Stochastic epidemic spreading: not all super-spreading processes are born equal, neither all lockdown strategies

Jhonatan Tavori and Hanoch Levy

Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel

ABSTRACT
We consider viral spreading processes, such as pandemics, in finite networks. For such processes, we propose and analyze a new model which combines two stochastic functions in the spreading intensity of a node, and accounts for two types of super-spreading. The first reflects personal properties of each node and the second reflects occasional spreads in the network. Consequently, studying the spreading dynamics requires the analysis of a stochastic process consisting of those two functions which drastically differ in their dynamics. One (personal) is biasly-modified throughout the process as infected nodes “leave the game” and the distribution of the susceptible population changes. The second (occasional) remains constant throughout the process. We show that the mix between these functions affects dramatically the threshold for the end of the spread (known as the Herd Immunity Threshold, or HIT). We address operational aspects and examine the effectiveness of control mechanisms that restrict the interaction among the population in order to suppress the spread (e.g. lockdowns). We reveal and establish that although different policies might have similar immediate impacts, not all lockdowns are “born equal”, and may drastically differ on the long-term impact: While some reduce the HIT, others may be counter-productive in the long-run and, perhaps surprisingly, increase the HIT.

1. Introduction

Epidemic processes refer to the spread of information through a network, and have been widely used as an abstraction for various real-world phenomena, such as human infections and the outbreak of epidemics,1[6,10] the spreading of computer viruses over the Internet,8,9 the spreading of rumors over online social networks,5,15 information broadcasts,7,19 and so on.
The effective reproduction number $R()$, which is the expected number of secondary cases generated by an infected individual, decreases as individuals become infected and then recover (i.e. develop immunity), or even removed, and the size of the susceptible population decreases. Once the Herd Immunity Threshold (HIT, measured in fractions of the population that contracted the spread) is surpassed, then $R()$ reduces below 1 and the number of new infection cases decreases. Such threshold phenomena have been observed empirically.

Numerous real-world networks consist of highly heterogeneous individuals, which may be infectious and susceptible to varying degrees. Recent works showed that the heterogeneity across individuals (recognized as “over-dispersion” or “super-spreading”), may have a drastic effect on the over-time reduction of the effective reproduction number, $R()$, and on the HIT. The ongoing COVID-19 pandemic has made it possible to observe the existence of such heterogeneity phenomenon conspicuously. The existence of “super-spreaders” yields that they are extremely likely to get infected and develop immunity in an early stage of the pandemic process. As a result, those works predicted that “super-spreading” disappears at early stages of the process and the traditionally predicted value of the HIT in heterogeneous networks should decrease dramatically in comparison to homogeneous networks.

Nonetheless, COVID-19 and its spread also taught us that the viral-spreading process is characterized by one more important type of “super-spreading”, a type that relates to occasional interactions between nodes in the network. For example, social gatherings which involve a large number of participants such as concerts, sports games or cultural gathering. This type cannot be tied to specific individuals, and thus does not disappear in the early stages of the process.

Hence, models that rely on the existence of super-spreaders (which we denote also as personal super-spreading) but disregard occasional super-spreading may be short of capturing the full behavior of the spreading process. In fact, we will show that the occasional super-spreading can have high impact on the process. In particular, the effective reproduction number $R()$ may remain relatively high, even after the “super-spreaders” are immune. Furthermore, as we will show, accounting in the model for both types of spreading allows one to evaluate the effect of control mechanisms (such as lockdowns) that focus on one of the two spreading types.

In this work, we propose and analyze a new stochastic model that accounts for mixed heterogeneity consisting of two-types of super-spreading. We propose that the spreading intensity of a node and its heterogeneity are composed of two components. The first component reflects
personal properties of each node and the second reflects occasional spreads which are common to all nodes.

More specifically, at each step $n$ of the spread, each individual $a$ is assigned with its susceptibility and infectiousness parameters (denoted $S(a,n)$ and $I(a,n)$, respectively). Each of them consists of the weighted sum of two random variables (which fundamentally differ in their evolution), as follows:

1. $S_p(a)$ and $I_p(a)$ which are the personal-trait (continual) susceptibility and infectiousness parameters of $a$, respectively. These values reflect personal traits and are drawn once for $a$ at the beginning of the pandemic and then remain constant throughout.
2. $S_e(a,n)$ and $I_e(a,n)$ which are the event-based (occasional) susceptibility and infectiousness parameters. These values reflect occasional event-based spreading, which are probabilistically redrawn for each individual at every step $n$ of the pandemic.

