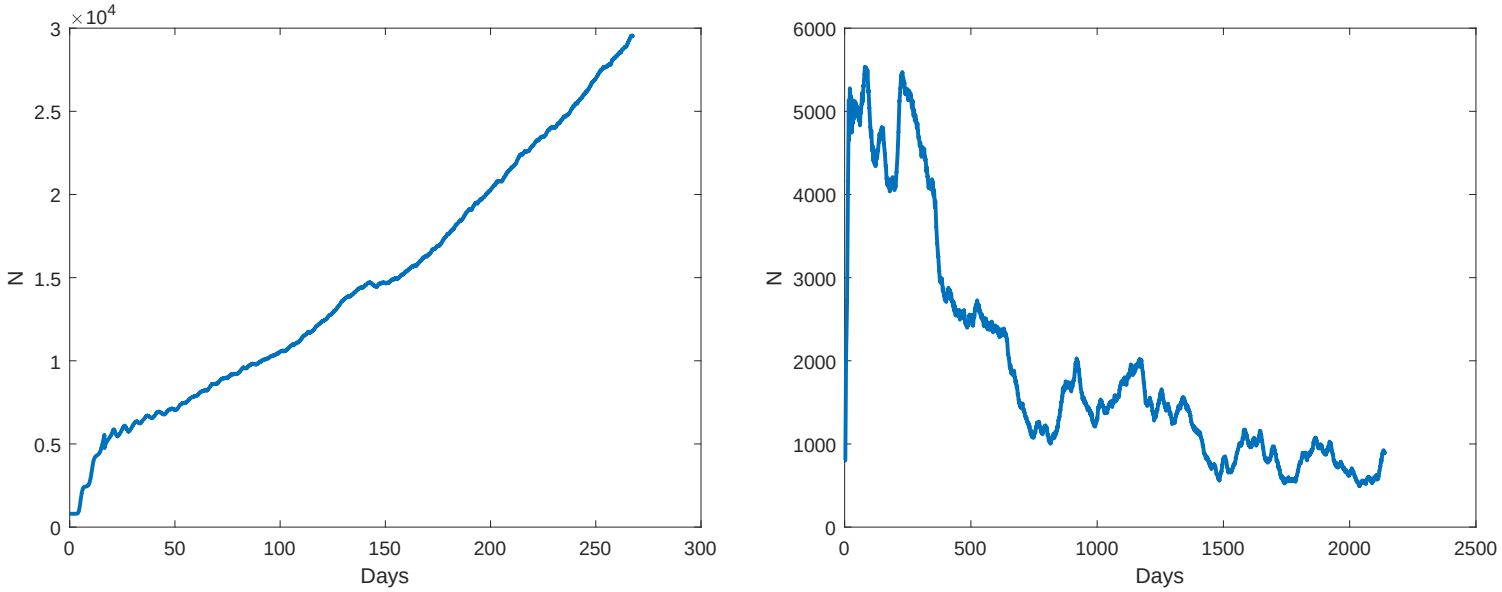
The actual number of potential infections is determined as a Poisson random number n centered around the infectiousness of the individual. The potential infection times are $T+T_j$, $j = 1 \ldots n$, where the T_j are normally distributed around $T+T_{\text{generation}}$. If a generated time stamp is earlier then the infection time of the initiating individual, we set it equal to the time of infection. If a generated time stamp is later then the recovery time of the initiating individual, we set it equal to the recovery time.

In step 2.1, the parent and each of the children of an individual initiating a Type 2 or 4 event has the potential to be traced. Let S be the time of the event and let m denote the number of people traced from this individual. The potential tracing times are $S + T_j$, $j = 1 \ldots m$, where the T_j is normally distributed around $S + T_{\text{delay}}$. If a generated time stamp is earlier then the time of diagnose of the initiating individual, we set it equal to the time of diagnose. Tracing happens when the parent/children satisfy two conditions: they are not already in isolation, and they were infected no earlier than time $S-T_{\text{recovery}}$ time before the event. If these conditions are met, the individual is traced with probability P_{trace}, in which case a Type 4 event is generated. The initiating individuals of the Type 4 events are always the individuals that are being found through tracing.

Note that we make a lot of minor assumptions here. There can be some debate as to how we choose the symptom onset times, the infection times and the tracing times, as well as the factor by which the infectiousness of super spreaders is increased. We feel that a normal distribution for the former three matches reality close enough. As for the factor of increased infectiousness for super spreaders, there is no way to reasonably determine a "correct" factor. Actually, each individual has a different infectiousness. This is reflected in the simulation model through the

number of potential infections of an individual being poison distributed around their expected infectiousness. For the scope of this project, we just want the super spreaders to be a group of people who are expected to be more infectious on average. This factor is therefore completely arbitrary.

The simulation terminates when one of three conditions is met: the cumulative number of individuals in *People* exceeds $C_{\max}$, the *Events* list is empty, or the first time stamp in the *Events* list exceeds $T_{\max}$. After termination of the simulation we are left with a list of *People* and a list of *Events*, which allows us to plot infective population curves over time, see for example Figure 5.



(a) $\varepsilon = 0.92$

(b) $\varepsilon = 0.93$

Figure 5: Simulated infective population size N over days for different values of ε. Other parameter values at standard values. These plots were made using version 3 of the simulation.

A.3 Adaptations

A simulation regularly employs the time-consuming procedure of ordering the potentially long *Events* list. Modifications to improve efficiency should focus on this part. By default, MatLab employs the Quicksort algorithm for ordering lists. This algorithm is of time complexity $\mathcal{O}(n \log(n))$, where in this case n is the number of events. We do not dispute the the strength of this algorithm. Instead, we look elsewhere for improvement. We can take two approaches. Either we reduce the amount of times we have to re-order the whole *Events* list, or we somehow make sure the length of the *Events* list stays within reasonable bounds.

Let us first consider the approach of reducing the frequency of re-ordering the *Events* list. We introduce a new element to the simulation; the *Queue* list of upcoming events. The idea is that, once a potential infection is accepted and new events are generated, rather then immediately adding them to the *Events* list, we place them in this *Queue* and re-order the *Queue*. Only

when the first event in the *Queue* has a time stamp earlier than the first event in the *Events* list do we merge the two and do a large re-ordering. The hope is that the *Queue* list will remain relatively short compared to the *Events* list. On the other hand, we do not want the *Queue* merging too often with the *Events* list.

Version 2 of the simulation model has the following schematic:

1. Take the first event in the *Events* list.
2. Compare the time stamp of this event with the first event in the *Queue*.
3. If the time stamp of first event in the *Events* list is lower than the first event in the *Queue*, proceed as follows; otherwise go to 4.
   - 3.1. Determine the type of event and process its outcome.
     - 3.1.1. If the event is Type 2 or 4, the initiating individual goes into isolation. Its isolation efficacy of the initiating individual is updated. We also apply bidirectional contact tracing: new Type 4 events initiated by the parent and children of the initiating individual are be generated with probability p_{trace} and sorted. If the lowest time stamp in the newly generated events is lower then the lowest time stamp in the *Queue* list, place all new events at the beginning of the *Queue* list. Otherwise, place them at the end.
     - 3.1.2. If the event is Type 1, the isolation efficacy ε of the initiating individual is consulted. Moreover, we consider potential updates to the basic reproduction number. A new individual is infected with probability $1-\varepsilon$. The new individual is generated with its i-state and added to the *People* list, and new Type 1, 2 and 3 events initiated by this individual are generated and sorted. If the lowest time stamp in the newly generated events is lower then the lowest time stamp in the *Queue* list, place all new events at the beginning of the *Queue* list. Otherwise, place them at the end.
     - 3.1.3. If the event is Type 3 the simulation knows the initiating individual is no longer infective.
   - 3.2. Remove the processed event from the *Events* list.
   - 3.3. Go back to 1.
4. If the time stamp of first event in the *Events* list is higher than the first event in the *Queue*, proceed as follows:
   - 4.1. Add the *Queue* to the *Events* list.
   - 4.2. re-sort the *Events* list.
   - 4.3. Go back to 1.

As an alternative to this version, we have Version 3 of the simulation model. In that version, whenever we add events to the *Queue* in step 3.1.1 or 3.1.2, we just put them at the end without sorting them beforehand. Instead, we sort the whole *Queue* after the new events were added. The advantage of Version 2 is that we only have to sort very short lists of events regularly. A disadvantage is the fact that the *Queue* is continually updated which is almost as time-consuming. It is of note that versions 2 and 3 do not require additional assumptions on the development of the epidemic compared to version 1.

Let us compare the performance of the three versions of the simulation model we have discussed in this Chapter. Both from what we expect and from experience it follows that the simulation takes much longer when a scenario is processed where the infective population grows exponentially, as the *Events* list will grow very large very quickly and is still growing when the simulation terminates. On the other hand, if the scenario leads to a controlled or extinct infective population, the simulation is much faster. Therefore we want to compare the performance of the three versions in both outcomes. To accomplish this, we set our parameters to values which we know for certain lead to the desired circumstance. From experience we see that if we take the regualr parameter values but set $\varepsilon = 0.5$, the outbreak will always cause an epidemic with exponential growth (outcome 1). On the other hand, if we choose $\varepsilon = 1$, the outbreak will always lead to extinction (outcome 2). These are therefore logical test examples to consider. There are two more factors that affect the simulation time: the simulation parameters N_{initial}, $T_{\max}$ and, in the case of exponential growth, $N_{\max}$. For this comparison we fix the following values for the simulation parameters:

$$N_{\text{initial}} = 20, \quad C_{\max} = 20000, \quad T_{\max} = 1000.$$

For each scenario we run each version of the simulation model 100 times and compare the averages of their run times. See the table below:

Version	Scenario 1	Scenario 2
1	67.606	$1.727 \cdot 10^{-2}$
2	18.284	$1.397 \cdot 10^{-2}$
3	25.137	$1.442 \cdot 10^{-2}$

Table 8: Performance comparison of the three versions of the simulation model. The data shown is the average simulation time in seconds over 100 simulations.

Considering theses performance results, we present version 2 as the definitive version.

A.4 Measuring the effective reproduction number

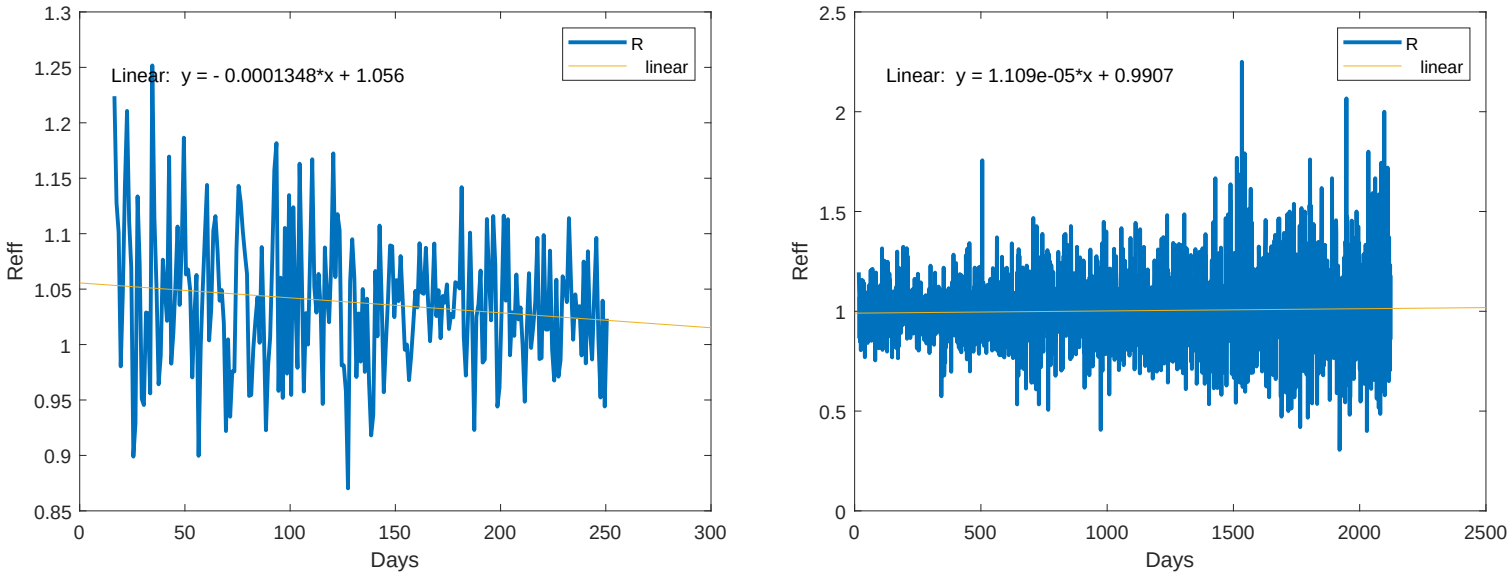
Given a scenario, we want to determine the resulting effective reproduction number. Let us first consider the level of a single simulation. We noted in Section A.1 that units of time correspond to days. Let T_{end} be the time stamp of the last processed event, then $T_{\text{end}} \leq T_{\max}$. We measure the effective reproduction per unit of time K, where

$$K = [k, k+1], \quad 0 \leq k \leq \lfloor T_{\text{end}} \rfloor,$$

by $R_{\text{eff}}^K = N_c/N_p$. Here N_p is the number of individuals with time stamps in K and N_c is the number of child cases originating from these individuals. We determine N_c by counting how often one of these individuals is listed as the parent in the i-state of another individual in the *People* list.

