# A note on variable susceptibility, the herd-immunity threshold and modeling of infectious diseases

Marcus Carlsson[1*], Jens Wittsten[2], Cecilia Söderberg-Nauclér[3,4]

**1** Centre for Mathematical Sciences, Lund University, Lund, Sweden

**2** Department of Engineering, Borås University, Borås, Sweden

**3** Department of Medicine, BioClinicum, Karolinska Institutet, Solna, Sweden

**4** Department of Neurology, Karolinska University Hospital, Stockholm, Sweden

* Corresponding author

Email: marcus.carlsson@math.lu.se (MC)

## Abstract

The unfolding of the COVID-19 pandemic has been very difficult to predict using mathematical models for infectious diseases. While it has been demonstrated that variations in susceptibility have a damping effect on key quantities such as the incidence peak, the herd-immunity threshold and the final size of the pandemic, this complex phenomenon is almost impossible to measure or quantify, and it remains unclear how to incorporate it for modeling and prediction.

In this work we show that, from a modeling perspective, variability in susceptibility on an individual level is equivalent with a fraction $\theta$ of the population having an "artificial" sterilizing immunity. Given that this new parameter $\theta$ can be estimated, we also derive formulas for $R_0$, the herd-immunity threshold and the final size of the pandemic. In the particular case of SARS-CoV-2, there is by now undoubtedly variable susceptibility due to waning immunity from both vaccines and previous infections, and our findings may be used to greatly simplify models. If such variations were also present prior to the first wave, as indicated by a number of studies, these findings can help explain why the magnitude of the initial waves of SARS-CoV-2 was relatively low, compared to what one may have expected based on standard models.

# Introduction

Since the fundamental works of Kermack and McKendrick [1–3] compartmental mathematical models (such as SIR, SEIR, etc.) are used to model the spread of infectious diseases. Among other things, these papers introduced the famous $R_0$-value and showed that, in contrast with human intuition, a very infectious disease will never infect the whole population, but start to decay at a level called the Herd-Immunity Threshold, for which they deduced the famous formula

$$1 - 1/R_0. \tag{1}$$

However, prior to the SARS-CoV-2 pandemic, there was no reliable data from a novel virus (affecting humans) on which this prediction could be tested. Unfortunately, this remains largely the case, since e.g. lockdowns and voluntary isolation (which the models can not predict) had a major effect on the spread. Despite this, data from places like Sweden, that did very little to stop community transmission, may indicate that the mathematical models have a tendency to overestimate the magnitude of the wave during a major outbreak.

Several factors are known to have a damping effect on model curves. One such example is variable susceptibility, see e.g. Ch. 1 and 3 in [4], and the articles [5–8]. Similar results have also been established numerically for other heterogeneities, such as age and activity [9]. Curiously, variable infectivity (super-spreaders) do not have any effect on the spread during a major outbreak. In any case, such conclusions are derived using heuristic arguments or by simply testing models, and the mechanisms behind these phenomena remain poorly understood. In particular, since variability in susceptibility is virtually impossible to quantify, it is unclear how to efficiently incorporate it into the models, wherefore predictions of future waves continues to be a major challenge. The same statement also applies to the next pandemic.

Concretely, suppose a novel infectious disease, whose transmission dynamics involves high variability in infectivity and/or susceptibility, is introduced in a well connected network like a large city, and suppose a major outbreak is about to take place. One may then estimate $R_0$, i.e. a rough estimate of the average number of new infections that one infective gives rise to, from the data series of early cases, using e.g. EpiEstim [10] or [11]. By a contact tracing study one may also estimate the generation time $T_{generation}$, which is the other parameter needed to run a SIR-model. In such a scenario, one can ask the question if the output of a simple SIR-simulation is a good first order approximation of what is about to come, in the absence of Non-Pharmaceutical Interventions? Is

the formula (1) a good indicator of when we may expect the outbreak to start to recede?

Of the prior studies on this topic, the article that comes closest to answering the above questions is Britton et. al. [9], where the authors prove that variations in activity patterns can significantly lower the herd-immunity threshold, in comparison with the *classical* estimate based on (1). An older publication with a similar message is [12]. In both cases, these are purely empirical observations based on models which are intended to illustrate the general effect of population heterogeneity. This effect is not established mathematically and it remains unclear how, and to what extent, different heterogeneities are manifested. In particular it remains unclear how to more accurately predict the herd-immunity threshold.

In the case of SARS-CoV-2, a number of factors such as genetic, cross-reactive immunity and innate immunity, have been shown to provide variation in susceptibility [13–15]. To what extent these factors affected the spread in reality is a complicated question that goes beyond the scope of this paper. However, it is beyond doubt that we presently have great variations in susceptibility due to waning immunity from previous infections (with different variants of concern) along with vaccinations (with various vaccines).

In this work we prove that, as expected, variations in susceptibility have a damping effect on the model curves, whereas variations in infectivity do not (as long as it is uncorrelated with the former, see [6]). More importantly, we also find that the distribution describing *how* susceptibility varies is not needed for accurate modeling; we show that a susceptibility heterogeneous model will behave almost identically to a standard (homogeneous) SIR-model where a portion of the population is totally immune, and that the precise shape of the susceptibility distribution only influences the value of the level of "immunity". Note that this immunity hence becomes like a mathematical mirage; even if everyone is susceptible to the virus, on a population level it will seem as if a portion of the population have sterilizing immunity. We will refer to such an immunity, needed for accurate mathematical modeling, as "Artificial Sterilizing Immunity" (ASI), and the fraction of the population having it as $\theta$. As long as $\theta$ can be estimated from available data, we show that the actual Herd-Immunity Threshold is indeed lower and given by

$$(1-\theta)\,(1-1/R_0)\,, \tag{2}$$

and the final size of the pandemic is also shrunk by the same factor $(1-\theta)$. We shall also demonstrate numerically that other population heterogeneities, such as those considered by Britton et. al. [9],

have an analogous effect, and hence the findings in this paper can be used to significantly reduce the amount of unknowns in a more realistic heterogeneous model for disease spread.

# 1 The mathematics of infectious disease spread dynamics

In order to explain the mathematical findings to a wider audience, we first give an overview of how the basic SIR-model works. SIR stands for Susceptibles, Infectives and Recovered, and is the simplest form of a "compartmental model" used in mathematical epidemiology [16]. In the model, $S$, $I$ and $R$ are functions of time $t$, and to illustrate how these are related we shall also introduce the (redundant) function $\nu$ describing the incidence, i.e. the amount of newly infected each day (not to be confused with $I$, which describes the prevalence). The formula for $\nu(t)$ is at the heart of the algorithm, and in the beginning we simply have $\nu(t) = \alpha I(t)$, where $\alpha$ is a constant that determines how many new cases an average infective gives rise to during a day. If $a$ is the average number of daily potentially infectious contacts by an average person, and $p$ is the probability that such a contact actually leads to transmission, then $\alpha = ap$.

As the amount of susceptibles gradually decreases, we have to modify this by multiplying with the fraction of the population that is still susceptible. If the total population is $N$ this fraction is $S(t)/N$ and the formula becomes

$$\nu = \frac{\alpha}{N} SI = \frac{ap}{N} SI. \tag{3}$$

To set up the remaining equations we also need the generation time $T_{generation}$, i.e. the average time it takes from infection to recovery. The remaining equations are then

$$\begin{cases} S' &= -\nu \\ I' &= \nu - \sigma I \\ R' &= \sigma I \end{cases} \tag{4}$$

where $\sigma = 1/T_{generation}$ and $'$ indicates differentiation. The equations are intuitively easy to understand, the incidence continuously gets withdrawn from $S$ and added to $I$, and at the same time there is a current of recovering individuals that leave $I$ at a rate $\sigma I$ and appear in $R$ instead.

The SIR-model, and our extensions thereof, are deterministic in the sense that if we run it twice, the output is the same. Such models are believed to work well during major outbreaks, where the law of large numbers applies [4]. All our findings pertain to this situation; for modeling of e.g. the

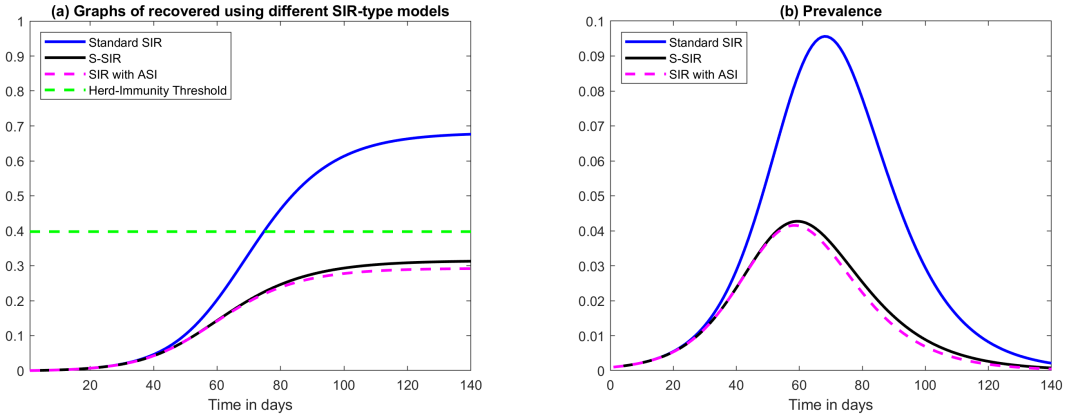



Fig 1: **Graphs of recovered $R$ and prevalence $I$.** (a) Graphs of recovered (as a fraction of the total population) for various SIR-models and a fixed value of $R_0 = 1.66$. First we display standard SIR, then S-SIR and finally SIR with Artificial Sterilizing Immunity (ASI) with parameters from (8). Note that they start out almost identically but that the latter two bend downwards much earlier than the first, which over-shoots the classical Herd-Immunity Threshold (HIT), whereas the second two stay closely together and level out below the classical HIT. (b) Corresponding curves for prevalence $I$.

initial phase or household transmission, other types of models are used.