The likelihood of $a$ to be infected is proportional to:
$$S(a,n) = p \cdot S_p(a) + (1 - p) \cdot S_e(a,n),$$
and the likelihood of $a$ to infect others is proportional to:
$$I(a,n) = p \cdot I_p(a) + (1 - p) \cdot I_e(a,n).$$

$p$ represents the relative weight of each spreading type in the network. Networks characterized by a high (low) level of occasional spreading will have a low (high) value of $p$. Precise modeling and further details appear in Section 2.

Using this model, we track the evolution of the population distribution throughout the spreading process. Mathematically, studying the spreading dynamics requires the analysis of a stochastic process consisting of two interacting stochastic functions which drastically differ in their evolution: One function (representing personal spreading) is dynamic – since its values are pre-determined, their distribution among the susceptible population is biasly-modified as individuals contract the spread and then “being removed” or develop immunity. In contrast, the second function (representing occasional spreading) is static – since its values are re-drawn at every step from the same distribution. To address this challenge, our analysis utilizes some techniques used in the analysis of a single-type heterogeneity model,[14] while deploying, in addition, additional probabilistic treatments needed to cope with tracking the interaction of the two heterogeneity types. Further probabilistic methodologies (e.g. stochastic dominance and likelihood ratio) are deployed to cope with the effects of control mechanisms.
The first part of our analysis derives $R(n)$, the expected value of the reproduction number, as a function of the number of infected individuals. This in turn allows us to predict the value of the HIT. Both approximations are derived as a function of the two spreading functions as well as their relative weight (i.e. the mix between them). We show that their relative weight, $p$, affects dramatically the dynamics of the spread across the network. We find that the HIT value is very sensitive to $p$ and thus different networks/societies may engender significantly different HITs; in fact, we show that under the spreading distribution and reproduction number attributed to the COVID-19 pandemic, depending on $p$, the HIT can vary from as low as 5% to as high as 66.66%. This difference is huge and it can have huge effect on the economics of a pandemic, as HIT estimations are a key factor in directing strategic decisions concerning the fight against it.

Having established the model and derived the HIT, we then investigate control mechanisms, such as lockdowns, which are used to suppress the spreading process by reducing the level of interaction between the individuals. We study their effect on the progression of the process, on $R(n)$ and on the HIT. Since our model accounts for both types of spreading, we may classify such lockdowns into two inherently different types, according to their main focus: (1) Personal-trait spreading targeted lockdowns; for example, closing or restricting workplaces. (2) Event-based spreading targeted lockdowns; for example, prohibition on cultural events. We compare a system subject to a “natural” evolution of the spreading process over the population (i.e. without any preventive-measures being used) to systems subject to personal-trait lockdowns, and to event-based lockdowns.

We establish stochastic monotonicity and dominance relations between the systems, resulting from the lockdown. Using these relations we reveal and establish that the effect of lockdowns on the long-term behavior of the pandemics, namely on HIT, is very sensitive to their focus: Event-based spreading targeted lockdowns reduce the HIT; In contrast, and somewhat surprisingly, personal-trait spreading targeted lockdowns act inversely, increase the HIT and may be counter-productive in the long-run.

Note that these results are general and hold for any relative weight of the spreading types in a network, and any spreading distributions, including various heterogeneity levels.

Preliminary short version of partial results was presented in Tavori and Levy.[17]

In light of the current public interest in the COVID-19 pandemic, and due to its tremendous importance to our daily life, our presentation will develop the model within this context. Nonetheless, the model and machinery developed in this work and possibly the results as well can prove useful...
in the analysis of computer virus spreading over computer networks or rumor spreading over social networks.