In Section A.2 we have noted that the initial population does not have parents that are themselves in the simulation. Therefore, the initial population cannot be found via forward contact tracing. This means that the measured effective reproduction number per unit of time at the start of the epidemic is overestimated. On the other hand, when the simulation terminates, there may still be individuals who will potentially infect others after the last run event. These new individuals and their corresponding events will never be generated. Due to this the measured effective reproduction number per unit of time at the end of the epidemic plummets

to 0 and is thus underestimated. Other than that we expect that the effective reproduction number per unit of time will average out to some constant. Experimentally it turns out that this number fluctuates a lot, especially in the case of a low total infective population. However, trends show that it is indeed almost constant, see Figure 6. The effective reproduction number predicted by a single simulation is then the average of the effective reproduction numbers per unit of time for the units K with $\lceil T_{\text{recovery}} \rceil \leq k \leq \lfloor T_{\text{end}} - T_{\text{recovery}} \rfloor$.



(a) $\varepsilon = 0.92$

(b) $\varepsilon = 0.93$

Figure 6: Simulated effective reproduction number R_{eff} over days for different values of ε. Other parameter values at standard values. These plots corrsepond to the population curves in Figure 5

A.5 Measuring epidemic control

For a given scenario, some simulations may predict growth of the infective population while others predict extinction. In fact, a third outcome may occur: after the initial outbreak, we could have a more or less stable population. After measuring the effective reproduction number as described in Section A.4, we still want to know in more detail what happened during the simulation. In particular, we want to distinguish between the three outcomes described above. Therefore, we save for each simulation the termination time T_{end}, the final number of infectives N_{end} and the peak number of infectives over the course of the simulation N_{max}. If a simulation has $N_{\text{end}} = 0$ we say that the simulation resulted in extinction. If $T_{\text{end}} = T_{\text{max}}$ or $N_{\text{end}} \leq 0.95 N_{\text{max}} \neq 0$, we say the simulation resulted in a stable population. In all other cases we say the simulation resulted in exponential growth of the infective population.

The factor 0.95 is somewhat arbitrary; we want to avoid marking growing populations as stable populations. It can be that around T_{end}, a large group of individuals has reached their recovery time. Thus, even for a population that is growing overall, it can be that $N_{\text{end}} \neq 0.95 N_{\text{max}}$. We feel that a 5% margin is safe here.

B Parameter updating

Notwithstanding our adaptations to the basic idea of the simulation model presented in Section A.3 and the resulting increased efficiency, a simulation still becomes significantly slower as the number of events increases. Moreover, as discussed in Section 3.2, the predictions made by a simulation are unreliable when we are in a critical region of parameter values. In this Appendix, we show how we use parameter updating in order to obtain a critical region for R_0 in a given scenario. This accomplishes function 2 of the simulation model.

B.1 Controlling the infective population

The idea is that we force a simulation into a state where the infective population size N is relatively stable. Ideally, we make sure N stays within some bound where the simulation can still handle the resulting amount of events in a reasonable time frame. The example in Section 3.2 shows that this will not work if we fix all our parameter values. Instead, we let R_0 be "free" in the sense that we update it depending on the state of the simulation. The idea behind this approach is twofold. On one hand, we circumvent the time complexity of the simulation model by controlling the infective population size. On the other hand, because the infective population size will controlled, it must be that R_{eff} will be close to 1, and this implies that R_0 is in its critical region. How we determine our updated values for R_0 will be discussed in Section B.2.

Since R_0 is no longer a constant parameter, let's use the notation $R_0(t)$, the basic reproduction number at the time t of an event. Let $\bar{R}_0$ be an upper bound for $R_0(t)$ which we know from experiments certainly induces epidemic growth under the given scenario. By the process described in Section A.2, the number of potential infections initiated by an individual is based on $\bar{R}_0$. However, the potential infection itself happens at time t, and it may be that $R(t) \neq \bar{R}_0$. We have to account for this. Considering how we set $\bar{R}_0$, we should certainly have $R(t) \leq \bar{R}_0$. When the basic reproduction number is equal to 1, the infective population is always controlled, hence we can also require $R(t) \geq 1$. Then, when a Type 1 event is processed at time t, we reject it with probability $1 - R(t)/\bar{R}_0$ to account for the fact that the basic reproduction number is now different from the one that the creation of the event was based on. Updates happen every predetermined time step σ. After an update, we want to give the simulation time to adjust before we do another update. Therefore, we set $\sigma = 15$, close to the standard recovery time.

In the simulation model, we can easily incorporate these updates as events. We will give them the Event Type 5 and similar to the chosen number of the parent in the initial population, we choose the "initiating individual" of these events to be -1. One Type 5 event needs to be added to the list of initial events, and a Type 5 event needs to trigger another Type 5 event one time step σ later. The schematic of the simulation model developed from version 2 then becomes:

1. Take the first event in the *Events* list.
2. Compare the time stamp of this event with the first event in the *Queue*.
3. If the time stamp of first event in the *Events* list is lower than the first event in the *Queue*, proceed as follows; otherwise go to 4.
   - 3.1. Determine the type of event and process its outcome.
     - 3.1.1. If the event is Type 2 or 4, the initiating individual goes into isolation. Its isolation efficacy of the initiating individual is updated. We also apply bidirectional

contact tracing: new Type 4 events initiated by the parent and children of the initiating individual are be generated with probability p_{trace} and sorted. If the lowest time stamp in the newly generated events is lower then the lowest time stamp in the *Queue* list, place all new events at the beginning of the *Queue* list. Otherwise, place them at the end.

3.1.2. If the event is Type 1, the isolation efficacy ε of the initiating individual is consulted. Moreover, we consider potential updates to the basic reproduction number. A new individual is infected with probability $(1-\varepsilon)(R_0(t)/\bar{R}_0)$. The new individual is generated with its i-state and added to the *People* list, and new Type 1, 2 and 3 events initiated by this individual are generated and sorted. If the lowest time stamp in the newly generated events is lower then the lowest time stamp in the *Queue* list, place all new events at the beginning of the *Queue* list. Otherwise, place them at the end.

3.1.3. If the event is Type 3 the simulation knows the initiating individual is no longer infective.

3.1.4. If the event is Type 5, update R_{new} and generate a new event of Type 5 with time stamp σ time after that of the initiating event. If the time stamp of the newly generated event is lower then the lowest time stamp in the *Queue* list, place it at the beginning of the *Queue* list. Otherwise, place it at the end.

3.2. Remove the processed event from the *Events* list.

3.3. Go back to 1.

4. If the time stamp of first event in the *Events* list is higher than the first event in the *Queue*, proceed as follows:

4.1. Add the *Queue* to the *Events* list.

4.2. re-sort the *Events* list.

4.3. Go back to 1.

When initiating a simulation, it is reasonable to set N_{initial} equal to N_{optimal}. If we wish to exclude the parameter updating from the simulation process, we need only remove the initial Type 5 event from the initial events.

B.2 How to update

In this Section we discuss how and how often we update R_0. As discussed in Section 3.3, we are basically attempting to control a delayed system. We make a small tangent by discussing the PID controller algorithm that is often used to control similar systems. Ultimately, we find that our simulation model is not suited for the implementation of this algorithm. Instead, we take another approach that can be used in two variations. We have not been able to find other studies that take a similar approach. However, our attempt works reasonably well and we therefore feel that it is unnecessary to take a different approach that requires a lot more effort.

Two phenomena are worth pointing out when considering the control of a delayed system. To start off, the controlling mechanism can be unstable, i.e. its output diverges. The risk of this happening is increased with larger delays. Secondly, even if stability is achieved, there may still be an initial overshoot in the system. In our setting, this translates to an infective population

that grows rapidly before it can be controlled. This can make the simulation slow despite our efforts.

As a preparation for the rest of this Section, let $N(t)$ denote the infective population size at the time t of an event. Moreover, let N_{optimal} be a predetermined number around which we want the infective population size to oscillate. This should be a number of infective individuals that can be handled by the simulation in a timely manner while its development is not too dependent on coincidence.

The PID controller is an example of an algorithm that is often used in similar delayed systems. It continually computes an error value $e(t)$ as the difference between a set point value, in this case N_{optimal} and a process variable, in this case $N(t)$, i.e. $e(t) = N_{\text{optimal}} - N(t)$. This error is then used to define the control function

$$u(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt}, \tag{22}$$

where K_p, K_i and K_d are the coefficients for the proportional, integral and derivative terms respectively. These coefficients must be tuned until the system reaches its reference in a reasonable time frame. However, there are some risks of degrading stability or increased overshoot [28]. We should then set

$$R_0(t) = R_0(t - \sigma) + u(t).$$

Unfortunately, the outputs of our simulation model do not allow us to easily approximate integrals or derivatives of the error function. We would have to add a whole new process to the model. Instead, we update R_0 using some easily determined factors. Experimentally, we observe the same setbacks regarding stability and overshoot.

The first updating variant is to use the predetermined number N_{optimal}. We let $R_0(0) = \bar{R}_0$. It is also natural to set $N_{\text{initial}} = N_{\text{optimal}}$. At the time t of an update, we increase or decrease $R_0(t-\sigma)$ by some factor $\rho_1 = \rho_1(t)$ depending on the ratio $N(t)/N_{\text{optimal}}$. Furthermore, we take into account the growth of the infective population at time t. If, for example, the population is growing while still below the optimal number, we want to mitigate the update by a second factor $\rho_2 = \rho_2(t)$ depending on the ratio $N(t)/N(t-\sigma)$. With this method, we set

$$R_0(t) = R_0(t - \sigma)\rho_1(t)\rho_2(t).$$

There can be some debate as to how we choose the settings of $\rho_1(t)$ and $\rho_2(t)$. We could naively set them to

$$\rho_1(t) = \frac{N(t)}{N_{\text{optimal}}} \quad \text{and} \quad \rho_2(t) = \frac{N(t)}{N(t-\sigma)}$$

respectively. However, experimentally, we find that the resulting updates are too severe; they become unstable. Due to the hard boundaries that we put on $R_0(t)$ in the previous Section, $R_0(t)$ will bounce between 1 and R_{initial}. After some more experiments, we find that setting

$$\rho_1(t) = \sqrt[32]{\frac{N(t)}{N_{\text{optimal}}}} \quad \text{and} \quad \rho_2(t) = \sqrt[32]{\frac{N(t)}{N(t-\sigma)}}$$

does result in a narrowing of the critical region, although some oscillation still occurs. We should also be somewhat careful when choosing $\bar{R}_0$. If we set it too high, the initial overshoot is so large that we either reach $N_{\max}$ and the simulation terminates before we reach a stable state, or the updating process overcompensates to the point where $R_0(t)$ is decreases too far and

the infective population goes extinct. On the other hand, if we set $\bar{R}_0$ too low, it may already be near the critical region in which case $R_0(t)$ may not have enough freedom to oscillate.

The second variant does away with N_{optimal}. We set $R_0(0) = R_{\text{initial}}$. It is convenient if R_{initial} is near the critical region for R_0. Furthermore, we set $\bar{R}_0 = 1.2R_{\text{initial}}$. At the time t of an update, we increase or decrease $R_0(t-\sigma)$ based only on the growth of the infective population at time t, i.e. by some factor $\rho = \rho(t)$ depending on the ratio $N(t)/N(t-\sigma)$. We set

$$R_0(t) = R_0(t-\sigma)\rho(t).$$

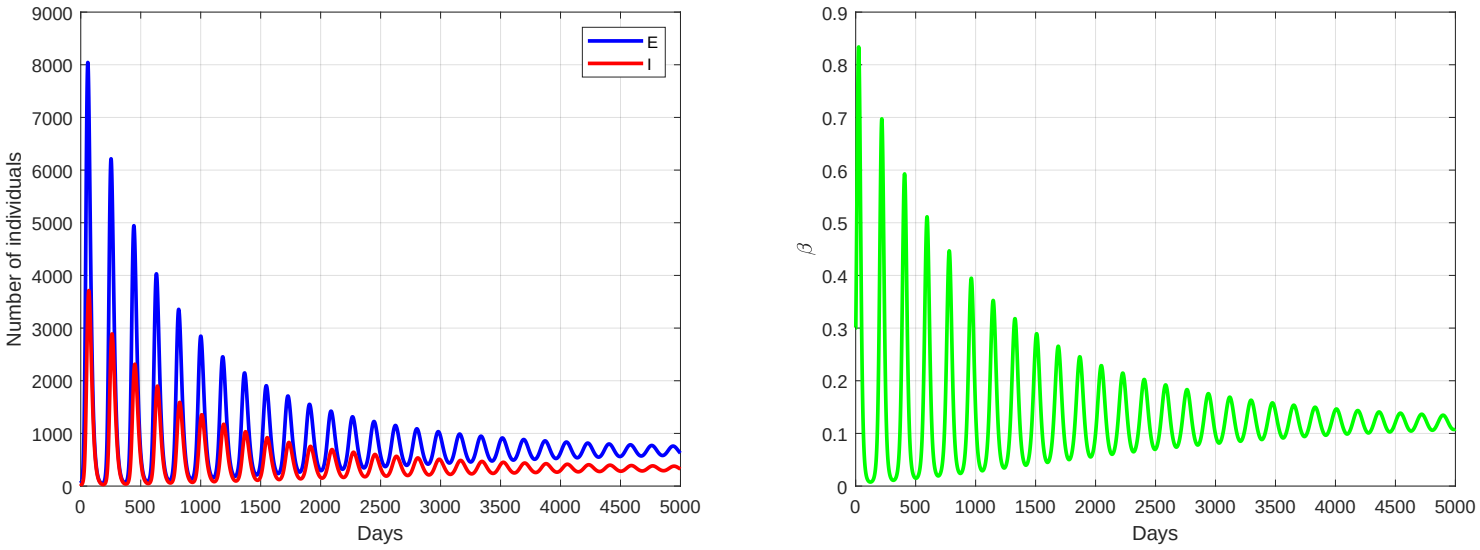
Experimentally, we find that

$$\rho(t) = \sqrt[8]{\frac{N(t)}{N(t-\sigma)}}$$

is optimal for achieving the goal of the simulation model. It is advised that the second variant of the updating process is used when the user has some idea of the critical region; for example if they have done some test simulations beforehand.