The most natural initial condition for a new disease is to start with $S(0) = N$ and $R(0) = 0$ (so everybody is initially susceptible and no-one has yet recovered), and set $I(0) = n$ where $n << N$ represents a small number of import cases arriving at time $t = 0$. The value of $n$ is completely irrelevant for the shape of the curves that follow, a low value of $n$ only gives the equation system a slower start so it takes a while longer for the outbreak to reach a certain value. Once this happens, the curves look exactly the same independent of $n$. See Fig. 1 for some typical examples of $R$-curves and $I$-curves. In this model, $R$ is always increasing and levels out on a number which is called "the final size of the pandemic" (see Fig 1a). $S$ approximatively looks like $N - R$, since the prevalence $I$ at any given time is small in comparison with the total population. The incidence $\nu$ typically looks just like $I$, albeit with a lower magnitude.

## 1.1 Contemporary models for COVID-19

Contemporary models used by professional modeling teams usually contain many more compartments than SIR, for instance relating to age stratification, variable activity levels, geographical regions, compartments for people who need ICU and compartments for people that die. It is common to add a compartment $E$ for "Exposed", incorporating the incubation time. However, as we show in the Section 3, this has a limited effect on the overall prevalence and the final size of the pandemic.

Moreover, compartments relating to severely ill and deaths have a marginal effect, simply because only a small fraction of the infected will end up in these compartments. For example, in Section 3 we show that a simple SIR-model with $R_0 = 1.66$ and $T_{generation} = 7$ peaks at a prevalence of 14%, whereas the corresponding SEIR-model peaks at 13% and the age-activity stratified SEIR by Britton et. al. peaks at 9% [9], i.e. a reduction of about 35% compared to SIR. This is consistent with the findings in [9], where a drop in the Herd-Immunity Threshold of around 30% is observed for the age-activity model, comparing with the prediction (1) based on SIR. Hence, we argue, as a first order approximation to more advanced models, a standard SIR or SEIR-model does a decent job.

For example, the model published by members of the Imperial College COVID-19 response team [17] also has at its root a basic SEIR (see p. 9 as well as Fig. S2 in the supplementary material of [17]), and the same goes for the model [18] used by a renowned Swedish modeling team, which managed to fit the ICU occupancy and deaths with high accuracy during the first wave in Sweden. The latter model also takes into account various regions and interaction patterns between these, but the in-region dynamics is basically a simple SEIR.

The list could be made longer, but the point we wish to make is that, since the output of SIR and SEIR are similar, a standard SIR-model is a good rough approximation to more advanced models, and hence understanding the dynamics of SIR, as we focus on in this paper, is valuable also for analyzing more advanced contemporary models.

## 1.2 Model versus reality mismatch?

Whether or not the more advanced models accurately describe the spread of COVID-19 is hard to determine, since one may always argue that Non-Pharmaceutical Interventions (NPI's) as well as voluntary behavioral changes have had a major impact. Without claiming to have a definite answer, the case of Sweden is interesting due to its relaxed strategy, which moreover was kept almost constant during 2020-2021. In particular, schools were kept open, people who could not work from home were encouraged to go to work, family members of infected households were obliged to work or go to school, and widespread face-mask use was never implemented, making the country ideal for comparing models with actual data. Due to insufficient testing, the time series of cases is of limited value, but measurements of sero-prevalence from blood samples give valuable information, since it has been established that most people who get COVID-19 also go on to develop anti-bodies [19], and that these antibodies remain for at least 9 months [20, 21]. Results published by the Swedish

Public Health Agency [22] indicate that roughly 11% had had COVID-19 in the Stockholm region[1] after the first wave 2020, which rose to around 22% in February 2021, following the second wave. Also among hospital staff in Sweden (not using face-mask), the prevalence was around 20% [21] after the first wave, in line with observations from infected households elsewhere [23].

However, the model by Sjödin et. al., referred to earlier, predicts a cumulative number of infected people of around 30% after the first wave, despite assuming a 56% decrease in contacts among people of age 0-59 and a 98% reduction among those aged 60-79[2] (see Fig. 2b, bearing in mind that the Stockholm region has 2.4 million inhabitants). Along the same line, Britton et. al. [9] estimated that the disease could level out at around 43% total infected, in a matter of months. While the authors stress that this is not an actual prediction, it is based on realistic parameters for COVID-19. The famous Report 9 by the Imperial College [24] predicted a total number of 81% infected in a "do-nothing" scenario, based on a more advanced so called "agent based model" that also treats household-contacts separately. According to table 3 in the report, the number of deaths and peak ICU capacity can be reduced by 50% and 81%, respectively, in the most effective NPI-scenario, which certainly goes beyond what was implemented in Sweden. However, as of February 2021, when the original Wuhan-strain was declining [25], these reduced predictions overestimate the actual figure by a factor of roughly 4 (deaths) and 10 (ICU) (when directly translated to Stockholm County).

The point here is not to criticize any particular model, and clearly the case of Sweden can not prove that models are right or wrong, as mentioned initially. However, based on the Swedish numbers compared with the various model outcomes described above, it is a legitimate question whether "contemporary models" have a tendency to overestimate the society spread and final size of the pandemic.

### 1.3 Pre-immunity, super-spreaders and other inhomogeneities

How can we alter the equation system (3)-(4) in order to dampen the curves? The simplest option is to assume that a certain fraction $\theta$ of the population have some form of sterilizing immunity so that they can not get infected by the virus. Mathematically, this is easily achieved by updating the

[1] we focus on the capital region since Sweden is a scarcely populated country and geographical constraints are more likely to have a major impact in less populated areas

[2] this is for scenario d) which accurately fitted ICU-occupancy and death.

initial conditions to

$$\begin{cases} S(0) & = \omega N \\ I(0) & = n \\ R(0) & = 0 \end{cases} \tag{5}$$

where $\omega = 1 - \theta$ is the fraction of initially susceptible. However, this is not very realistic since immunity is usually not binary, i.e. either 100% or 0%. The hypothesis that some people are more susceptible than others is then far more plausible than a binary immunity. In the particular case of SARS-CoV-2, when it was first introduced, the hypothesis that certain individuals had some form of pre-immunity was suggested as an explanation for the, at least for some, unexpectedly mild initial infection waves [26]. This paper lists a number of studies showing that some people had some a priori T-cell immunity. Since then, different articles have demonstrated various mechanisms that make certain individuals more or less susceptible to SARS-CoV-2, e.g. [13–15]. It is also well established that infectivity levels vary dramatically, as mentioned earlier (see e.g. [27]). In addition, this seems uncorrelated to how sick they become; many individuals with very high viral loads are even asymptomatic. In light of this, the most probable assumption is that also the way the virus enters the human is subject to large individual variations.

To make a more realistic model for the spread of COVID-19, or any infectious disease for that matter, it is reasonable to divide the compartments $S$ and $I$ into a number of subcompartments $S_1, \ldots, S_J$ and $I_1, \ldots, I_K$ where people in each compartment have a different level of susceptibility/infectivity. To see how to set up a corresponding equation system for disease spread, recall that $a$ was the amount of daily contacts by one individual. We now let $p_{jk}$ be the probability that such a contact leads to transmission when an individual in $S_j$ meets one in $I_k$. The incidence $\nu_j$ coming from the group $S_j$ then becomes

$$\nu_j = \frac{S_j}{N} \left( a p_{j1} I_1 + \ldots + a p_{jK} I_K \right) \tag{6}$$

(c.f. (3)). The remaining equations in (4) are easily modified to this new vector setting, see SM Sec. 1 for details. In Section 3 we also discuss other extensions such as SEIR and variable activity levels.

## 2 Main results

The main point of this research is that extensions to both SIR and SEIR of the type mentioned above yield overall curves that are only marginally different from basic SIR, given that a level of

Artificial Sterilizing Immunity (ASI) is included. First of all, after setting up the details in Section 1, we prove in SM Proposition 1.1 that the division of $I$ into various sub-compartments have no effect whatsoever. In other terms, the existence of "super-spreaders" do not in any notable way affect the dynamics of disease spread. As mentioned initially, this should not come as a surprise to experts, since similar conclusions are scattered in the literature, see e.g. [4, 8, 28]. Removing this layer of complexity, the equations (6) simplify to

$$\nu_j = \frac{ap_j}{N} S_j I \tag{7}$$

where $p_j$ is the probability of transmission when a susceptible in group $S_j$ encounters an "average" infectious individual. We refer to SM (13)-(15) for the full system of equations, which we label S-SIR for "Susceptibility-Stratified SIR". It is a very curious fact that the division of $S$ into subcompartments can not, in contrast to $I$, mathematically be further reduced to a simpler equation system. However, and this is the key result of this paper, we can prove mathematically that the overall behavior of S-SIR (in terms of prevalence $I$ and recovered $R$) differs only marginally from the basic SIR (3)-(4) upon including ASI to the initial conditions, as we did in (5). This is the essence of Theorem 2.1, which is found in SM Section 2. Given probabilities $p_1, \ldots, p_J$, the theorem also provides formulas for suitable values of the transmission coefficient $\alpha$ (used to compute the incidence $\nu$ in (3)) and artificial sterilizing immunity $\theta$ (used in the initial conditions (5)), as follows:

$$\alpha = a\frac{\sum_{j=1}^{J} w_j p_j^2}{\sum_{j=1}^{J} w_j p_j}, \quad \omega = \frac{\left(\sum_{j=1}^{J} w_j p_j\right)^2}{\sum_{j=1}^{J} w_j p_j^2}, \tag{8}$$

where $\omega = 1 - \theta$ and $w_j$ is the fraction of the population initially belonging to $S_j$; $w_j = S_j(0)/N$. A simple illustration of these results is found in SM Section 1.3. It is important to be careful with the interpretation of $\theta = 1 - \omega$ as a fraction of people who actually have sterilizing immunity, since there is, in reality, not a division of $\theta N$ immune and $\omega N$ susceptible, which is why we have chosen the acronym ASI; *artificial* sterilizing immunity. These results are illustrated in Figures 1 and 2.