The paper is organized as follows. Our stochastic spreading model is presented in Section 2. We then (Section 3) move to analyze the spread dynamics and its effect on the distribution of the population, and derive the Herd Immunity Threshold; the analysis is supported by Monte-Carlo simulations that provide excellent fit with the analytic HIT prediction. In Section 4 we address operational aspects and examine control mechanisms which slow down the spread by limiting the interaction between the network entities (i.e. lockdowns). We study the long-term effects of different lockdown policies on the spread of the disease, and on the HIT. In Section 5 we evaluate the model both numerically and via a simulation and discuss the results. Concluding remarks are given in Section 6.

2. The model

We present our model for the spreading process, accounting for mixed (two-type) heterogeneity of susceptibility and infectiousness (spreading). Mathematically, the types differ in the way their distributions evolve through the process, and the way they affect each other and the infection process. We adopt the Susceptible-Infected-Recovered (SIR) model, a standard model of epidemic processes.\cite{10,13,15}

Let \( A_0 = \{a_1, a_2, \ldots, a_{N_0}\} \) be the set of the network’s nodes (i.e. population), where \(|A_0| = N_0\). Our analysis begins with a certain number of infective nodes, where all others are assumed to be initially susceptible (S). As a result of an infection, susceptible nodes become infective (I). After an infectious period, infected nodes “leave the game”, that is, recover or removed (R).

We follow\cite{14} and index the spread across the network as a function of the number of nodes which contracted it. Namely, the event whereby the \( n \)th infection occurs is called the \( n \)th step of the disease. Hence, the number of steps is upper-bounded by \( N_0 \).

At step \( n \), each node \( a \) is assigned with susceptibility and infectiousness parameters, \( S(a, n) \) and \( I(a, n) \); The construction of which will be further described. The values of \( S \) and \( I \) are in the range \([0, 1]\). \( I(a, n) \) represents the probability of \( a_i \) to spread the disease to others when it is in Infective mode. \( S(a, n) \) represents the probability of \( a_j \) to attract the infection. Assuming that \( a_i \) is infective at step \( n \), the probability that it will infect the susceptible node \( a_j \) is: \( \Pr[a_i \text{ infects } a_j \text{ at step } n] = I(a_i, n) \cdot S(a_j, n) \). Note that \( S(a, n) \) and \( I(a, n) \) implicitly include the social interaction level of \( a \) (probability of meeting other individuals) as well as any personal or physical properties (such as biological features, social distancing obedience or...
The duration of the infectious period is absorbed in the spreading parameters.\textsuperscript{5}

Throughout our analysis we will assume that no individual infects (directly) a significant fraction of the population (while allowing heterogeneity of the spreading parameters). This is a common assumption in the epidemiics literature.\textsuperscript{[2,11,14]} See further discussion at Section 3 and Appendix B.

\section{Stochastic spreading functions}

We classify spreading into two inherently different types: (1) \textit{Personal-Trait-Spreading} (continual) and (2) \textit{Event-Based-Spreading} (occasional). The first (Personal-Trait) stems from the personal traits of an individual, while the second (Event-Based) relates to occasional spreads (e.g. social events) in which every individual may participate, regardless of their personal traits.

As an example of these two types, consider a supermarket cashier compared to an academic researcher. In any given day, the cashier interacts with tens or hundreds of people, and therefore has high personal-trait spreading degree. In contrast, the researcher stays in the laboratory or interacts with a small research group. However, both may, occasionally, participate in a social-gathering event such as a concert, a wedding, or “just” a family birthday party. During such an event, both may engage in a similar amount of social interaction (which may be quite large), and therefore they both have the same event-based spreading degree.\textsuperscript{6}

From mathematical point of view they also inherently differ from each other, as they follow inherently different stochastic progressions along an epidemic.