We have tested both variants of the updating process using a deterministic SEIR model (see Section 2.2, equation (10)). Here, we update the parameter β which is proportional to R_0. We find that $R_0(t)$ as well as $N(t)$ converge for both variants. Figure 7 below shows the resulting population curves and the parameter β over days under the updating process. Since our simulation model is stochastic rather then deterministic, we cannot expect complete convergence of $N(t)$ and $R_0(t)$ to some fixed value. However, as noted before, we will be able to determine a critical region for R_0. Since the parameter updating is multiplicative, the critical regions will widen as the average R_0 increases.

Figures 10 and 11 in Appendix C show plots of $R_0(t)$ for different scenarios with bidirectional contact tracing. From these Figures, we can infer critical regions for R_0. These plots were made using the first variant of parameter updating. Figures 8 and 9 show the corresponding population curves. Initial overshoots in the infective population can clearly be observed. It should be noted that the majority of the simulations producing these plots runs until the termination time $T_{\max} = 5000$. This does not mean that we expect an epidemic to persist for this period of time. It is just for simulation purposes that we continue a simulation for such an extended time. First, it gives the the system time to settle after the intitial overshoot. Second, it provides us with more data which means the critical region for R_0 can be determined more precisely.

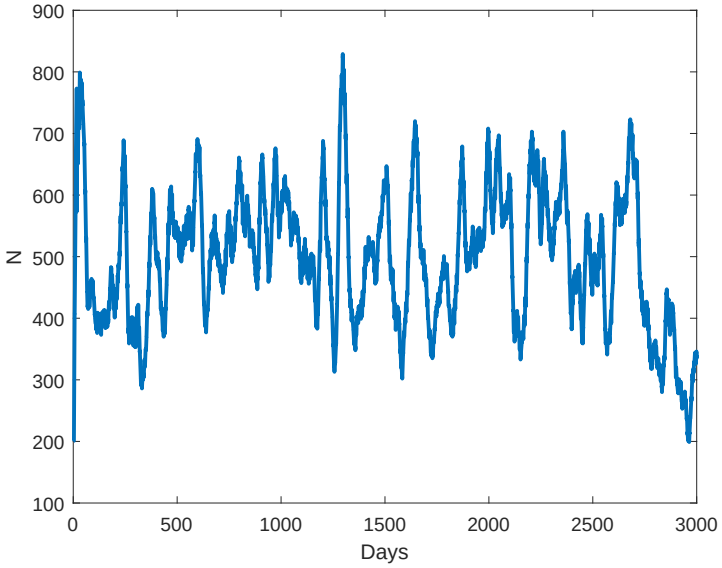


(a) Number of exposed and infectious individuals over days.

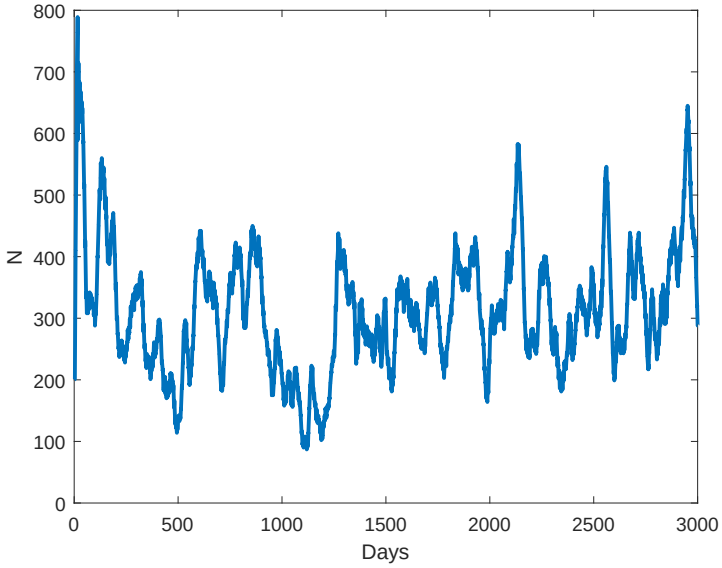
(b) Parameter $\beta(t)$ over days.

Figure 7: The first variant of parameter updating described in this Section induces convergence of population curves and the free parameter β in the deterministic SEIR model.

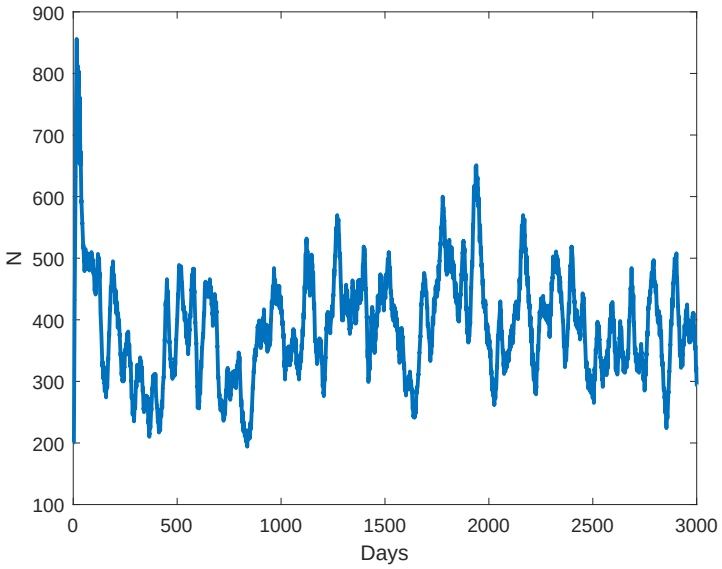
C Figures



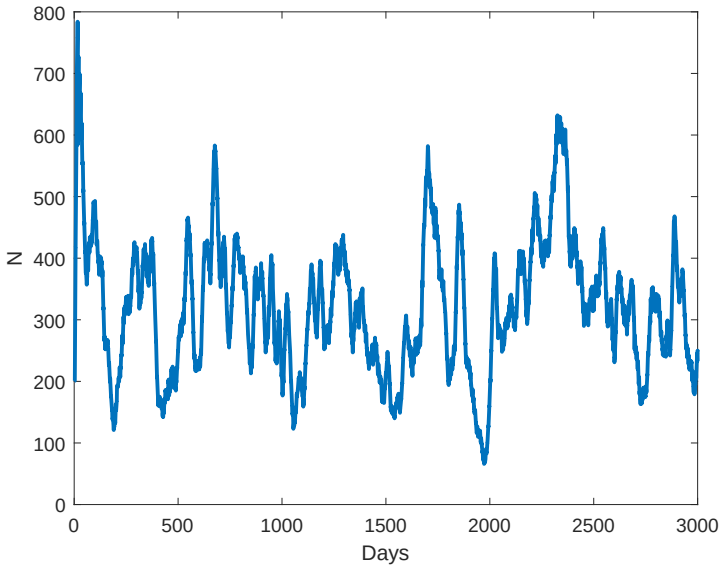
(a) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

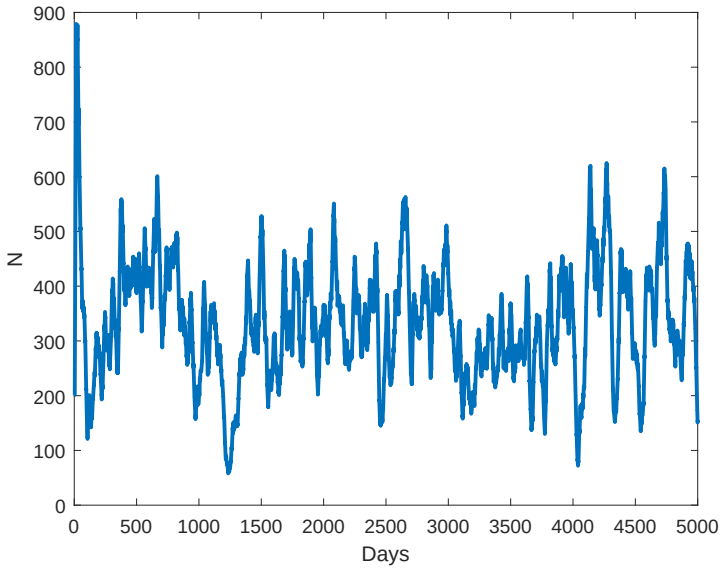


(c) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

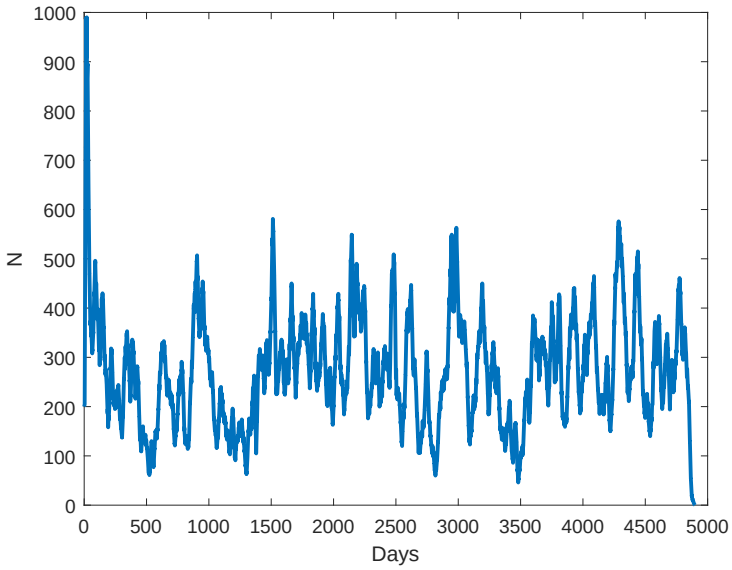


(d) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

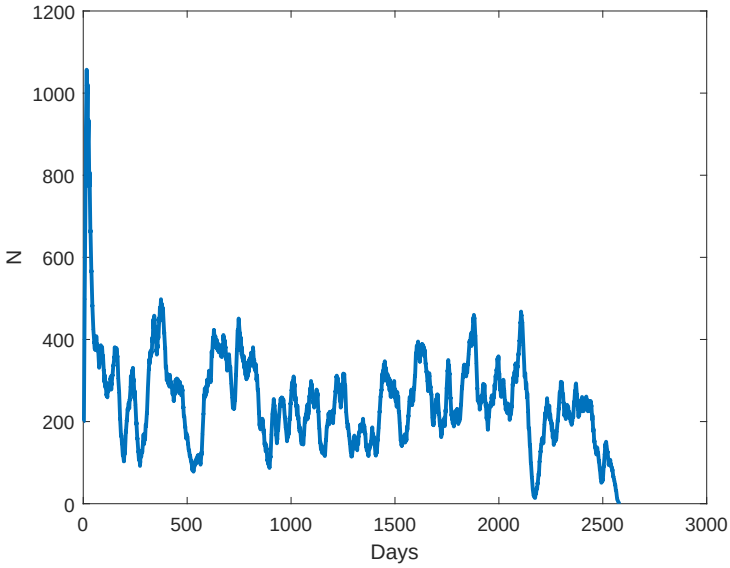
Figure 8: Simulated infective population size N over days for different settings of the measure parameters. Other parameter values at standard values. Bidirectional contact tracing.