We have observed the same phenomenon also when modeling with SEIR and also when including e.g. different age groups and variable activity levels, following [9]; models with many such layers produce output which seem practically indistinguishable from the output of (3)-(5). We briefly discuss these results in Section 3, which we leave as a numerical observation. In particular, given an observed level of ASI $\theta$ in a society, it is mathematically impossible to draw any conclusions

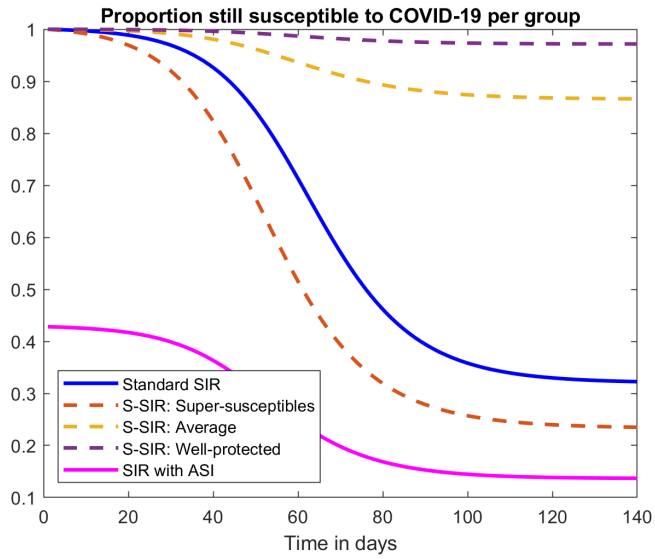



Fig 2: **Graphs of susceptibles.** $S-$curves corresponding to the 3 graphs in Figure 1. The S-SIR model has three subgroups $S_1$, $S_2$, $S_3$ corresponding to $p_1 = 1$, $p_2 = 0.1$ and $p_3 = 0.02$ (labeled "super-susceptibles", "normal" and "well-protected"). Note how the spread in the latter two sub-groups level out as soon as it levels out in the super-susceptible group. This rather unexpected behavior is typical. We also remark that the SIR with ASI starts out with 57% immune, which is why the initial value for the pink curve is 43%. This number was chosen using the formula (8).

about how much of $\theta$ is caused by inhomogeneities in age and behavior, and how much comes from variations in susceptibility.

Incidentally, at the end of each paper [1–3], Kermack and McKendrick stress that a weakness in their model is that they assume uniform susceptibility, which they consider unrealistic in many cases. However, it seems that they never got around to address this issue, and we have not found a rigorous mathematical analysis of how to deal with this situation elsewhere in the literature either. In particular, the formula $1 - 1/R_0$ for the Herd-Immunity Threshold (HIT), which stems from their seminal papers, may very well be inaccurate, as demonstrated numerically in [9]. In the coming section we derive a refined version of this formula taking ASI into account.

## 2.1 Formulas for $R_0$ and the Herd-Immunity Threshold

It is easy to see that the generation time $T_{generation}$ (introduced below (3)) coincides with the average time an infected individual remains infective. Since $\alpha$ is the infection rate, we conclude that $R_0 = \alpha T_{generation}$ for the standard SIR (3)-(4), assuming a fully susceptible population. However, in the presence of ASI $\theta$, the actual infection rate is only $(1 - \theta)\alpha$ and hence the correct formula for

the $R_0$-value

$$R_0 = (1-\theta)\alpha T_{generation} = \omega\alpha T_{generation}. \tag{9}$$

The above value for $R_0$ is the value that would be estimated by e.g. EpiEstim [10] or [11] from a real time series generated by the model (3)-(4) with initial data (5).[3] Mathematically, $R_0$ is defined as the number of new infections that one infected individual gives rise to, before disease induced immunity starts to build up. (To compute this, first solve $I'(t) = -\sigma I(t)$, given $I(0) = 1$, recalling that $\sigma = 1/T_{generation}$, and then integrate the resulting incidence $\nu$, as given by (3), while keeping $S(t)$ fixed at $S(0) = \omega N$.) Similarly, one sees that the effective $R$-value, denoted $R_e(t)$, in the above model is

$$R_e(t) = \frac{S(t)}{N}\alpha T_{generation} = \frac{S(t)}{S(0)}R_0.$$

The term "herd-immunity" carry a variety of meanings [29]. In mathematical epidemiology, given a certain model and a novel virus, the Herd-Immunity Threshold is defined as the total number of infective and recovered needed to achieve $R_e(t_0) = 1$. Since $I'(t) = (R_e(t) - 1)\sigma I(t)$, (recall (4)), we see that this coincides with the point at which the wave of infectious naturally starts to recede. Under such circumstances the disease will naturally die out and, assuming the disease has been temporarily eradicated, any import cases will not spark new outbreaks. For an new outbreak to a known virus, e.g. caused by a mutation or reintroduction in combination with waning immunity, the herd-immunity threshold signifies the proportion that need to get infected during the new wave, in the absence of NPI's, before the disease start to decline. We denote this value by $H_{IT}$.

In the SIR-model, it is assumed that individuals mix homogeneously and that recovered individuals have protective antibodies. While it is known that anti-bodies wane over time, at least for SARS-CoV-2, this waning happens much more slowly than the duration of an outbreak [20], and hence this assumption is reasonable for the discussion of the herd-immunity threshold. However, we wish to stress that the waning means that herd-immunity is never a stable condition, but will fade with time.

Assume now that a SIR-model with a certain level of ASI accurately describes a given outbreak. The Herd-Immunity Threshold $H_{IT}$ then equals $S(0)/N - S(t_0)/N$ where the latter is given by solving $R_e(t_0) = 1$ (i.e. the difference between the fraction $S(t_0)/N$ susceptible at the time $t_0$ when herd-immunity is reached, and the fraction of susceptibles initially). In the SIR-model with ASI,

[3] up to a minor numerical error caused by the discrepancy between theory and practice

solving $R_e(t_0) = 1$ yields the equation $S(t_0)/N = 1/\alpha T_{generation}$, and so we deduce

$$H_{IT} = \omega - \frac{1}{\alpha T_{generation}} = \omega(1 - 1/R_0),$$

where we used the earlier formula (9) as the definition of $R_0$. This is the formula for the Herd-Immunity Threshold presented in equation (2) in the introduction. It implies that the classical formula (1), given an estimate of $R_0$ from e.g. EpiEstim, is over-estimating the Herd-Immunity Threshold. More importantly, it allows to predict $H_{IT}$, given that the ASI parameter $\theta = 1 - \omega$ can be estimated from available data.

That the classical formula may be misleading has been pointed out before [12], and a more recent contribution indicating that the $H_{IT}$ could be significantly lower than the value (1) is [9]. These works illustrate this by simply testing models that involve heterogeneities (primarily social mixing patterns, not variable susceptibility), and therefore it offers little guidance for actual estimation of $H_{IT}$. Formula (2) is, to our knowledge, the first time this effect has been mathematically proven and quantified.

To sum up, the results in the previous section imply that a Susceptibility-stratified SIR will behave like SIR with ASI, and the above formula hence applies to this model as well, with $\omega$ given by (8). It is crucial to note that this applies under the assumption that the immunity is achieved by natural spread. The herd-immunity threshold for vaccinating is still given by the classical formula (1) (assuming the vaccine gives sterilizing immunity). This is shown in SM Section 1.2.

## 2.2 Damping and the final size of the pandemic

As mentioned earlier, several works have established that variable susceptibility have a damping effect on the prevalence. By the above results, this can now can be quantified. Suppose $(S_0, I_0, R_0)$ is a solution to SIR in a homogenous and fully susceptible population (so $S(0) = N$), and let $\alpha_0$ be the corresponding transmission rate. Given a fixed value of ASI $\theta$, it is then easy to see that $(S, I, R) = (\omega S_0, \omega I_0, \omega R_0)$ is a solution to (3)-(5), where $\omega = 1 - \theta$ and $\alpha = \alpha_0/\omega$. Hence the effect of ASI is really nothing but a rescaling of standard SIR curves. Due to formula (9), the $R_0$-value equals $\alpha_0 T_{generation}$ in both cases.

It is well known that the final size of the pandemic $\pi_0 = R_0(\infty)/N$ in the usual SIR (as well as SEIR) is given by solving $1 - \pi_0 = e^{-R_0\pi_0}$ (see [8] and Chapter 3 of [4]). Combining this with the

above we see that the final size of the pandemic $\pi$ in SIR with ASI is given by solving

$$1 - \pi/\omega = e^{-R_0\pi/\omega}.$$

Hence, in combination with our main result about reduction of Susceptibility-Stratified SIR to SIR with ASI, we deduce that the above solution $\pi$ is a good approximation to the final size of the pandemic for S-SIR with $\omega$ given by (8).