\subsection{Personal-trait spreading}

Let \(S_p\) and \(I_p\) denote the \textit{personal-trait susceptibility and infectiousness distributions} of the network. Where \(\text{Supp}(S_p) = \text{Supp}(I_p) = [0,1]\). At the beginning of the process each node \(a \in A_0\) is assigned with values \(S_p(a)\) and \(I_p(a)\), reflecting the personal traits of \(a\). These values accompany individual \(a\) throughout the entire process. The ensemble of the values over all \(a \in A_0\) forms the distribution \(S_p\) and \(I_p\).\textsuperscript{7}

\subsection{Event-based spreading}

Let \(S_e\) and \(I_e\) denote the \textit{event-based susceptibility and infectiousness distributions} of the network. Where \(\text{Supp}(S_e) = \text{Supp}(I_e) = [0,1]\). These values reflect occasional spreads across the network. By nature, these are random for each individual and their realizations are redrawn for each individual at
every step. Let $S_e(a,n)$ and $I_e(a,n)$ be random variables denoting the event-based susceptibility and infectiousness values assigned to $a$ at step $n$; $S_e(a,n)$ and $I_e(a,n)$ denote realizations of those random variables. Since these parameters are event based, these variables are redrawn for $a$ (from $S_e$ and $I_e$, which are common to the entire population) at every step $n$. This reflects the fact that at one step the realizations of $S_e(a,n)$ and $I_e(a,n)$ may be low since $a$ stayed home, while at another step they may be high since $a$ went to a music concert. Note that this redrawing is in contrast to $S_p(a)$ and $I_p(a)$ which remain constant throughout the process.

**Remark 2.1.** Since our indexing method (i.e. steps) is based on the number of infection cases, we follow it and, to carry the analysis, we let the event-based parameters to be redrawn at each step too. Note that this redrawing method does not capture the very precise behavior of an individual at an event, say a wedding, since in real life this individual redrawing is based on time periods, for example, a wedding lasts 2–3h. Nonetheless, our redraw approach allows capturing well, in one model, the population-wide behavior of event-based (including weddings) and personal-based spreading. In particular, it captures well the fact that the level of personal-based spreading “dies out” quickly along an epidemic process since the super-spreaders are likely to get infected early and go out of the game; and that, in contrast, event-based spreading remains highly active as long as “the weddings are on”. To this end, we will validate and demonstrate our analytic results using numerical simulations, in which we will redraw the event-based parameters based on time periods (e.g., days or hours). The comparison of these simulations results to the analysis, which finds excellent fit, is provided in Section 3.4 and in Figure 1.

**Notation 2.2** (Spreading parameters). The susceptibility parameter (realization) of $a$ at step $n$, is modeled as a linear combination of its personal-trait and event-based parameters, as follows:

$$S(a,n) = p \cdot S_p(a) + \bar{p} \cdot S_e(a,n),$$  \hspace{1cm} (1)

The infectiousness parameter of $a$ at step $n$, is:

$$I(a,n) = p \cdot I_p(a) + \bar{p} \cdot I_e(a,n).$$  \hspace{1cm} (2)

where $0 \leq p \leq 1$, $\bar{p} = 1 - p$, and $S_p, I_p, S_e, I_e$ are produced as described above.

The value of $p$ represents the relative weight of each spreading type in the network. Networks characterized by a low (high) level of social gatherings will have a high (low) value of $p$. In Sections 3 and 5 we will examine the effects of the value of $p$ on the spreading process and will demonstrate
that it has a significant effect on the reduction of the effective reproduction number. Thus, deviation in its value might affect dramatically the HIT.

2.2. Monotonicity of spreading

In many cases it is logical to assume that the susceptibility and infectiousness of each individual are equal, as susceptibility and infectiousness levels are both proportional to the level of interaction or other properties of $a$. Yet, the results developed in this work are based on a significantly lighter assumption – *monotonicity of spreading*: more susceptible individuals are stochastically more infectious. Formally, denote by:

**Notation 2.3** (Expected conditional personal-trait infectiousness).

$$
\varphi(s) := \mathbb{E}_{a \in A_n} \left[ I_p(a) \mid S_p(a) = s \right].
$$

We say that $S_p$ and $I_p$ hold *monotonicity of spreading* if their sampling is positively correlated such that $\varphi(s)$ is monotonically non-decreasing in $s$. Throughout this work, we assume that $S_p$ and $I_p$ hold the monotonicity of spreading property. Off course, the case of equality $I_p(a) = S_p(a)$ is a special case of monotonicity of spreading.