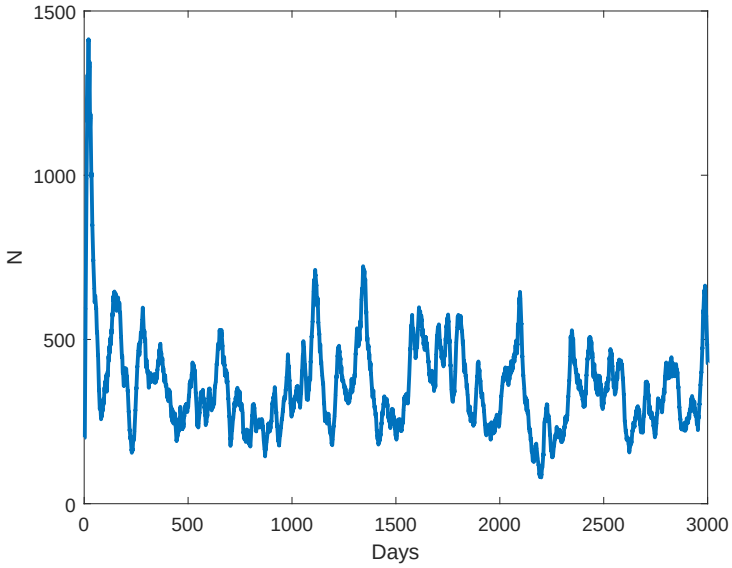
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

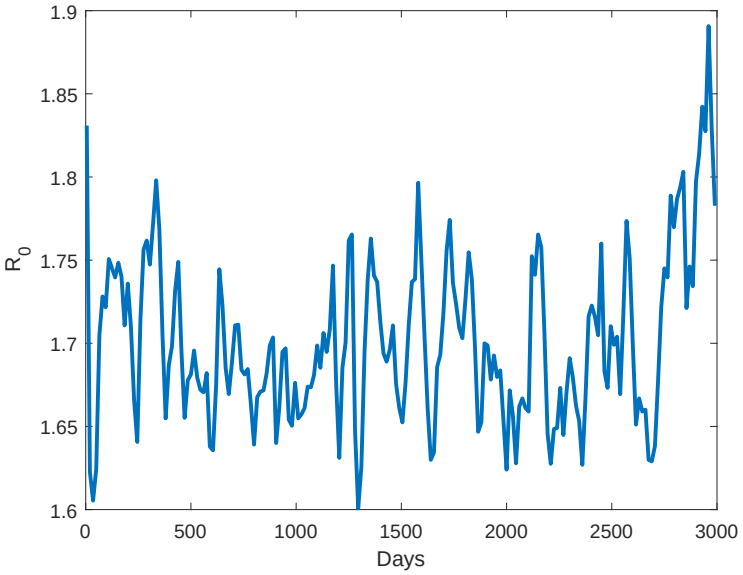


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

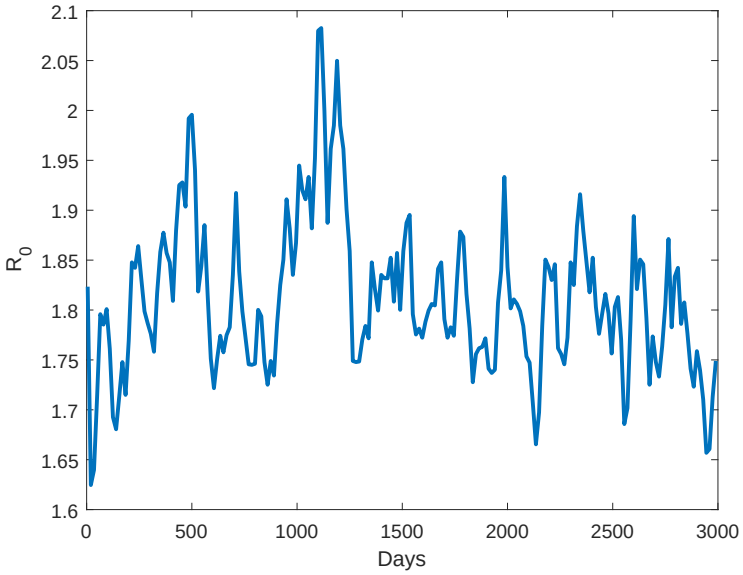


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

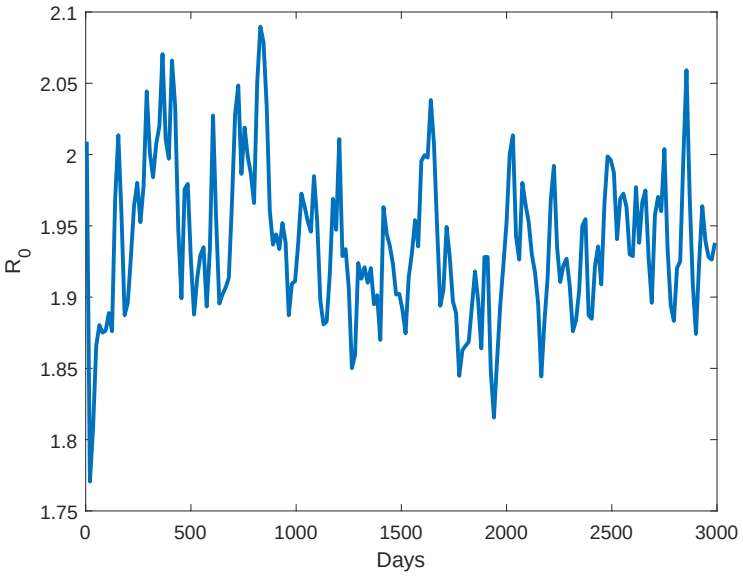
Figure 9: Simulated infective population size N over days for different settings of the measure parameters. Other parameter values at standard values. Bidirectional contact tracing.



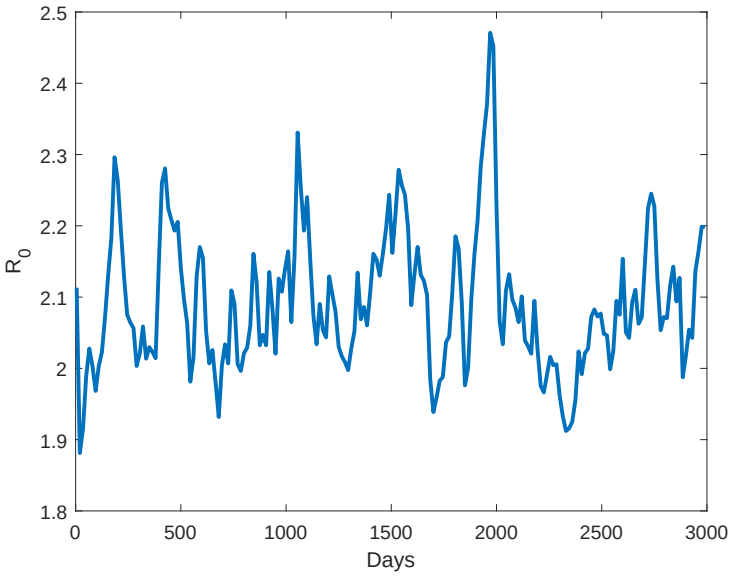
(a) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

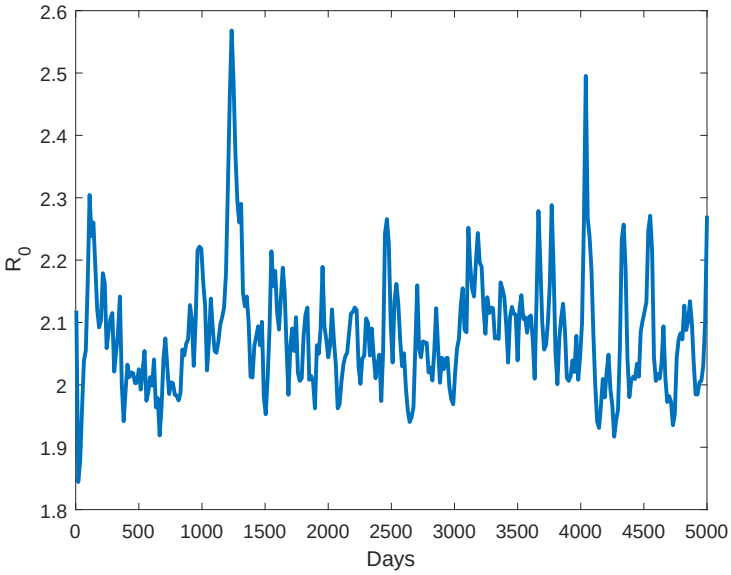


(c) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

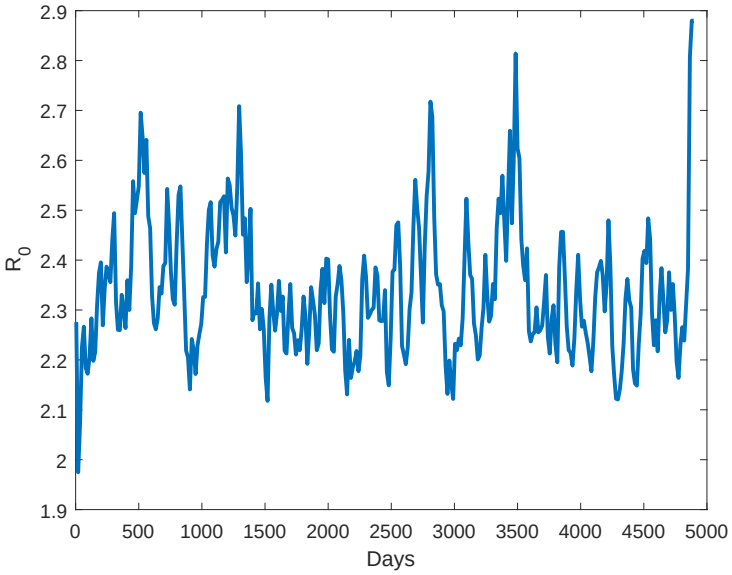


(d) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

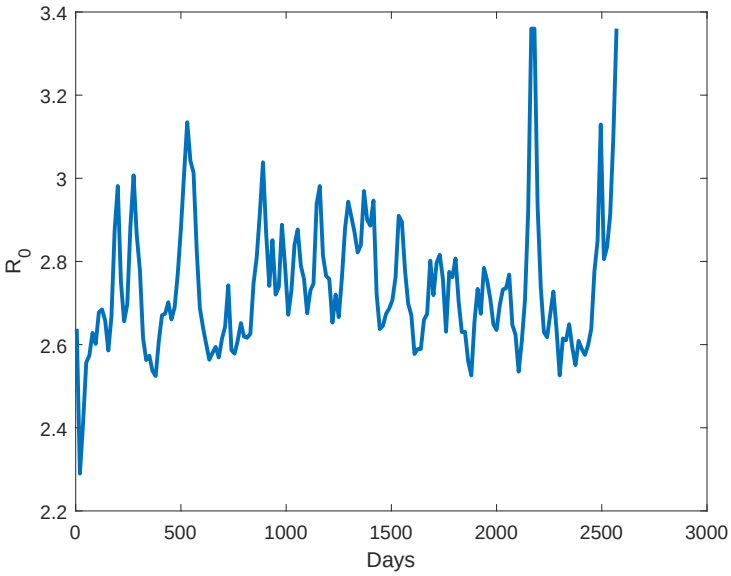
Figure 10: Basic reproduction number R_0 resulting from updates over days for different settings of the measure parameters. Other parameter values at standard values. Bidirectional contact tracing. These figures correspond to those in Figure 8.