# 3 Extension to more general models

For a disease like COVID-19, with a short incubation period followed by an even shorter infectious period, there is only a marginal difference between modeling using SIR and using SEIR, and hence we believe that the key conclusions of this paper extend to this model as well. Similarly, we have found numerically that more advanced SEIR-models taking variable age and activity levels into account, behave just like SIR if we incorporate ASI. We leave the formal verification of these observations as an open conjecture, and content ourselves with showing some examples.

## 3.1 SEIR

SEIR has two key parameters apart from $R_0$, namely $T_{infectious}$ and $T_{incubation}$, where the former is the average time that a person is infectious and the latter is the time from when a person becomes infected until he or she becomes infectious. Estimates for these vary, we here follow Britton et. al. [9] and set $T_{incubation} = 4$ and $T_{infectious} = 3$. It then follows that the *generation time* equals

$$T_{generation} = T_{infectious} + T_{infective} = 7,$$

where the generation time is the average time it takes from a person getting infected until that person infects others (see equation (5) in [25]). Note that this is consistent with the choice of $T_{generation}$ in previous sections.

The reason why SEIR and SIR give almost identical output for COVID-19 is that both are primarily determined by the values of $T_{generation}$ and $R_0$. To wit, during a major outbreak, it does not matter if a person is sick for 7 days and infect $R_0$ people during those 7 days, or if he undergoes incubation for 4 days and then infect $R_0$ people during the remaining 3 days. As an example, consider Fig. 3 (a); we see a very similar behavior by choosing parameters for SIR and SEIR in

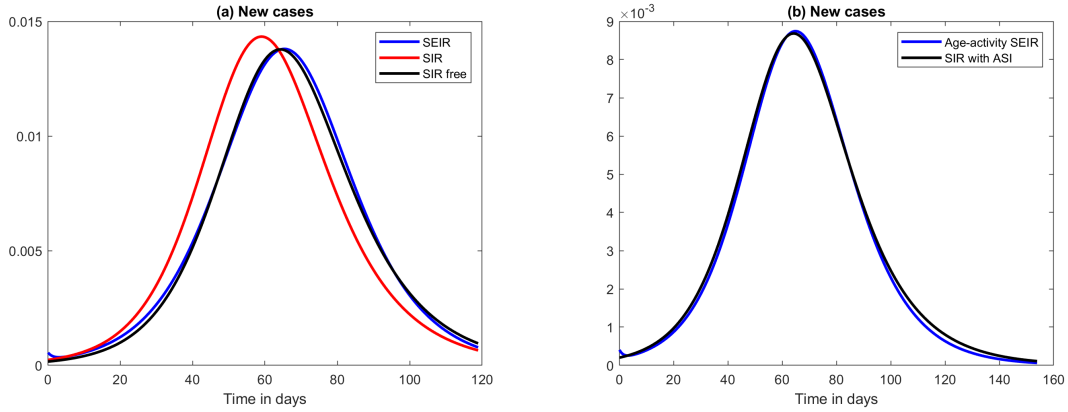



Fig 3: **Approximations using SIR with ASI.** (a) SEIR with $R_0 = 1.66$ and $T_{infectious} + T_{infective} = 7$, SIR with the same $R_0$ and $T_{generation} = 7$ and finally SIR with a 1% lower $R_0$, same $T_{generation}$ and a lower initial amount of infectives. (b) Age-activity stratified SEIR with $R_0 = 1.66$ and $T_{infectious} + T_{infective} = 7$ (blue); SIR using the same $T_{generation}$ but an ASI of 25% and slightly different $R_0$ (black).

accordance with the above formulas (with $R_0$ fixed). Moreover, by allowing free parameters, SIR can be made to behave almost identically as SEIR (even without involving ASI). Since the exact value for the input parameters are unknown in reality, we argue that it is irrelevant whether one uses SIR or SEIR. Therefore, the observations of this paper should extend to SEIR as well, even if we have not been able to establish this mathematically.

## 3.2 Heterogeneous models

Variable susceptibility is not the only type of population heterogeneity which could manifest itself as ASI on a macro level. In [9] the authors develop a heterogeneous SEIR model taking variable interaction pattern between different age-groups into account, as well as the fact that people in each age-group have varying amount of contacts. We implemented their model and then sought parameters for SIR, allowing for ASI, that would yield a similar output. The result is seen in Figure 3 (b). Again, the difference is so fine that it would be impossible to spot in practice. Henceforth, what may appear as a certain level of population (pre-)immunity in mathematical models may in fact be a mix of various population heterogeneities, in which variable susceptibility is only one ingredient.

# 4 Discussion

There could be many reasons for why certain people are more susceptible than others to infection by a novel virus, ranging from innate and adaptive immunity to cross-reactive immunity from other known viruses as well as genetic differences. For a novel disease, sterilizing pre-immunity, i.e. individuals which are completely immune to without ever having had the virus, most likely does not exist. The key point of this study is that sterilizing individual immunity is not needed in order to observe what looks like sterilizing immunity on a population level, which we have coined ASI; artificial sterilizing immunity. We show mathematically that, in order to have ASI, we only need moderate variation in susceptibility. Moreover, we demonstrate numerically that other types of population heterogeneities, such as variable social mixing patterns, also manifest themselves as ASI. The findings in this paper do not limit themselves to SARS-CoV-2, but basically shows that classical formulas for the herd-immunity threshold and the models for spread of infectious diseases with roots in the famous paper by Kermack and McKendrick [1] are inapt to model any infectious disease subject to large variability in susceptibility, and need to be modified as described in Section 2.1.

The estimation of the herd-immunity threshold $H_{IT}$ is crucial for efficient disease control management and planning. For example, if a society decides to make a lock-down before $H_{IT}$ is reached, it is almost certain that the disease will re-emerge, unless strict control of import cases is implemented. The classical formula (1) is still very much in use, despite the fact that it is known to rely on a number of oversimplifying assumptions which may lead to an erroneous indication. We have established a new formula which we prove applies when variable susceptibility is present. Since we show that our simplified model, SIR with ASI, also seems to be a good substitute for models that involve variable social mixing patterns, it is possible that (2) applies more generally than what we are able to prove mathematically.

# Supporting information

**S1 Supplementary Material.**

## Acknowledgments

We thank Erik Wahlén for fruitful discussions. The research of CSN was supported by the Swedish Medical Research Council (2019-01736).

## References

1. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proceedings of the royal society of London Series A, Containing papers of a mathematical and physical character. 1927;115(772):700–721.

2. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics II. The problem of endemicity. Proceedings of the Royal Society of London Series A, containing papers of a mathematical and physical character. 1932;138(834):55–83.

3. Kermack WO, McKendrick AG. Contributions to the mathematical theory of epidemics III. Further studies of the problem of endemicity. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character. 1933;141(843):94–122.

4. Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious disease dynamics. In: Mathematical Tools for Understanding Infectious Disease Dynamics. Princeton University Press; 2012.

5. Gerasimov A, Lebedev G, Lebedev M, Semenycheva I. COVID-19 dynamics: a heterogeneous model. Frontiers in Public Health. 2021;8:911.

6. Hickson R, Roberts M. How population heterogeneity in susceptibility and infectivity influences epidemic dynamics. Journal of Theoretical Biology. 2014;350:70–80.

7. Miller JC. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. Physical Review E. 2007;76(1):010101.

8. Miller JC. A note on the derivation of epidemic final sizes. Bulletin of mathematical biology. 2012;74(9):2125–2141.

9. Britton T, Ball F, Trapman P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science. 2020;369(6505):846–849.

10. Thompson R, Stockwin J, van Gaalen RD, Polonsky J, Kamvar Z, Demarsh P, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;29:100356.

11. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. American journal of epidemiology. 2013;178(9):1505–1512.

12. Fox JP, Elveback L, Scott W, Gatewood L, Ackerman E. Herd immunity: basic concept and relevance to public health immunization practices. American journal of epidemiology. 1971;94(3):179–189.

13. Dee K, Goldfarb DM, Haney J, Amat JA, Herder V, Stewart M, et al. Human rhinovirus infection blocks SARS-CoV-2 replication within the respiratory epithelium: implications for COVID-19 epidemiology. Journal of Infectious Diseases. 2021.

14. Ng KW, Faulkner N, Cornish GH, Rosa A, Harvey R, Hussain S, et al. Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. Science. 2020;370(6522):1339–1343.

15. Zeberg H, Pääbo S. A genomic region associated with protection against severe COVID-19 is inherited from Neandertals. Proceedings of the National Academy of Sciences. 2021;118(9).

16. Brauer F, Castillo-Chavez C, Feng Z. Mathematical models in Epidemiology. Springer; 2019.

17. Walker PG, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low-and middle-income countries. Science. 2020.

18. Sjödin H, Johansson AF, Brännström Å, Farooq Z, Kriit HK, Wilder-Smith A, et al. COVID-19 healthcare demand and mortality in Sweden in response to non-pharmaceutical mitigation and suppression scenarios. International journal of epidemiology. 2020.

19. Gudbjartsson DF, Norddahl GL, Melsted P, Gunnarsdottir K, Holm H, Eythorsson E, et al. Humoral immune response to SARS-CoV-2 in Iceland. New England Journal of Medicine. 2020;383(18):1724–1734.

20. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science. 2021.

21. Bred immunitet efter 9 månader. Danderyds hospital press-release. Webbadress: www.ds.se/jobba-hos-oss/mot-oss/bred-immunitet-efter-nio-manader/

22. Folkhälsmyndigheten. Påvisning av antikroppar efter genomgången COVID-19 i blodprov från öppenvården.

23. Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household transmission of SARS-CoV-2: a systematic review and meta-analysis. JAMA network open. 2020;3(12):e2031756–e2031756.

24. Ferguson N, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College London. 2020;10(77482):491–497.

25. Carlsson M, Hatem G, Söderberg-Nauclér C. Mathematical modeling suggests pre-existing immunity to SARS-CoV-2. medRxiv. 2021.

26. Doshi P. Covid-19: Do many people have pre-existing immunity? Bmj. 2020;370.

27. Jones TC, Biele G, Mühlemann B, Veith T, Schneider J, Beheim-Schwarzbach J, et al. Estimating infectiousness throughout SARS-CoV-2 infection course. Science. 2021.

28. Britton T, Trapman P, Ball F. The risk for a new COVID-19 wave–and how it depends on $R_0$, the current immunity level and current restrictions. arXiv preprint arXiv:201004574. 2020.

29. Fine P, Eames K, Heymann DL. "Herd immunity": a rough guide. Clinical infectious diseases. 2011;52(7):911–916.

30. Yang Q, Saldi TK, Gonzales PK, Lasda E, Decker CJ, Tat KL, et al. Just 2% of SARS-CoV-2-positive individuals carry 90% of the virus circulating in communities. Proceedings of the National Academy of Sciences. 2021;118(21).

31. Brauer F. The final size of a serious epidemic. Bulletin of mathematical biology. 2019;81(3):869–877.

32. Artstein Z. Continuous dependence on parameters: On the best possible results. Journal of Differential Equations. 1975;19(2):214–225.

# Supplementary Material

## 1 Modeling of variable susceptibility and infectivity

We first build a compartmental model taking variable infectivity and susceptibility into account, and then prove that variable infectivity is irrelevant, under the assumption that the two are not correlated (otherwise the statement is false, see e.g. [6]). This is not surprising, see e.g. [8], but it is still worthwhile to prove within our modeling framework. Section 2 is the key contribution where we prove that the resulting susceptibility-stratified model can be further reduced to a basic SIR with ASI.

When modeling variable infectivity and super-spreaders, it is important to distinguish between super-spreaders in the sense that they have an extremely high viral load (possibly combined with lack of symptoms [27, 30]), and *social* super-spreaders, which are persons whose work or behavior puts them in contact with an unusually high number of people. The latter type was discussed in Section 3.2, so we focus here on the former type. We refer to [4] and [31] for the basics of mathematical modeling of infectious diseases.

Assume that the susceptibles can be divided into $J$ subgroups, and that the infectives can be divided into $K$ subgroups. Let $\mathbf{S} = (S_1, \ldots, S_J)^T$ be a vector such that $S = \sum_{j=1}^{J} S_j$ is the total amount of susceptibles, where the symbol $T$ denotes transpose, i.e. the operation that turns a row-vector into a column-vector. (All vectors considered here are to be column-vectors for the matrix vector formalism to function.) Similarly $\mathbf{I} = (I_1, \ldots, I_K)^T$ will represent the various compartments of $I$ where $I = \sum_{k=1}^{K} I_k$. All the above quantities implicitly depend also on time $t$ which is a continuous variable. Hence $S(t)$ and $I(t)$ denotes the total amount of susceptibles and infectives at a given time $t$, whereas $\mathbf{I}(t)$ and $\mathbf{S}(t)$ contain the amount in each respective compartment. Note that there is no need to introduce various compartments for $R$, since once you recover neither infectivity level nor susceptibility level is important. Let $a$ be the daily amount of potentially infectious contacts by an average infective during his period of illness and let $p_{j,k}$ be the probability that an infectious encounter between a member of $I_k$ and one in $S_j$ leads to transmission, and let $M_P$ denote the corresponding matrix. The rate of new infections in group $S_j$ is then easily seen to be $\nu_j(t) = \frac{S_j(t)}{N} \sum_{k=1}^{K} p_{j,k} a I_k(t)$, or in vector form

$$\boldsymbol{\nu} = \frac{a}{N} \mathrm{diag}_{\mathbf{S}} M_P \mathbf{I} \tag{10}$$

where $\text{diag}_{\mathbf{S}}$ denotes the diagonal matrix with $\mathbf{S}(t)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_J)$. Since we assume that susceptibility and infectivity are uncorrelated, the total amount of new infections are distributed over the various $I_k$'s in proportion to the fraction of the population that belong to respective group. More precisely, let $x_k$ denote the fraction of the total population that would end up in infectivity group $I_k$, if falling ill, and let $\mathbf{x} = (x_1, \ldots, x_K)^T$ denote the corresponding vector. The corresponding SIR-equation system becomes

$$\begin{cases} \mathbf{S}' &= -\boldsymbol{\nu} \\ \mathbf{I}' &= \nu\mathbf{x} - \sigma\mathbf{I} \\ R' &= \sigma I(t) \end{cases} \tag{11}$$

where $\nu = \sum_{j=1}^{J} \nu_j$ (so it is not a misprint that one is bold and the other not). The initial conditions become

$$\begin{cases} \mathbf{S}(0) &= \mathbf{w}N \\ \mathbf{I}(0) &= \tilde{\mathbf{x}}n \\ R(0) &= 0 \end{cases} \tag{12}$$

where $\mathbf{w} = (w_1, \ldots, w_J)^T$ is the fraction of the population in the respective susceptibility group at the onset of the pandemic, $n$ is the total number of import cases to the population of $N$ individuals, and $\tilde{\mathbf{x}}$ a vector containing the fraction of these in each subgroup of $I$.

These are the equations we propose for disease spread in a model with variable (uncorrelated) susceptibility and infectivity. A weak point of such a model is clearly the difficulty of measuring reasonable values of the transmission probabilities $p_{j,k}$ in practice. However, this is not needed since the main point of this paper is that the output of any such model is almost identical to the output of the most basic possible SIR if we include ASI. Given a fixed value of $\sigma$ (i.e. the inverse of the generation time $T_{generation}$), the basic SIR then depends only on two parameters, ASI $\theta$ and transmission rate $\alpha$. Hence, as long as we believe that there exist numbers such that the above system (10)-(12) would be a good model for accurately forecasting disease spread, we can instead run a basic SIR-model with only two unknown parameters, which are (more) easily fitted to real data. In conclusion, we do not need to know the actual transmission probabilities $p_{j,k}$, or the contact rate $a$. We now begin to reduce these equation to simpler ones.

## 1.1 The initial conditions are irrelevant

Even if $\tilde{\mathbf{x}}$ is different from $\mathbf{x}$, it will not remain so for long, for the equation for $\mathbf{I}'$ in (11) ensures that at any time point the new cases are distributed according to $\mathbf{x}$. Hence it will only take a few

generations before the distribution of infected very closely resembles that of $\mathbf{x}$. This will happen in the very early stages of the disease progression, and therefore we might as well assume that $\tilde{\mathbf{x}} = \mathbf{x}$ from the beginning.

More rigorously, we can argue as follows; Given the function $\nu$ describing the total incidence, the solution for $I_j$ is given by

$$I_j(t) = x_j \int_0^t e^{\sigma(s-t)} \nu(s)\, ds + \tilde{x}_j e^{-\sigma t},$$

by which it follows that $\tilde{x}_j$ soon becomes irrelevant, due to the exponential decay of the corresponding term $\tilde{x}_j e^{-\sigma t}$.

## 1.2 Variable infectivity is irrelevant; S-SIR

We now reduce the above equations to the Susceptibility-stratified model (S-SIR) used in the main text. Recall that $\mathbf{I}$ denotes the vector of all infectivity compartments and $I$ denotes their sum.

**Proposition 1.1.** *Given any solution* $(\mathbf{S}, \mathbf{I}, R)$ *to the system* (10)-(12) *with* $\tilde{\mathbf{x}} = \mathbf{x}$*, the triple* $(\mathbf{S}, I, R)$ *solves the system*

$$\boldsymbol{\nu} = \frac{a}{N} \mathbf{p} \cdot \mathbf{S} I \tag{13}$$

$$\begin{cases} \mathbf{S}' &= -\boldsymbol{\nu} \\ I' &= \nu - \sigma I \\ R' &= \sigma I(t) \end{cases} \tag{14}$$

$$\begin{cases} \mathbf{S}(0) &= \mathbf{w} N \\ I(0) &= n \\ R(0) &= 0 \end{cases} \tag{15}$$

*where* $\cdot$ *denotes componentwise multiplication,* $\nu = \sum_{j=1}^{J} \nu_j$ *and* $\mathbf{p} = M_P \mathbf{x}$*.*

*Proof.* Let $(\mathbf{S}, \mathbf{I}, R)$ be given and let $\boldsymbol{\nu}$ be defined by (10). Given a scalar solution to the equation $i' = \nu - \sigma i$, $i(0) = n$, it follows that $I_j(t) = x_j i(t)$ is a solution to the corresponding row in the equation $\mathbf{I}' = \nu \mathbf{x} - \sigma \mathbf{I}$. By the uniqueness of such solutions, we conclude that $\mathbf{I}(t) = \mathbf{x} i(t)$. Hence, since $I$ denotes the sum of $\mathbf{I}$, we see that $i = I$. With $\mathbf{p} = M_P \mathbf{x}$ it is easy to see that (10) reduces to (13), and that the equation for $\mathbf{I}$ in (11) reduces to the middle equation of (14).