![Figure 1. HIT values as a function of $p$. We use population of size $N_0 = 50000$, and Gamma distribution with shape parameter $k = 0.1$. The simulations run multiple times both for $R_0 \approx 3$, and $R_0 \approx 9$ and for various values of $p$ from 0 to 1.](image_url)
2.3. Tracking the spreading distribution

While the event-based spreading parameters are redrawn at every step (from the non-changing distributions $S_e$ and $T_e$), the personal-trait spreading parameters are determined once and are not re-drawn. Hence, the distribution of the personal-trait susceptibility and infectiousness parameters across the susceptible population may change throughout the spread, as infected nodes are removed from the (susceptible) population.

Let $A_n$ denote the susceptible population at step $n$ (a subset of the population at step $n - 1$). $A_n$ is a random variable distributed over all possible scenarios of infection.

Notation 2.4 (Personal-Trait Susceptibility Density at Step n). We denote by $\rho(s, n)$ the Probability Density Function (pdf) of the personal-trait susceptibility of $A_n$, the susceptible population at stage $n$. Namely, if an individual $a$ is picked at random at stage $n$, its susceptibility is distributed according to $\rho(s, n)$.

Note that $\rho(s, 0)$ is the pdf of the personal-trait susceptibility of $A_0$, namely the pdf of $S_p$. This implies that the distribution is a discrete distribution (with support equaling the set formed by the $N_0$ values drawn initially). The support of $\rho(s, n)$ is a subset of that of $\rho(s, 0)$, following the epidemics process (at which individuals are removed from the susceptible population).

3. Evolution of the spread and the HIT

In this section we track the “natural” evolution of the process over the population, without using any preventive measures (e.g. lockdowns). In subsection 3.1 we analyze the change in the distribution (composition) of the susceptible population as individual contract the spread and become infected/recovered. In subsection 3.2 we track the decrease of $R(n)$ along the process, and provide an approximation of $R(n)$. In subsection 3.3 we provide the resulting derivation of the Herd Immunity Threshold (HIT). In subsection 3.4 we numerically evaluate and validate our result.

3.1. Tracking the spreading process

Let $A_n$ be a realization of the susceptible population at step $n$. We start by deriving $\Pr[a \in A_{n+1} \mid a \in A_n]$, where $A_{n+1}$ is a random variable distributed over all the possible susceptible populations at step $n + 1$, given $A_n$. That is, according to the infection probability.

Since the likelihood of getting infected is proportional to the susceptibility value of the individual, it holds that the probability that $a \in A_n$ is the next infected is
where $S(b, n)$ is the realization of the susceptibility parameter of individual $b$ at step $n$. Therefore,

$$\Pr[a \in A_{n+1} | a \in A_n] = \left( 1 - \frac{S(a, n)}{\sum_{b \in A_n} S(b, n)} \right).$$

(5)

Note that $\Pr[a \in A_{n+1} | a \in A_n] = \Pr[a \in A_{n+1}]/\Pr[a \in A_n]$, and by applying Eq. (5) recursively over $1, ..., n$ we have a probability for an individual $a \in A_0$ to be susceptible at step $n$:

$$\Pr[a \in A_n | a \in A_0] = \prod_{i=0}^{n-1} \left( 1 - \frac{S(a, i)}{\sum_{b \in A_i} S(b, i)} \right).$$

(6)

Eq. (6) holds for any $a \in A_0$, without any assumptions on the susceptibility parameters of $a$. Hence, it can be used to develop the ratio between $\rho(s, n)$ and $\rho(s, 0)$.

Recall that $\rho(s, 0)$ is the pdf of the personal-trait susceptibility of the initial population, $A_0$. In other words, the personal-trait susceptibility of the population, $S_p$, is distributed according to $\rho(s, 0)$. $\rho(s, n)$ is the pdf of the personal-trait susceptibility of the (susceptible) population at stage $n$. In other words, if an individual $x$ is picked at random at stage $n$, its personal-trait susceptibility, $S_p(x)$, is distributed according to $\rho(s, n)$. Hence,

$$\rho(s, n) = \frac{\rho(s, 0) \cdot \Pr[x \in A_n | x \in A_0 \text{ and } S_p(x) = s]}{\int \rho(\sigma, 0) \Pr[x \in A_n | x \in A_0 \text{ and } S_p(x) = \sigma] d\sigma}.$$  

(7)