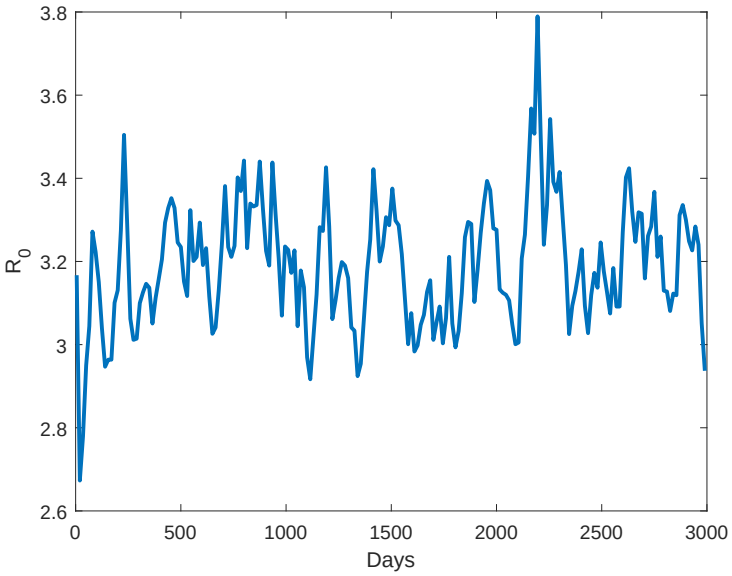
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

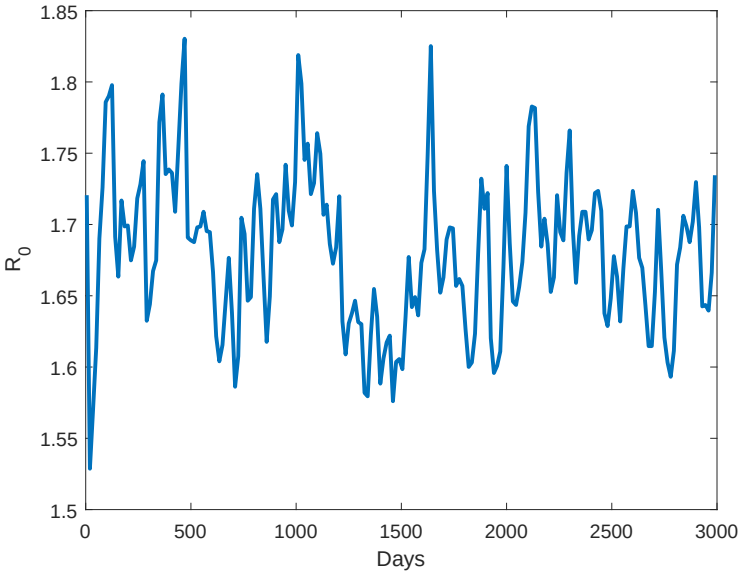


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

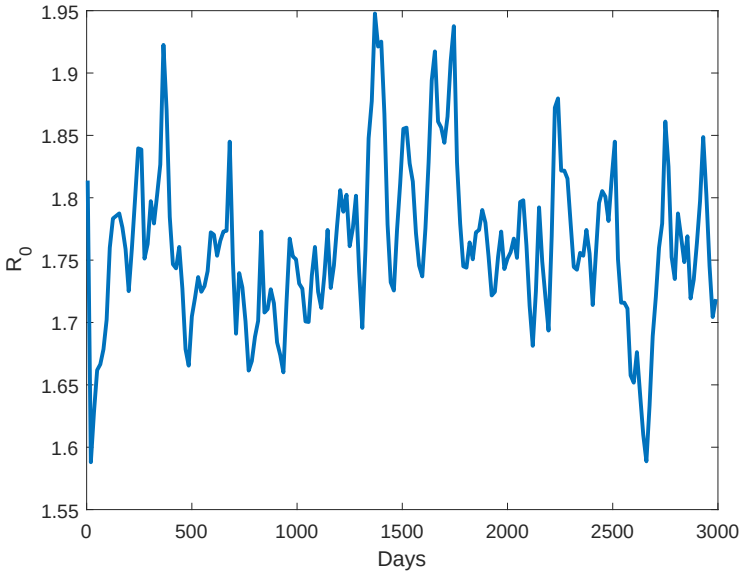


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

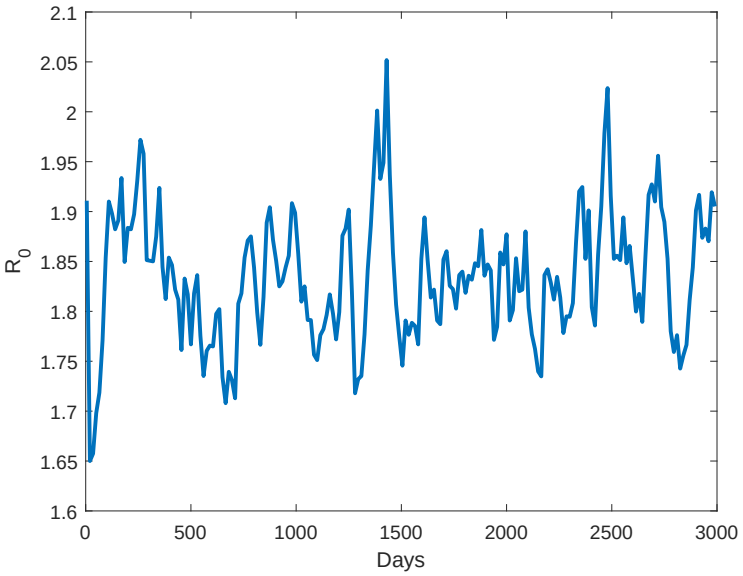
Figure 11: Basic reproduction number R_0 resulting from updates over days for different settings of the measure parameters. Other parameter values at standard values. Bidirectional contact tracing. These figures correspond to those in Figure 9.



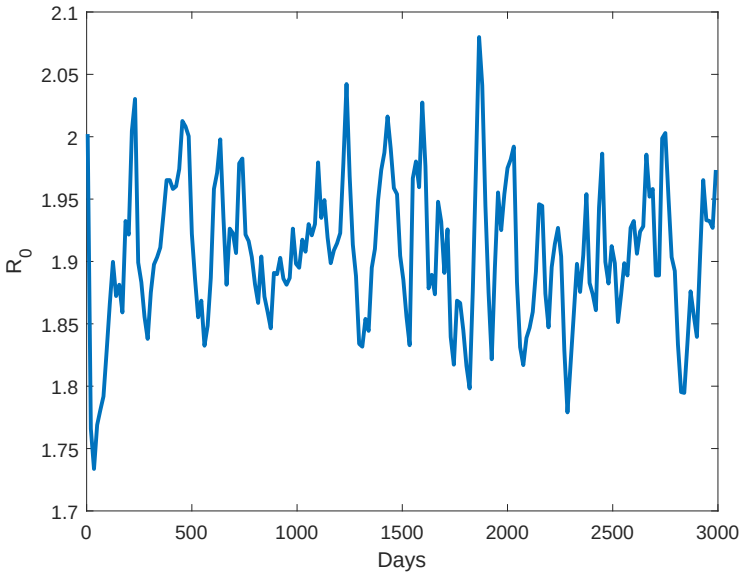
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

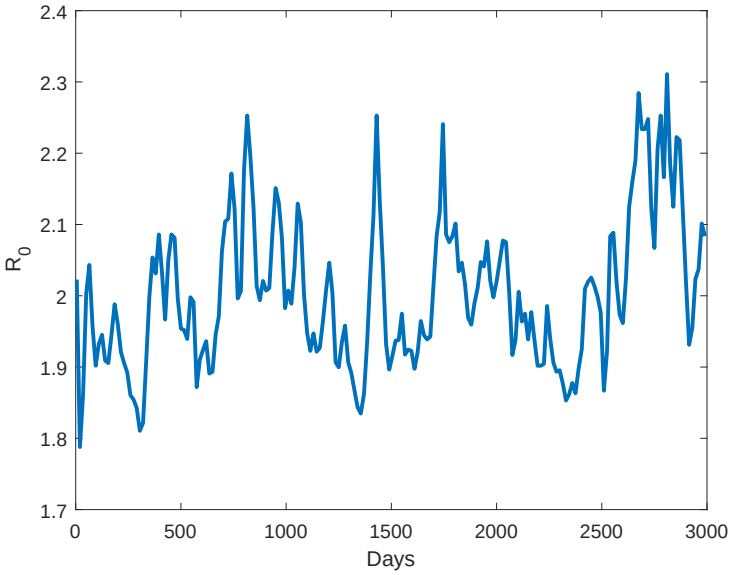


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

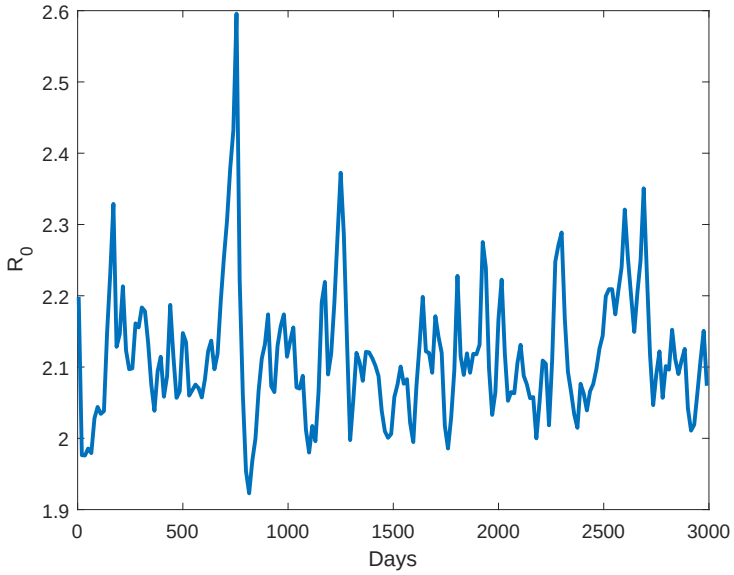


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

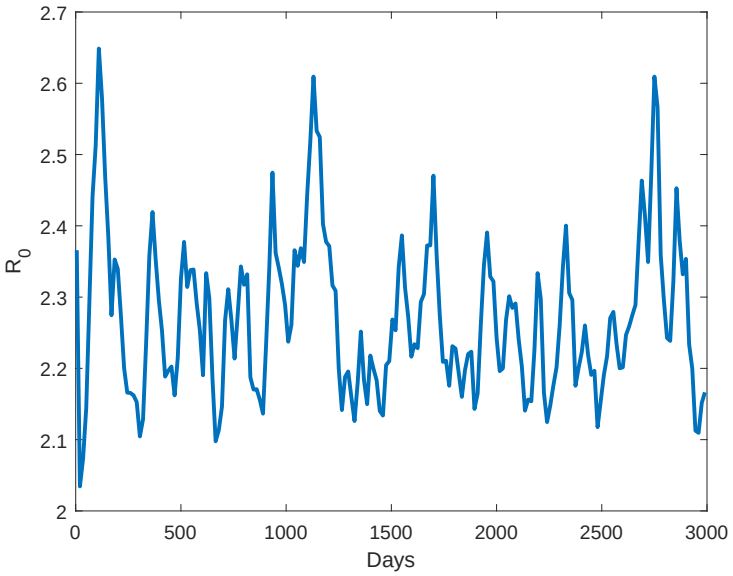
Figure 12: Basic reproduction number R_0 resulting from updates over days for different settings of the measure parameters. Forward contact tracing only. Other parameter values at standard values.



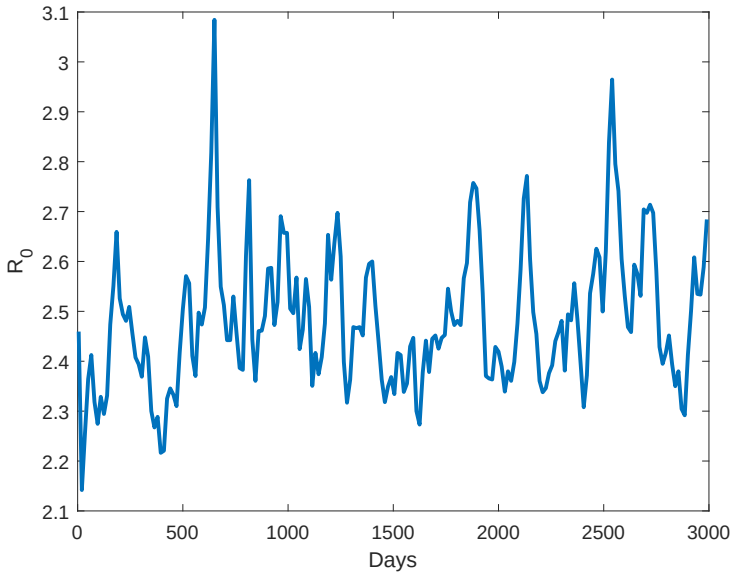
(a) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