□

Note that the numbers $p_1, \ldots, p_J$ in $\mathbf{p}$ are between 0 and 1 and can be interpreted as the probability of disease transmission when an "average" infective meets a susceptible in group $S_1, \ldots, S_J$, respectively. We can also compute $R_0$ given these transmission probabilities via the formula

$$R_0 = a\mathbf{w}^T M_P \mathbf{x} T_{generation} = a\langle \mathbf{w}, \mathbf{p} \rangle T_{generation}, \tag{16}$$

which should be compared with $R_0 = \alpha T_{generation}$ for the basic SIR (see Section 2.1). The simplest way to derive this is to note that at $t = 0$ we have $\mathbf{S}/N = \mathbf{w}$ so the equation for $I$ then reads $I'(0) = a\langle \mathbf{w}, \mathbf{p} \rangle I - \sigma I$. This means that an average infective infects $a\langle \mathbf{w}, \mathbf{p} \rangle$ new individuals every day, which in average goes on for $T_{generation}$ amount of days. Multiplying these two quantities gives (16). The same formula can also be deduced more rigorously using the formal mathematical definition of $R_0$ as the spectral radius of the next-generation matrix, see e.g. Sec. 5 in [16].

In the S-SIR model (13)-(15), herd-immunity arise when the derivative $I'(t)$ becomes negative, i.e. at the point $t_0$ when the equation $a\langle \frac{\mathbf{S}(t_0)}{N}, \mathbf{p} \rangle - \sigma = 0$. Since it is not possible to theoretically compute the distribution of susceptibles at this point, we can not give a closed formula for the herd-immunity threshold. This underlines the value of the result in the coming section, implying that the value is approximately

$$H_{IT} \approx \frac{\left(\sum_{j=1}^{J} w_j p_j\right)^2}{\sum_{j=1}^{J} w_j p_j^2}(1 - 1/R_0)$$

(combine (2) and (8)) The fact that the value is lower than the classical formula (1) is clearly related to the fact that highly susceptible individuals get infected to a greater extent than the other compartments. However, during a vaccination campaign there will be an equal fraction of immunity in in each respective group, i.e. a constant times $\omega$. Letting $H_{IT,vac}$ denote the vaccine induced herd immunity threshold, we thus get the equation

$$a\langle (1 - H_{IT,vac})\mathbf{w}, \mathbf{p} \rangle - \sigma = 0$$

which in combination with (16) implies the classical formula $H_{IT,vac} = 1 - 1/R_0$.

## 1.3 Examples with variable susceptibility

To illustrate, we note that if $J = 3$, then (13) becomes

$$\begin{bmatrix} \nu_1(t) \\ \nu_2(t) \\ \nu_3(t) \end{bmatrix} = \begin{bmatrix} p_1 S_1(t) \\ p_2 S_2(t) \\ p_3 S_3(t) \end{bmatrix} \frac{aI(t)}{N} \tag{17}$$

and now the total amount of newly infected becomes

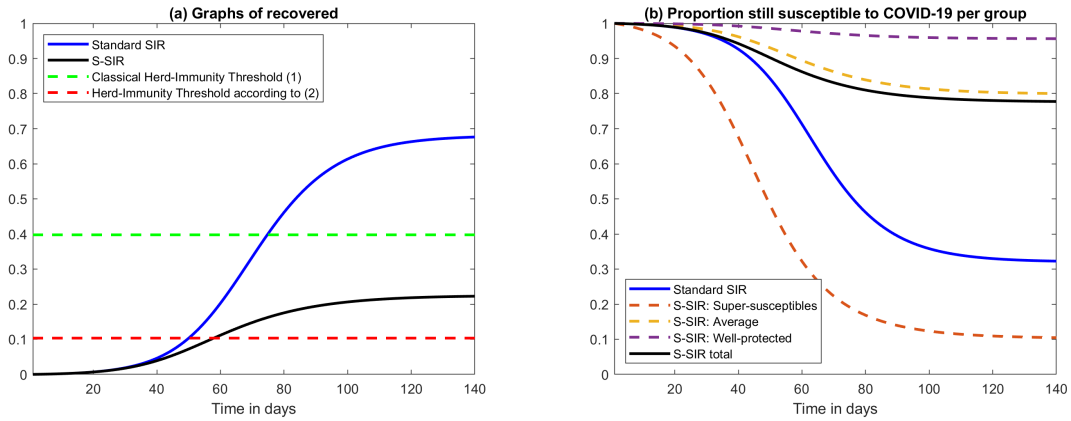
$$\nu(t) = \nu_1(t) + \nu_2(t) + \nu_3(t) = \frac{p_1 S_1(t) + p_2 S_2(t) + p_3 S_3(t)}{N} aI(t). \tag{18}$$

To see why this can not be further reduced to a basic SIR, in a similar way as the proof of Proposition 1.1, note that $\nu_1$ gets withdrawn from $S_1$, $\nu_2$ gets withdrawn from $S_2$ and $\nu_3$ gets withdrawn from $S_3$. If individuals in $S_1$ are much more susceptible than individuals in $S_2$, who are much more susceptible than those in $S_3$ (so $p_1 \gg p_2 \gg p_3$) then the fraction $\nu_1(t)/S_1(t)$ is much larger than $\nu_2(t)/S_2(t)$ and $\nu_3(t)/S_3(t)$, which means that the group of "super-susceptibles" $S_1$ will become depleted faster. But if $p_1 \gg p_2 \gg p_3$ then the super-susceptibles contribute to the main part of the coefficient in front of $I(t)$ in (18), and hence it turns out that the disease starts to recede when $S_1$ becomes sufficiently small, way ahead of what is expected from the standard (homogenous) SIR-model. In other words, while a standard SIR model needs to get above the classical herd-immunity threshold $1 - 1/R_0$ before it dies out, the model here can die out as soon as the super-susceptibles are depleted. Thus, once the super-susceptibles are depleted, the disease can die out by itself at fractions much lower than predicted by the classical HIT.

This phenomenon is clearly visible in Figure 2 in the main text. For that experiment, we assumed that $S_1$ consists of the super-susceptibles, for which $p_1 = 1$, and further that these constitute 30%, i.e. $w_1 = 0.3$. Similarly we let $S_2$ be 60% average group where only one in 10 contacts leads to disease transmission, i.e. $p_2 = 0.1$, and finally we supposed that the remaining 10% are well-protected with $p_3 = 0.02$. Clearly, the group of super-suseptibles have a huge attack rate of around 80% (compared with 72% for the whole population in the standard SIR using the same $R_0 = 1.66$), but as a consequence, the other two groups have attack rates of around 15% and 5%, respectively. This illustrates why traditional estimates for Herd-Immunity Threshold may be overly pessimistic.

To underline that the behavior of S-SIR seen in Figure 1 and 2 is representative, we provide a second example with only 10% super-spreaders and 30% well protected (i.e. we flip the amount in

group $S_1$ with the amount in group $S_3$). The result is displayed in Supp. Fig. 1. As is plain to see, not much changes; the tiny group of super-susceptibles still "protects" the rest, in the sense that once they are depleted the disease dies out naturally. This group now has a 92% attack rate, as opposed to previously 80% in Fig. 2. Looking at the graph of recovered, we also see that it looks almost as either of the curves in Fig. 1a, only the final size of the pandemic changes to 23%, from 33% for the previous S-SIR and 72% for standard SIR, respectively. Intuitively, it should be possible to find a setting of parameters where the disease burns out among the super-susceptibles, but lives on a bit longer in the "average" group. However, the main finding of this paper is that this intuition is wrong, as we establish in the next section.




Supp. Fig 1: **Comparison graphs of $R$ and $S$ for alternative compartment sizes.** (a) Graphs of recovered as fraction of the population for SIR and S-SIR with a fixed value of $R_0 = 1.66$, along with estimates for the Herd-Immunity Threshold. As expected, S-SIR reaches a final size of the pandemic way below the classical HIT, but about twice that of the HIT as computed according to formula (2) suggested in this paper. (b) Corresponding graphs for susceptibles, which should be compared with Figure 2 in the main text. The model assumes 10% super-spreaders and 30% well protected, as opposed to the situation with 30% super-spreaders and 10% well protected depicted in Figure 2.