According to Definition 1, the susceptibility parameter at step $i$ of individual $x$, for which $S_p(x) = s$, is given by $S(x, i) = p \cdot s + \tilde{p} \cdot S_e(x, i)$ where $S_e(x, i)$ is drawn independently from $S_e$ (and does not depend on $S_p(x) = s$ or on $i$). Denote by $\delta(n, s)$ the probability that an individual $x \in A_0$ such that $S_p(x) = s$ is part of the population $A_n$ (i.e. “survives up to step $n$”). Following Eq. (6) (full probabilistic analysis is provided in Appendix A):

$$\delta(n, s) \approx \prod_{i=0}^{n-1} \left( 1 - (p \cdot s + \tilde{p} \cdot \mathbb{E}[S_e]) \cdot \mathbb{E}[\delta(i)] \right)$$

(8)

where $S_e$ is a random variable drawn from $S_e$, and

$$\delta(i) = \frac{1}{\sum_{b \in A} S(b, i)}.$$

(9)
According to Eq. (7):

$$\rho(s, n) = \frac{\rho(s, 0) \cdot \delta(n, s)}{\int \rho(\sigma, 0) \cdot \delta(n, \sigma) d\sigma}$$

(10)

Remark 3.1. By Eqs. (8) and (10), for any $n$, the ratio $\frac{\rho(s, n)}{\rho(s, 0)}$ is decreasing in $s$. In addition, for any $n$, $\frac{\rho(s, n+1)}{\rho(s, n)}$ is decreasing in $s$ as well.

3.2. Basic and effective reproduction rate

The basic reproduction number, $R_0$, is a measure of how transferable a disease is. It is defined as the expected number of secondary cases produced by a single (typical) infection in a completely susceptible population (of size $N_0$).

As the spread continues, varying proportions of the population are recovered or removed at any given time. Hence, we will measure the effective reproduction number, $R(n)$, which is defined as the expected number of infections directly generated by the $n$th infected individual (e.g. $R_0 = R(0)$):

$$R(n) = E[\text{Number of infections generated by the } n\text{th infected individual}].$$

$R(n)$ can be approximated as follows:

$$R(n) \approx (N_0 - n) \int \rho(s, n) \cdot r(s) \, ds. \tag{11}$$

where $r(s) := (p \cdot s + \bar{p} \cdot E[S_e])(p \cdot \varphi(s) + \bar{p} \cdot E[I_e])$. Full analysis is given in Appendix B.

Remark 3.2 (Accuracy of the approximation in Eq. (11)). The approximation is based on assuming that one may infect itself (and adding this event to the count of expected infections, in the derivation of Eq. (11), which is provided in Appendix B). Hence the error is given by the probability of such an event, yielding a bound on the relative error by $O\left(\frac{\max_{a} S(a, n)}{\sum_{b} S(b, n)}\right)$.

Recall that we assume that no individual infects a significant fraction of the population, namely that $\sqrt{N_0} \gg O\left(\frac{\max_{a} S(a, n)}{E[S(b, n)]}\right)$, as was assumed in Oz et al., implying that $O\left(\frac{\max_{a} S(a, n)}{\sum_{b} S(b, n)}\right)$ is negligible. A further discussion is provided in Appendix B, and validation of the accuracy of the approximation is provided in Subsection 3.4.

Note that, intuitively, $R(n)$ is continuously decreasing as individuals become infected and then removed, and the size of the susceptible population decreases. Yet, the pdf of the personal-trait susceptibility, $\rho(s, n)$,
changes through the spreading process, which may cause a non-linear reduction of $R(n)$ in $n$.

### 3.3. Derivation of the herd immunity threshold

The Herd Immunity Threshold (HIT) is reached when $R(n)$ reduces below 1, and its value is measured in the fraction of the population that contracted the spread. Hence, in this subsection we bind the reduction of $R(n)$ to the value of $n$. We provide an approximation of the Herd Immunity Threshold by connecting $R(n)$ to $n$ using $\rho(\cdot, 0)$, as follows.