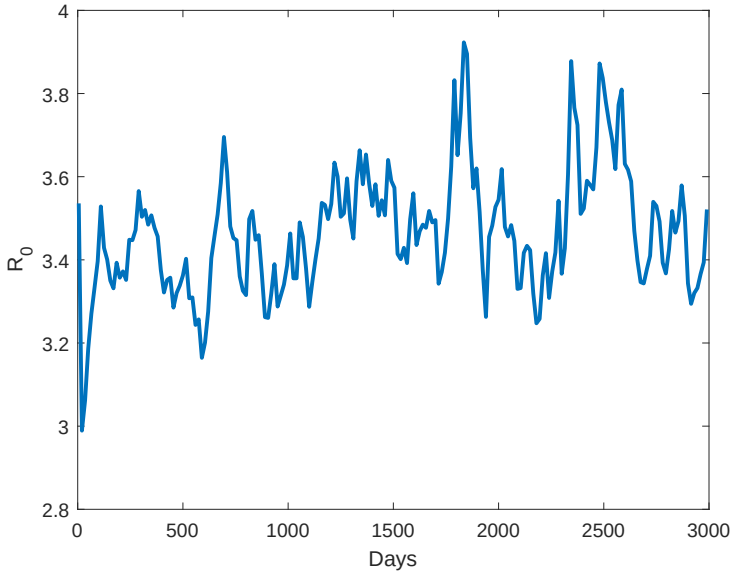


(c) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

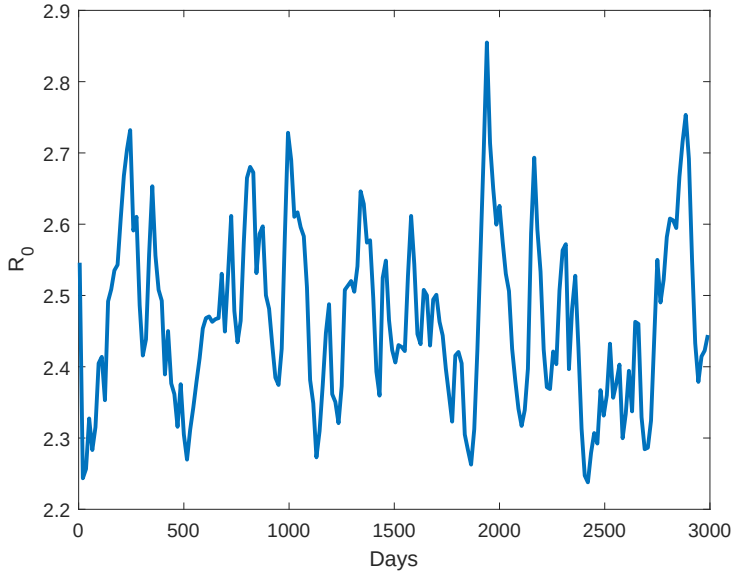


(d) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

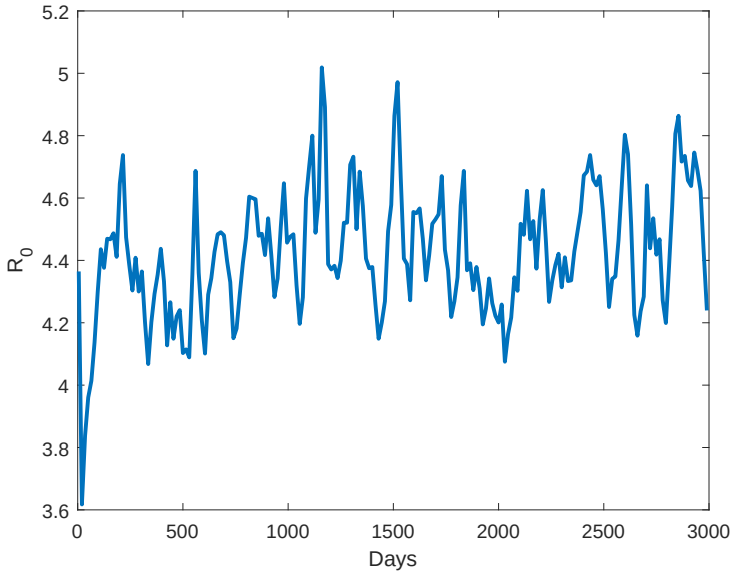
Figure 13: Basic reproduction number R_0 resulting from updates over days for different settings of the measure parameters. Forward contact tracing only. Other parameter values at standard values.



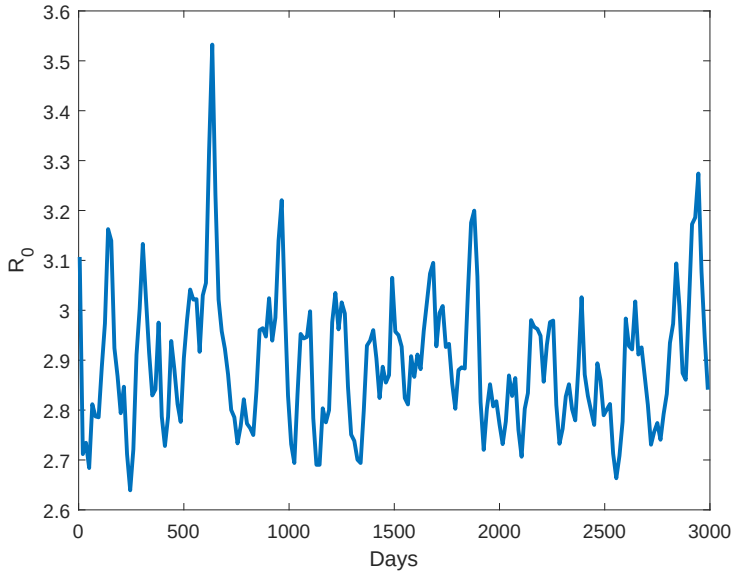
(a) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$. Bidirectional contact tracing.



(b) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$. Forward contact tracing only.

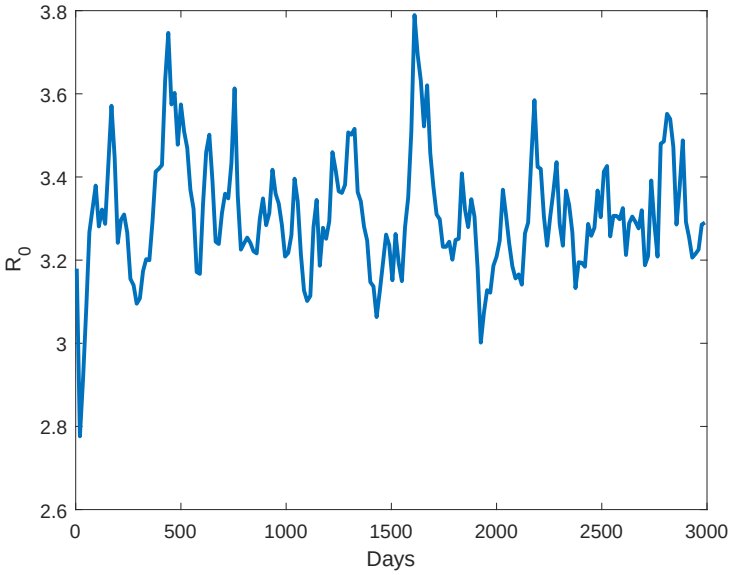


(c) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$. Bidirectional contact tracing.

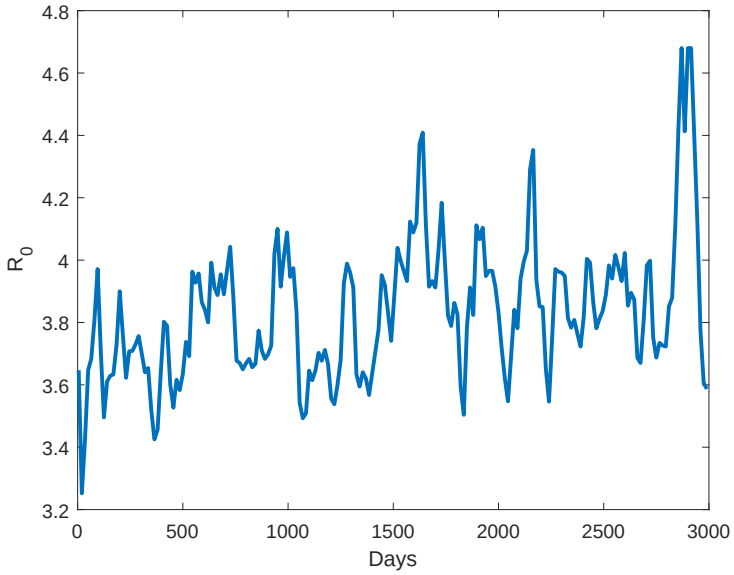


(d) Measure parameters at $\varepsilon = 0.8$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$. Forward contact tracing only.

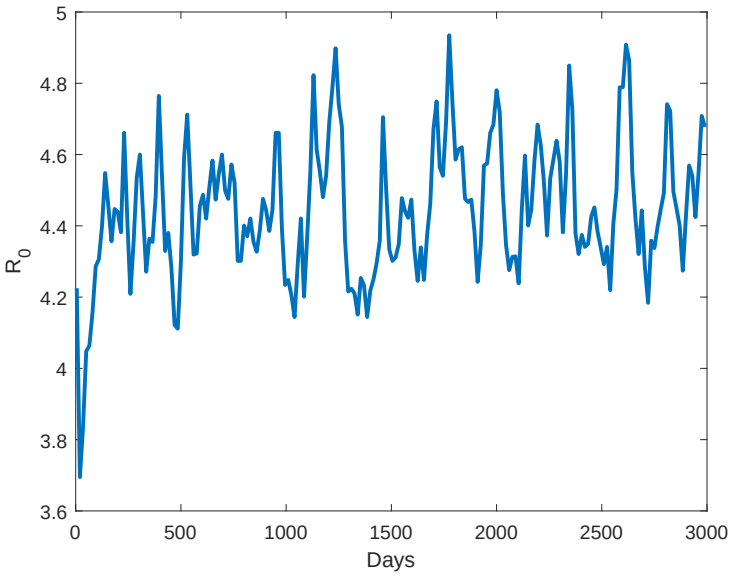
Figure 14: Basic reproduction number R_0 resulting from updates over days for extreme settings of the measure parameters. Other parameter values at standard values.



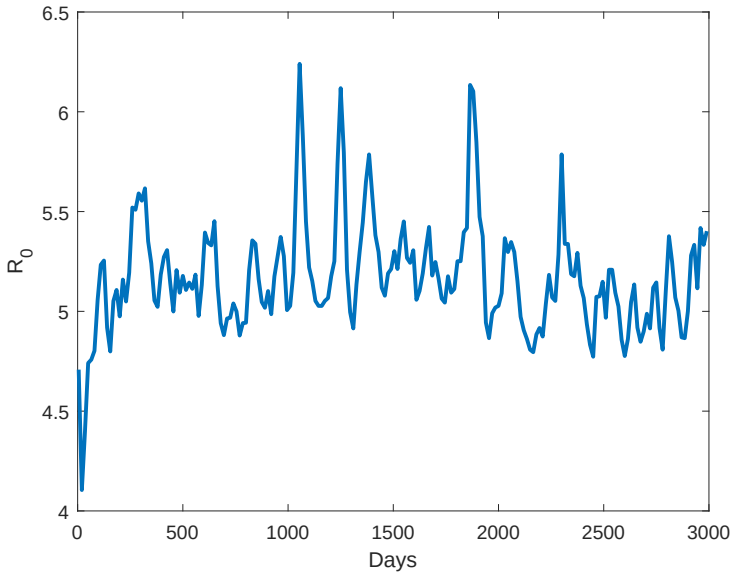
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

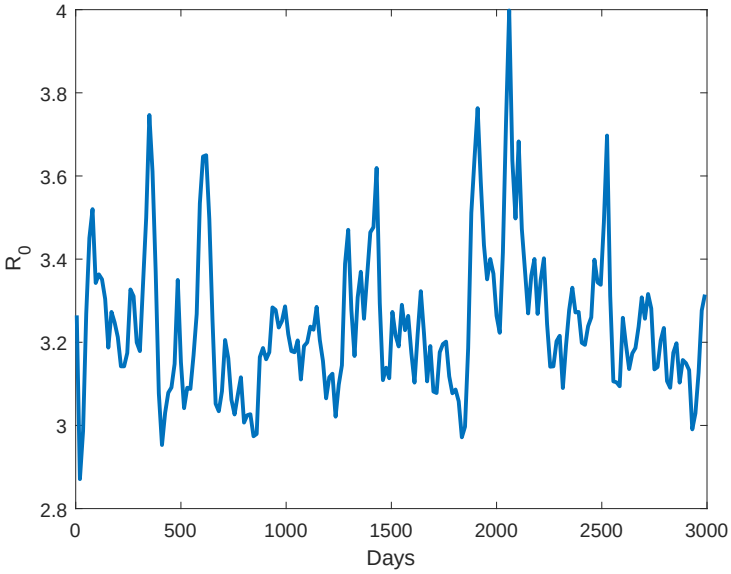


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

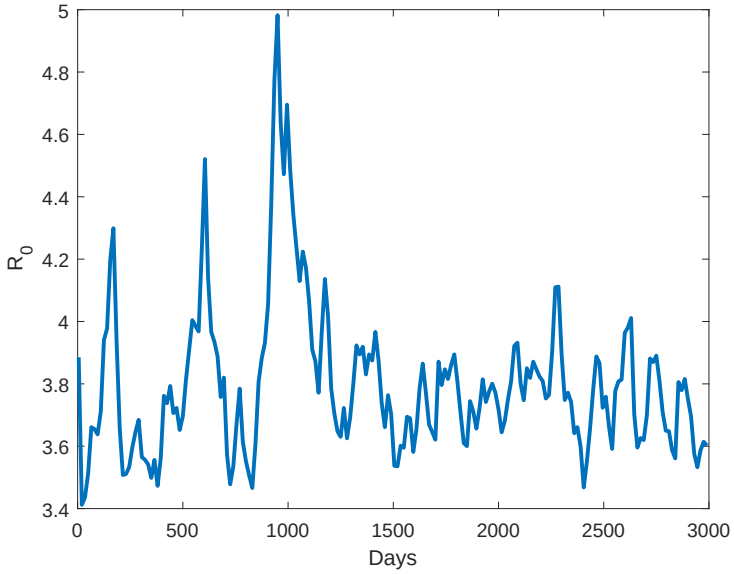


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

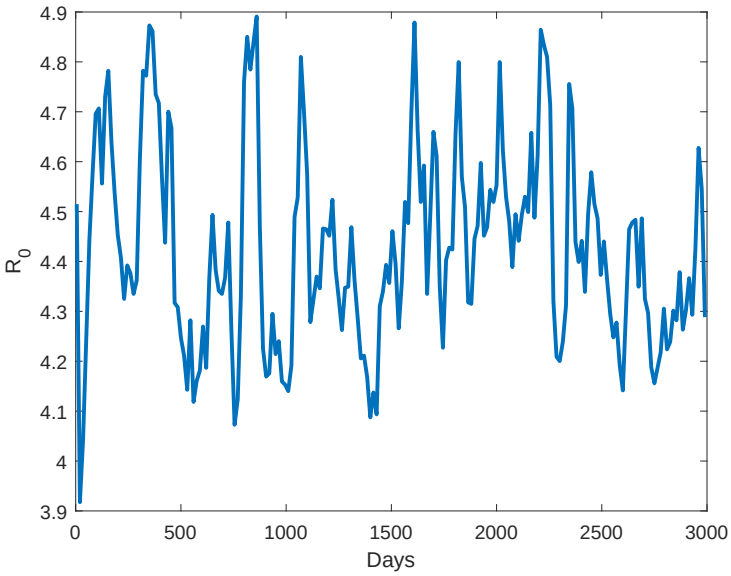
Figure 15: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0$ and $P_{\text{a}} = 0$. Other parameter values at standard values. Bidirectional contact tracing.