## 2 Variable susceptibility is almost equivalent to ASI

We now motivate the claims made in Section 2, in particular formula (8), in the more general setting of the Susceptibility-stratified SIR (13)-(15). We first introduce some notation and explain the intuition behind the formula (8). First consider standard SIR (3)-(5), and let $(S_0, I_0, R_0)$ be a particular solution. We introduce the normalized functions $s_0(t) = S_0(t)/N$, $i_0(t) = I_0(t)/N$,

$r_0(t) = R_0(t)/N$. These functions solve the system

$$\begin{cases} s' &= -\alpha s i \\ i' &= \alpha s i - \sigma i \\ r' &= \sigma i \end{cases} \tag{19}$$

with $s(0) = \omega$, $i(0) = n/N =: \varepsilon$ and $r(0) = 0$. Note that $s_0(t) + i_0(t) + r_0(t) = \omega + \varepsilon$. Next, we reduce (19) by using the well-known fact that $t$ can be replaced by $r$ as the independent variable. Let $\tilde{s}_0$ and $\tilde{i}_0$ be functions defined via $s_0(t) = \tilde{s}_0(r_0(t))$ and $i_0(t) = \tilde{i}_0(r_0(t))$. By the chain rule the functions $(\tilde{s}_0, \tilde{i}_0)$ are solutions to the system

$$\begin{cases} ds/dr &= -\alpha s/\sigma \\ di/dr &= \alpha s/\sigma - 1 \end{cases} \tag{20}$$

which is easily solved. Since $r_0(0) = 0$ we have that $\tilde{s}_0(0) = s_0(0) = S_0(0)/N = \omega$, by which we infer that

$$\tilde{s}_0(r) = \tilde{s}_0(0)e^{-\alpha r/\sigma} = \omega e^{-\alpha r/\sigma} = \frac{\sigma}{\alpha} f_0(r) \tag{21}$$

where $f_0(r) := \frac{\alpha\omega}{\sigma} e^{-\alpha r/\sigma} = R_0 e^{-R_0 r/\omega}$. Returning to (20) we see that

$$\frac{d\tilde{i}_0}{dr} = \frac{\alpha \tilde{s}_0}{\sigma} - 1 = f_0 - 1.$$

Integrating both sides and using the initial condition $\tilde{i}_0(0) = i_0(0) = \varepsilon$ gives the solution

$$\tilde{i}_0(r) = \varepsilon + F_0(r) - r, \tag{22}$$

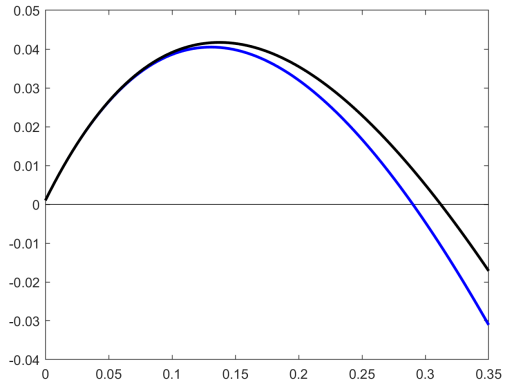
where $F_0(r) = \int f_0 = \omega(1 - e^{-R_0 r/\omega})$. See Supp. Fig. 2 for an illustration.

We now solve the equation for $r_0(t)$ in (19), which is separable and can be written

$$\frac{dr_0/dt}{\sigma \tilde{i}_0(r_0(t))} = 1.$$

Integrating both sides with respect to $t$ and making the change of variables $x = r_0(t)$ gives

$$\int_0^{r_0(t)} \frac{dx}{\sigma \tilde{i}_0(x)} = t.$$




Supp. Fig 2: **Graphs of $\tilde{\imath}_0$ and $\tilde{\imath}_1$.** The figure shows the solution $\tilde{\imath}_1$ (black) for the same example as in Fig 2 and $\tilde{\imath}_0$ (blue) using the choice of $\alpha$ and $\omega$ stipulated by (8).

Letting $t_0$ be the primitive function of $1/(\sigma\tilde{\imath}_0)$ satisfying $t_0(0) = 0$ we thus obtain the implicit solution $t_0(r_0(t)) = t$. Hence, $t_0$ is also the inverse of $r_0$ (which exists since $r_0'(t) > 0$ for all $t > 0$), and so

$$r_0(t) = t_0^{-1}(t), \quad t_0(r) := \int_0^r \frac{dx}{\sigma\tilde{\imath}_0(x)}. \tag{23}$$

By differentiating (22) we see that $\tilde{\imath}_0$ is concave. Let $r_\infty$ be the positive solution of the equation $\tilde{\imath}_0(r) = 0$, which thus is unique. By standard calculus this gives rise to a non-integrable singularity in (23), and hence $\lim_{r\to r_\infty} t_0(r) = \infty$. Since $r_0$ is the inverse of $t_0$, we see that $r_\infty$ equals the limit $\lim_{t\to\infty} r_0(t)$, known as the final size of the pandemic. We therefore refer to this number simply as $r_0(\infty)$ in what follows. Summing up, $t_0$ is only defined for values of $r$ in $[0, r_0(\infty))$. Note also that as a consequence we also obtain

$$\lim_{t\to\infty} i_0(t) = 0 \tag{24}$$

since $i_0(t) = \tilde{\imath}_0(r_0(t))$ so the above limit becomes $\tilde{\imath}_0(r_0(\infty)) = 0$.

Applying the same arguments to (13)–(15) we obtain first a system in terms of normalized variables

$$\begin{cases} s_j' &= -ap_j s_j i \quad 1 \le j \le J \\ i' &= \sum_{j=1}^J ap_j s_j i - \sigma i \\ r' &= \sigma i \end{cases} \tag{25}$$

and then an equivalent reduced system of the form

$$\begin{cases} ds_j/dr &= -ap_j s_j/\sigma \quad 1 \le j \le J \\ di_1/dr &= \sum_{j=1}^J ap_j s_j/\sigma - 1. \end{cases} \tag{26}$$

Let $(\tilde{s}_1, \ldots, \tilde{s}_J, \tilde{\imath}_1)$ be a solution to (26) with $\tilde{\imath}_1(0) = \varepsilon$ and $\tilde{s}_j(0) = S_j(0)/N = w_j$ for $j = 1, \ldots, J$. Similarly, we will write $r_1(t)$ to denote the function described by (23) with $\tilde{\imath}_0$ swapped for $\tilde{\imath}_1$. As in (21)–(22) we obtain the solutions

$$\tilde{s}_j(r) = w_j e^{-ap_j r/\sigma} \quad 1 \le j \le J \tag{27}$$

$$\tilde{\imath}_1(r) = \varepsilon + \sum_{j=1}^J w_j(1 - e^{-ap_j r/\sigma}) - r \tag{28}$$

See Supp. Fig. 2 (black curve) for an illustration using the same values as in SM Section 1.3, in particular note that it meets the $x$-axis at the final size of the pandemic $r_1(\infty)$, which is seen slightly above 0.3 for the black graph in Figure 1.

The philosophy behind Section 2 and (8), is that we want to pick the parameters $\alpha$ and $\omega$ so that (22) becomes approximately equal to (28). Then, due to (23) we will have $r_1(t)$ approximately equal to $r_0(t)$ and therefore $i_1(r_1(t))$ approximately equal to $i_0(r_0(t))$. Now, by Taylor's formula we have

$$\tilde{\imath}_0(r) = \varepsilon - r + \omega\left(\frac{\alpha}{\sigma}r - \frac{1}{2}\frac{\alpha^2}{\sigma^2}r^2 + O(r^3)\right)$$

and

$$\tilde{\imath}_1(r) = \varepsilon - r + \sum_{j=1}^J w_j\left(\frac{ap_j}{\sigma}r - \frac{1}{2}\frac{(ap_j)^2}{\sigma^2}r^2 + O(r^3)\right)$$

where $O$ is the big O notation of Bachman-Landau. Comparing Taylor coefficients we set $\omega\alpha = \sum_{j=1}^J w_j a p_j$ and $\omega\alpha^2 = \sum_{j=1}^J w_j (ap_j)^2$ so that the Taylor polynomials of order two of $\tilde{\imath}_0$ and $\tilde{\imath}_1$ coincide. Solving for $\alpha$ and $\omega$ gives (8), i.e.

$$\alpha = a\frac{\sum_{j=1}^J w_j p_j^2}{\sum_{j=1}^J w_j p_j}, \quad \omega = \frac{\left(\sum_{j=1}^J w_j p_j\right)^2}{\sum_{j=1}^J w_j p_j^2}.$$

Note here that we always have $\omega < 1$. In fact, since $\sum_{j=1}^J w_j = 1$ and $0 < w_j, p_j \le 1$ for all $j$ it

follows that $\sum_j w_j p_j \leq 1$, which implies that

$$\sum_{j=1}^{J} w_j p_j^2 > \sum_{j=1}^{J} w_j p_j \geq \left( \sum_{j=1}^{J} w_j p_j \right)^2 \quad \Longleftrightarrow \quad 1 > \frac{(\sum_{j=1}^{J} w_j p_j)^2}{\sum_{j=1}^{J} w_j p_j^2} = \omega \tag{29}$$

(with strict inequality since at least one $p_j$ is strictly less than one). These formulas can be put in somewhat neater form by introducing $R_1 = a/\sigma$, i.e. the $R_0$-value one would have at the beginning of the pandemic if everyone were a super-susceptible. We then get

$$R_0 = \frac{\omega\alpha}{\sigma} = \frac{a \sum_{j=1}^{J} w_j p_j}{\sigma} = R_1 \sum_{j=1}^{J} w_j p_j.$$

Introducing $f_1(r) = R_1 \sum_{j=1}^{J} w_j p_j e^{-R_1 p_j r}$ and $F_1 = \int f_1 = \sum_{j=1}^{J} w_j (1 - e^{-R_1 p_j r})$, we can write

$$\tilde{\imath}_1(r) = \varepsilon + F_1(r) - r, \tag{30}$$

which should be compared with (22). Note that $F_1$ is engineered to have same first three Taylor coefficients as $F_0$. Moreover, introducing

$$t_1(r) := \int_0^r \frac{dx}{\sigma \tilde{\imath}_1(x)}, \tag{31}$$

we obtain $r_1$ as $t_1^{-1}$ and subsequently $r_1' = \sigma i_1$.