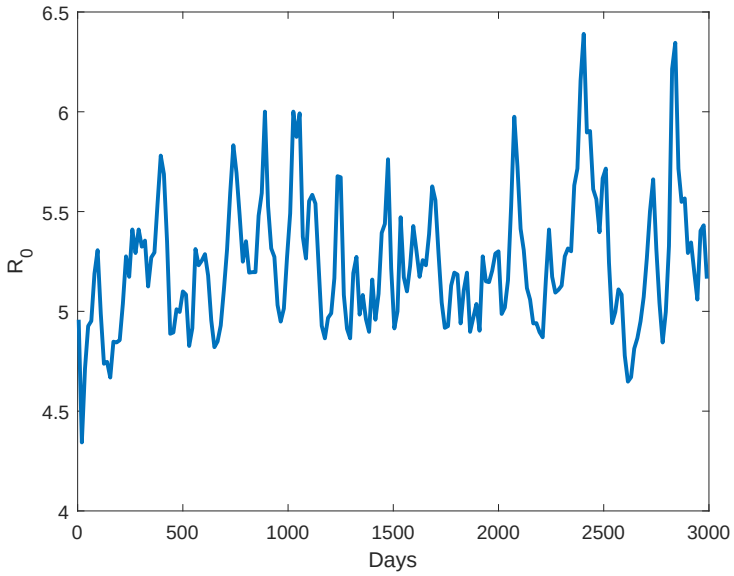
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

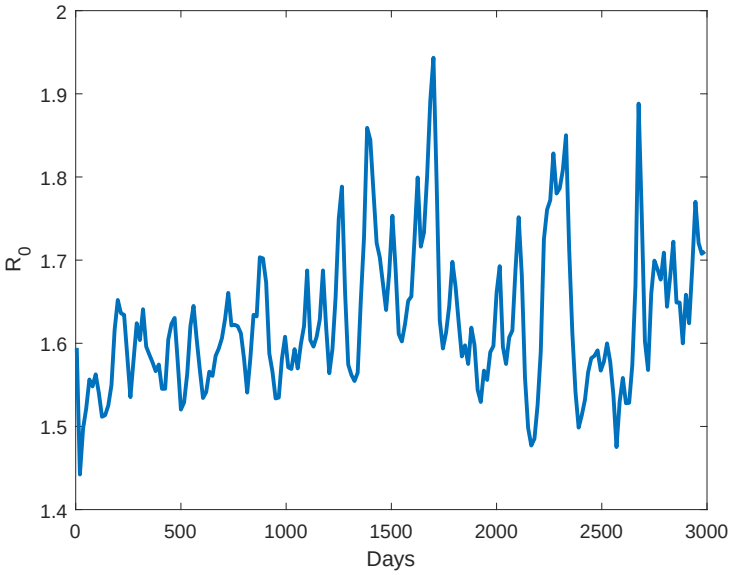


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

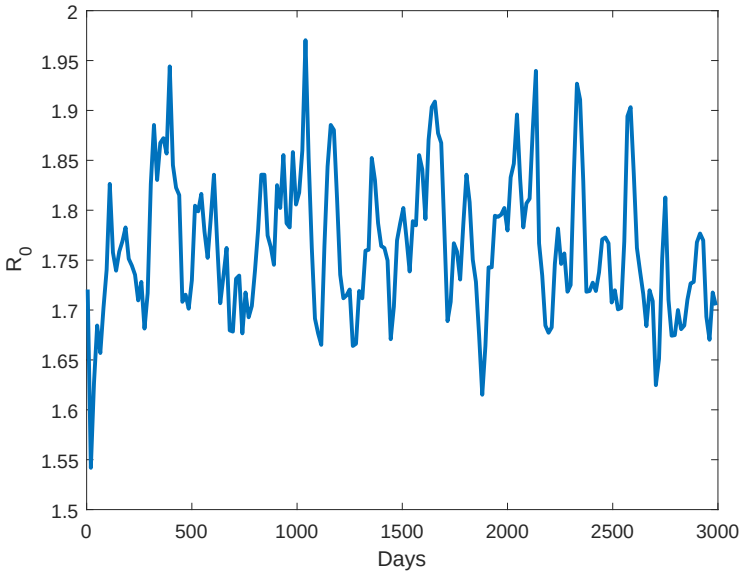


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

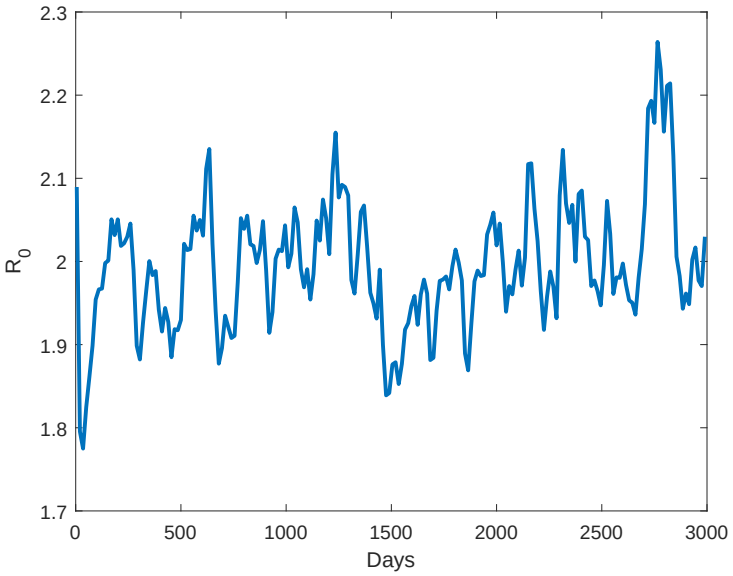
Figure 16: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0.25$ and $P_{\text{a}} = 0$. Other parameter values at standard values. Bidirectional contact tracing.



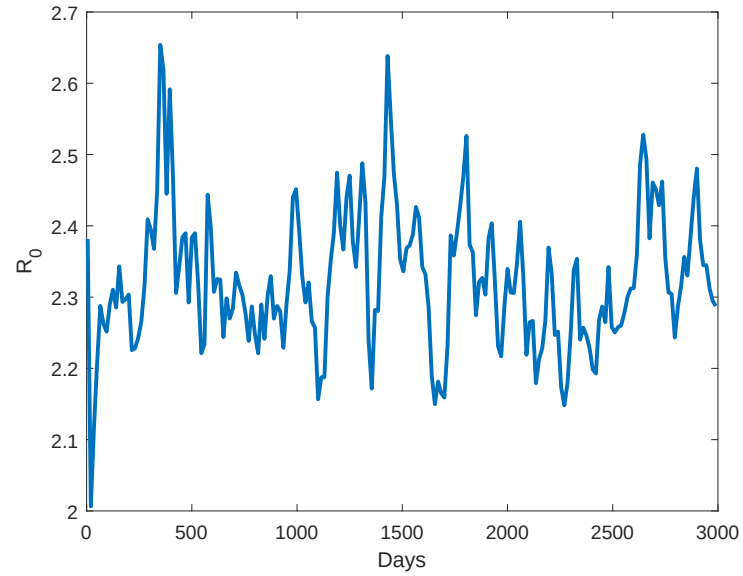
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

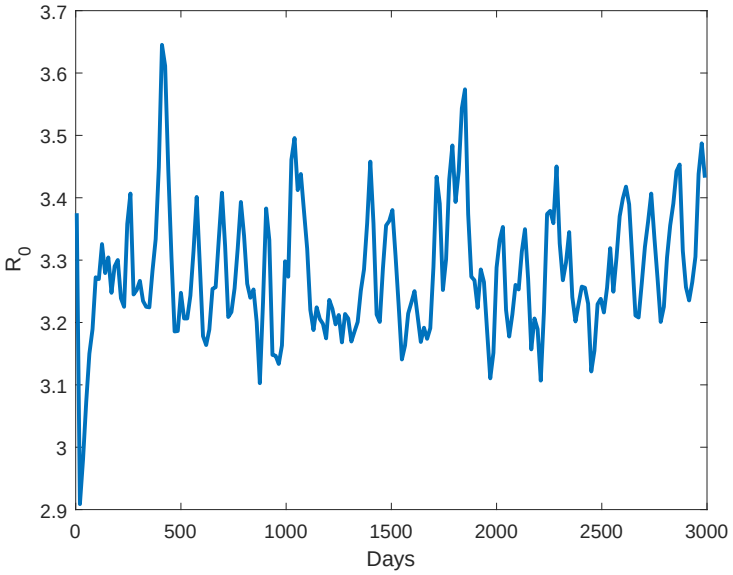


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

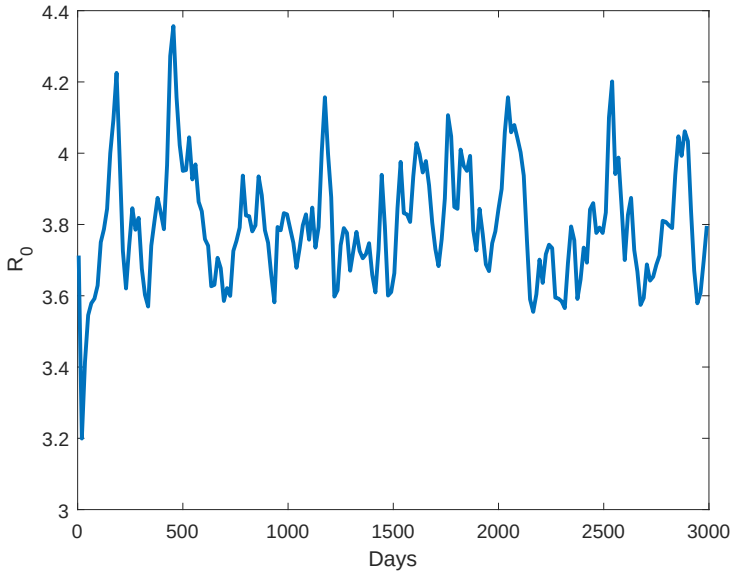


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

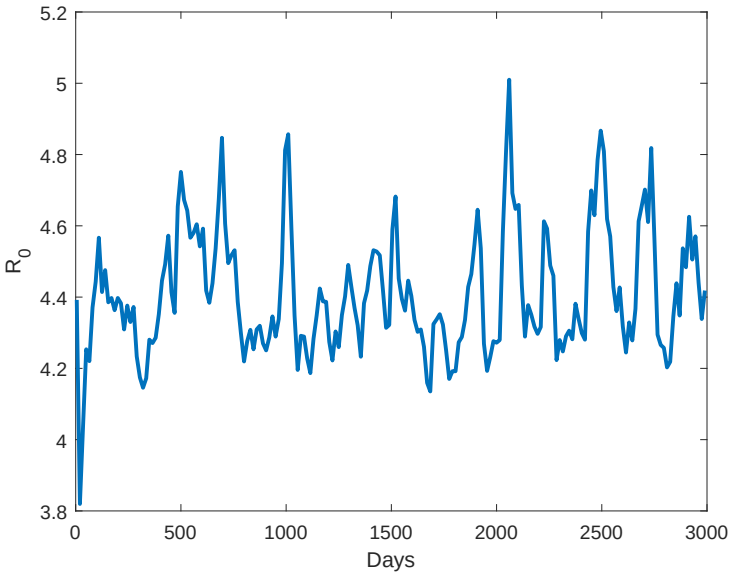
Figure 17: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0$ and $P_{\text{a}} = 0.5$. Other parameter values at standard values. Bidirectional contact tracing.