To summarize this section, we argue that if $F_0 \approx F_1$ and $r_0 \approx r_1$, then it should follow that $i_0 \approx i_1$ since $i_j(t) = \tilde{\imath}_j(r_j(t))$, $j = 0, 1$, and $\tilde{\imath}_0$ and $\tilde{\imath}_1$ are given by (22) and (30), respectively. To formalize this statement a bit, consider now the more general ODE

$$r'(t) = G(r(t), \delta), \quad r(0) = 0, \tag{32}$$

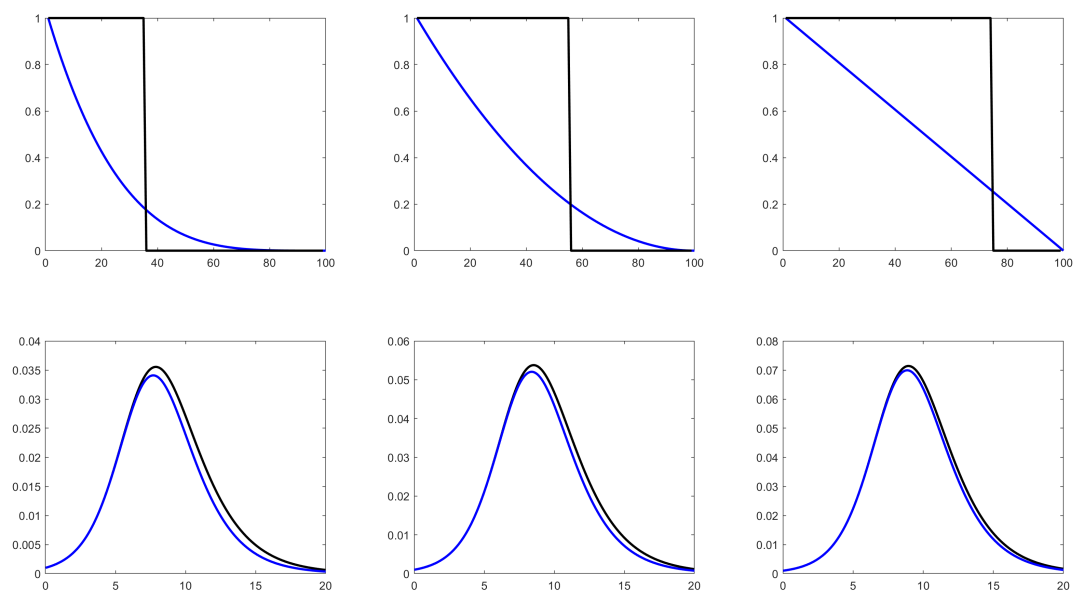
where

$$G(r, \delta) = \sigma(\varepsilon + (1 - \delta) F_0(r) + \delta F_1(r) - r), \quad \delta \in \mathbb{R}.$$

Then $r_0(t)$ is the unique solution to (32) for $\delta = 0$, while $r_1(t)$ is the unique solution to (32) for $\delta = 1$. Since $G(x, \delta)$ is independent of $t$ and continuous in $\delta$ for every fixed $x$, it follows that the solution $r(t) = r(t, \delta)$ of (32) is continuous in $\delta$, uniformly in $t \in [0, T]$ for any $T > 0$, see for example [32] and the references therein. We collect these observations in the following theorem.

**Theorem 2.1.** *Let $r(t,\delta)$ be the unique solution of* (32)*. Then for any $T > 0$, the solution $r(t,\delta)$ is a continuous function of $\delta$, uniformly in $t \in [0,T]$. Moreover, we have $r(t,0) = r_0(t)$ and $r(t,1) = r_1(t)$.*

Our numerical simulations have confirmed that indeed, the difference between $i_0$ and $i_1$ is small. To illustrate, we divided the population into 100 (equal size) subgroups (as opposed to 3), and assigned probabilities according to the formula $p_j = \left(1 - \frac{j-1}{100}\right)^\rho$ where $\rho$ is a parameter to be chosen. This means that we assume a monomial distribution of individual transmission probabilities, and see which exponent $\rho$ that seems to give a good fit. In Supp. Fig. 3 we show the curves for $\rho = 4,\ 2$ and 1 ($p_j$-curve in embedded picture) with respective values of ASI (obtained using (8)) equal to 64%, 45% and 25%. Note that basic SIR with ASI becomes a special case of S-SIR if we choose a function $p$ that is either 0 or 1 with a sharp drop at $100\omega$. These curves are also seen in the embedded pictures.



Supp. Fig 3: ***p*-curves.** Top: Various $p$-curves (blue) and corresponding binary immunity curves (black) with ASI obtained from (8). Bottom: corresponding waves of prevalence for S-SIR and SIR with ASI, using parameters obtained from (8).

## 2.1 Further results and a conjecture

It remains to establish proper estimates proving that the resulting solution $(i_0, r_0)$ is near $(i_1, r_1)$. We first establish the following important inequality, which is the explanation behind that the black

curve (from S-SIR) always seems to be a bit higher than the blue one (from SIR).

**Proposition 2.2.** *It holds that $f_1 \geq f_0$. Consequently $F_1 \geq F_0$, $\tilde{\imath}_1 \geq \tilde{\imath}_0$ and $d\tilde{\imath}_1/dr \geq d\tilde{\imath}_0/dr$.*

*Proof.* The choice of $\omega$ and $\alpha$ is designed so that $f_1(0) = f_0(0)$ and $f_1'(0) = f_0'(0)$. Moreover we have

$$f_1(r) - f_0(r) = \frac{1}{\sigma}\left(\sum_{j=1}^{J} w_j a p_j e^{-ap_j r/\sigma} - \omega\alpha e^{-\alpha r/\sigma}\right).$$

Multiplying with $e^{\alpha r/\sigma}$ we obtain $\sum_{j=1}^{J} w_j a p_j e^{(\alpha - ap_j)r/\sigma} - \omega\alpha$ which is a strictly convex function (unless $\alpha = ap_j$ for all $j$, which means that all probabilities are the same and the model collapses to the standard SIR model, and there is nothing to prove) that in addition equals zero and has derivative zero at $r = 0$ by construction. By convexity it follows that the function is non-negative, giving $d\tilde{\imath}_1/dr \geq d\tilde{\imath}_0/dr$ with strict inequality for $r > 0$. Convexity also implies that the function and its derivative are increasing. The desired inequality follows by integrating $d\tilde{\imath}_1/dr \geq d\tilde{\imath}_0/dr$, keeping in mind that $\tilde{\imath}_1(0) = \tilde{\imath}_0(0)$. □

**Proposition 2.3.** *Let $r_0$ be the solution to* (19) *and $r_1$ the solution to* (25)*. Then $r_1 \geq r_0$.*

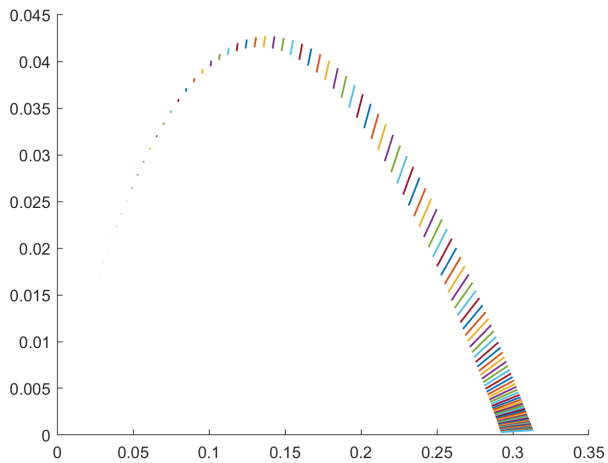
*Proof.* Let $g(x) = \sigma\tilde{\imath}_0(x)$. Then $r_0'(t) = g(r_0(t))$ while $r_1'(t) \geq g(r_1(t))$ by Proposition 2.2. Since $r_1(0) = r_0(0)$, standard theory of ordinary differential equations then dictates that $r_1(t) \geq r_0(t)$ for $t \geq 0$. □

However, we believe there is more to say than what is captured in the above two propositions. Supp. Fig. 4 gives an illustration of how the pairs $(i_0(t), r_0(t))$ and $(i_1(t), r_1(t))$ evolve with time. Note how the distance seems to be gradually increasing and always points in the north-east direction.

We have not been able to prove the following observation, which in a more concrete way would imply that $i_1$ and $i_0$ are close.

**Conjecture 2.4.** *We have $i_1 \geq i_0$ and furthermore, $s_0 - \sum_{j=1}^{J} s_j$ is an increasing function.*

To understand the implications of the conjecture, recall that the final size of the pandemic solves $\tilde{\imath}_j(r_j(\infty)) = 0$, $j = 1, 2$, and that these two points tend to be quite close, see for example Supp. Fig. 2 where the difference $r_1(\infty) - r_0(\infty)$ is around 0.02 or 2%. Since $s_0 - \sum_{j=1}^{J} s_j = \omega - 1 + (i_1 + r_1) - (i_0 + r_0)$ and $r_1(\infty) - r_0(\infty)$ is the limit of $(i_1 + r_1) - (i_0 + r_0)$ as $t \to \infty$, the conjecture implies that both the difference $r_1(t) - r_0(t)$ as well as $i_1(t) - i_0(t)$ are bounded by $r_1(\infty) - r_0(\infty)$.




Supp. Fig 4: **Envelope foliation.** Line between pairs $(i_0(t), r_0(t))$ and $(i_1(t), r_1(t))$ for a grid of $t$-values, using the same parameters as in Fig. 2 for S-SIR and parameters from (8) for SIR. Note that the inner and outer envelope is given by $\tilde{\imath}_0$ and $\tilde{\imath}_1$ (cf. Fig. 2), and that the slopes are decreasing but positive.