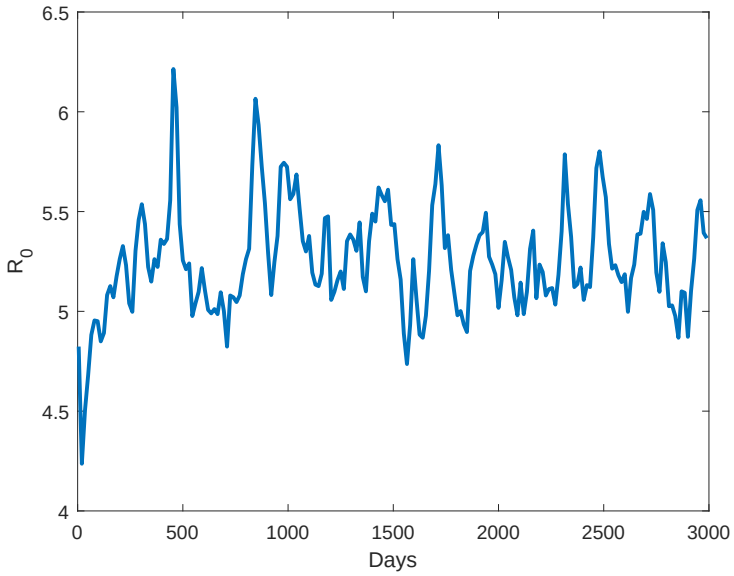
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

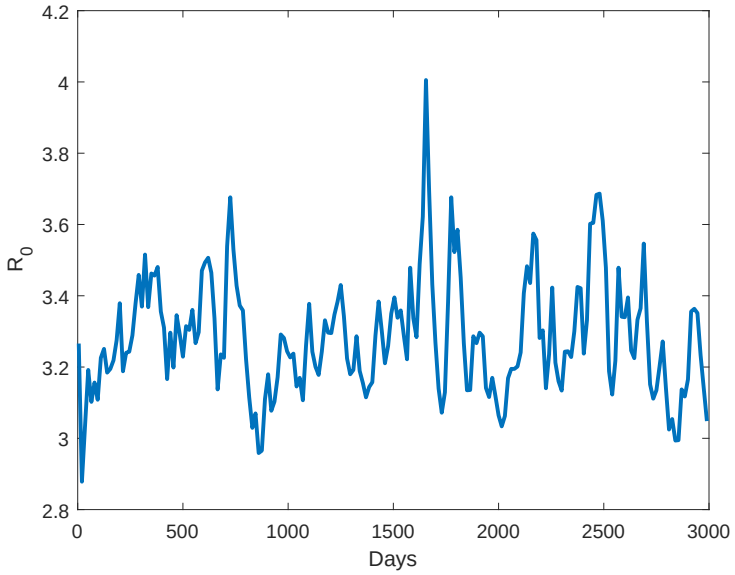


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

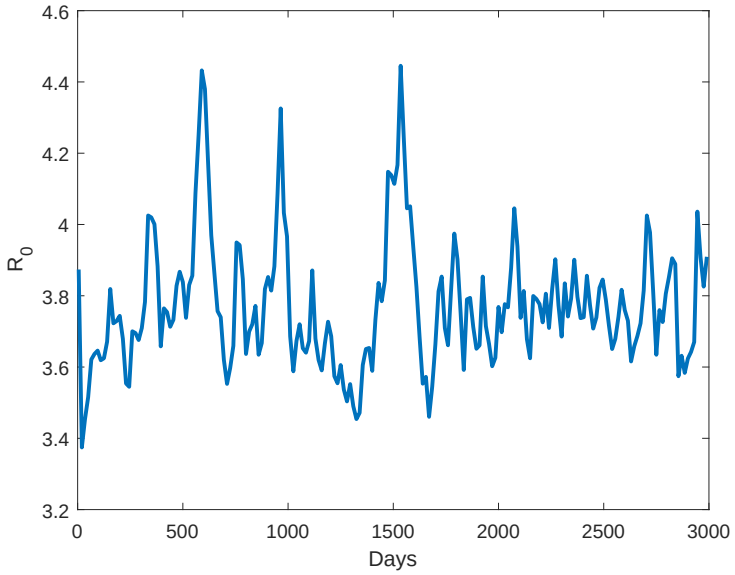


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

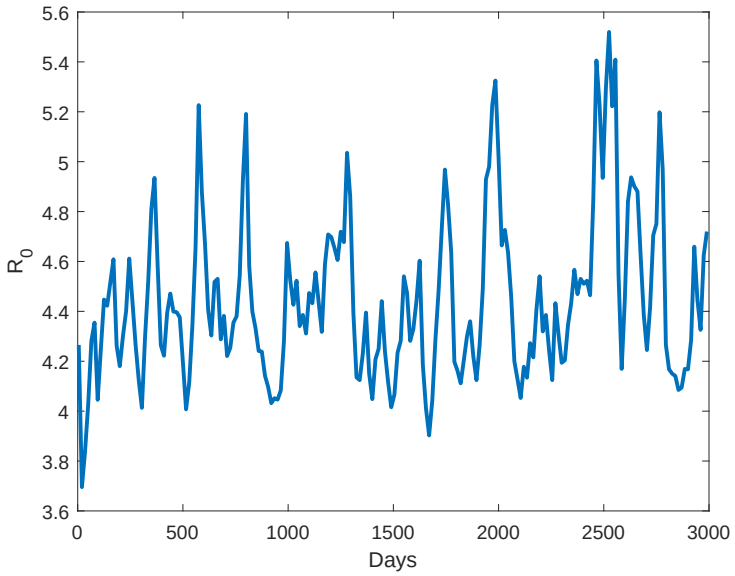
Figure 18: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0$ and $P_{\text{a}} = 0$. Other parameter values at standard values. Forward contact tracing only.



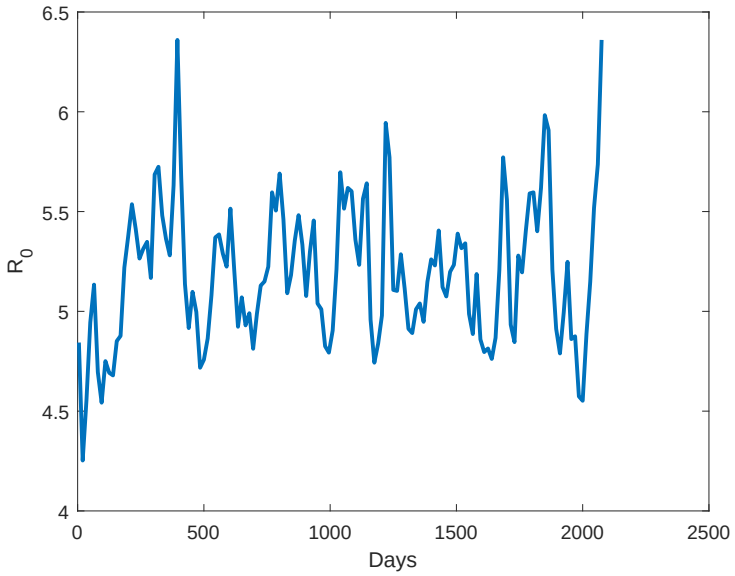
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.

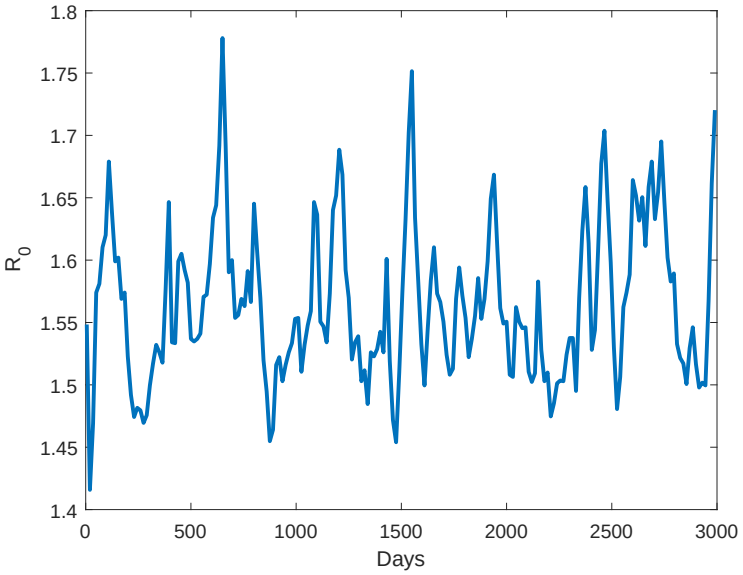


(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.

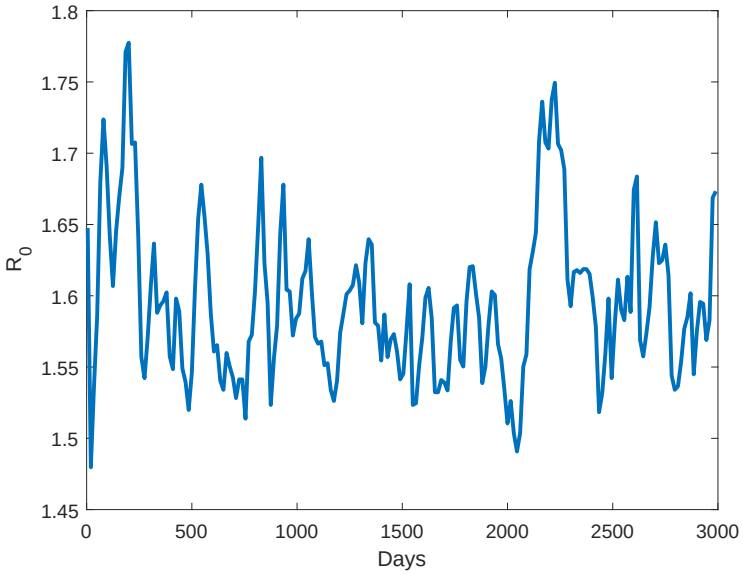


(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

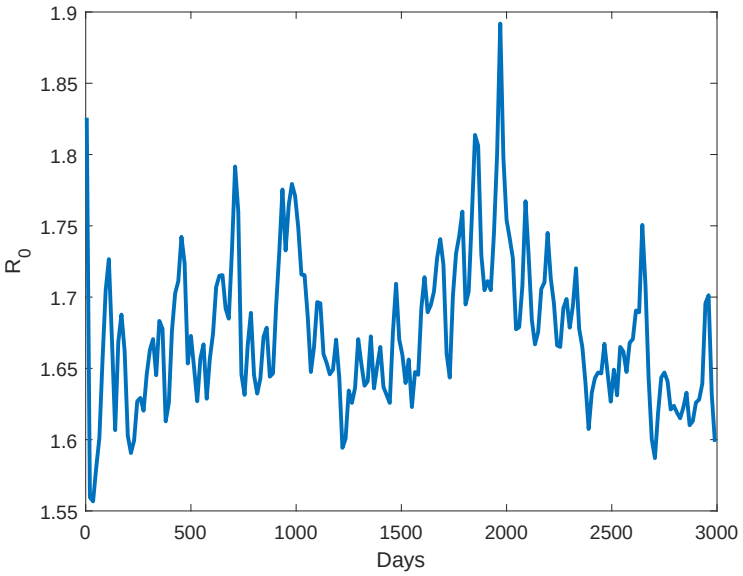
Figure 19: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0.25$ and $P_{\text{a}} = 0$. Other parameter values at standard values. Forward contact tracing only.



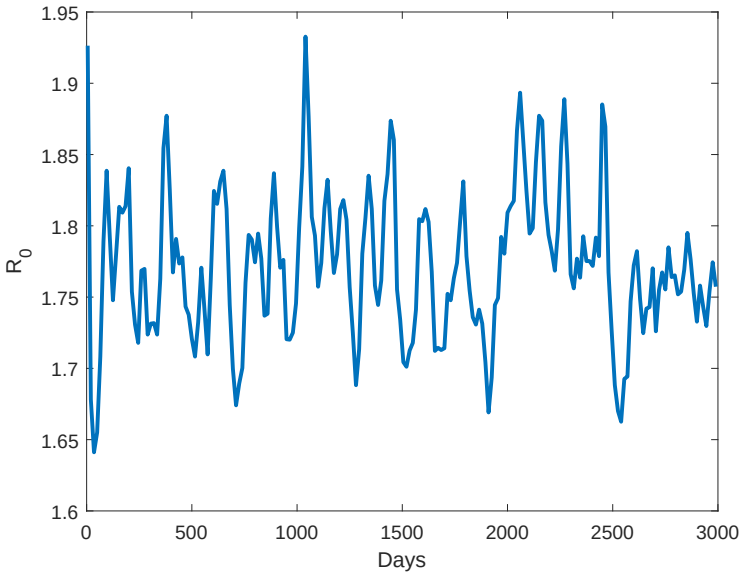
(a) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.5$ and $T_{\text{delay}} = 2$.



(b) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.6$ and $T_{\text{delay}} = 1.5$.



(c) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.7$ and $T_{\text{delay}} = 1$.



(d) Measure parameters at $\varepsilon = 0.95$, $P_{\text{trace}} = 0.8$ and $T_{\text{delay}} = 1$.

Figure 20: Basic reproduction number $R(t)$ resulting from updates over days for different settings of the measure parameters with $P_{\text{s}} = 0$ and $P_{\text{a}} = 0.5$. Other parameter values at standard values. Forward contact tracing only.