We begin with providing an approximation of Eq. (6). Using algebraic manipulations of Eq. (6), we show that:

$$
\Pr\left[ x \in A_n \mid x \in A_0 \text{ and } S_p(x) = s \right] \approx \exp\left( -\mathbb{E}[\Delta(n)](p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \right)
$$

where (recall Eq. (9))

$$
\Delta(n) = \sum_{i=0}^{n-1} \delta(i) = \sum_{i=0}^{n-1} \frac{1}{\sum_{b \in A} S(b, i)}
$$

The full derivation is given in Appendix B.

**Remark 3.3** (Accuracy of approximation in Eq. (12)). The relative error of Eq. (12) is analyzed in Appendix B and is demonstrated to be negligible.

Noting that:

$$
|A_n| = N_0 - n = N_0 \int \rho(s, 0) \cdot \Pr\left[ x \in A_n \mid x \in A_0 \text{ and } S_p(x) = s \right] \, ds
$$

and substituting Eq. (12) yields:

$$
\frac{n}{N_0} \approx 1 - \int \rho(s, 0) \cdot \exp\left( -\mathbb{E}[\Delta(n)](p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \right) ds. \tag{14}
$$

Using Eqs. (7) and (12), and similarly to Eq. (10), we have that

$$
\rho(s, n) \approx \frac{\rho(s, 0) \cdot \exp\left( -\mathbb{E}[\Delta(n)](p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \right)}{\int \rho(\sigma, 0) \cdot \exp\left( -\mathbb{E}[\Delta(n)](p \cdot \sigma + \bar{p} \cdot \mathbb{E}[S_e]) \right) d\sigma}. \tag{15}
$$

By Eqs. (11) and (15), combined with Eq. (14), the value of the effective reproduction number $R(n)$ is given by:

$$
R(n) \approx N_0 \cdot \int \rho(s, 0) \cdot \exp\left( -\mathbb{E}[\Delta(n)](p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \right) \cdot r(s) \, ds. \tag{16}
$$

Since we connected $R(n)$ to $n$ using $\Delta(n)$ (Eqs. (14) and (16)), we have that for any $\Delta > 0$, when
\[ \frac{n}{N_0} \approx 1 - \int \rho(s,0) \cdot \exp \left( -\Delta \cdot (p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \right) ds \]  

\[ \frac{\int \rho(s,0) \cdot \exp \left( -\Delta \left( p \cdot s + \bar{p} \cdot \mathbb{E}[S_e] \right) \right) \cdot r(s) ds}{\int \rho(s,0) \cdot r(s) ds} \]  

The threshold for Herd Immunity (HIT) is when the value of the effective reproduction number is \( R(n) \leq 1 \) (and hence the number of infection cases decreases).

**Remark 3.4.** The analysis in Oz et al.\cite{14} provides a result for the HIT which is a special case of our analysis whereby \( p = 1 \) (and \( \bar{p} = 0 \), that is, only pre-determined spreading type).

### 3.4. Numerical evaluation

**3.4.1. Validation of the approximation accuracy**

We validate the accuracy of the approximation provided in the above analysis in two ways. First, we use numerical simulations that provide excellent fit with the analytic HIT prediction. Second, we validate our assumption (see Remark 3.2) and demonstrate that under the spreading distribution attributed to COVID-19 it holds that \( \sqrt{N_0} \gg O\left( \max_{\mathbb{E}[S(a,n)]} \right) \).

The analysis allows to derive the HIT numerically (and in special case - analytically); Eq. (17) can be used to derive \( \Delta \), whose substitution into Eq. (18) yields the corresponding \( R(n) \), from which the HIT can be derived.

To validate the quality of the approximation presented in Eqs. (17) and (18) we compare their HIT prediction with that of Monte-Carlo simulations and find excellent fit. For the sake of completeness, the technical details of the simulations are provided in Appendix C.

Figure 1 depicts the value of the HIT derived as described above, along with the results of the simulations. The figure further demonstrates that the HIT is highly sensitive to \( p \), namely to the mix between the personal-trait and event-based spreading functions.

Next, we validate our approximation assumptions (described in Remarks 3.2 and 3.3). Recall that we assume that no individual infects a significant fraction of the population. In particular, we assumes that \( \sqrt{N_0} \gg O\left( \max_{\mathbb{E}[S(a,n)]} \right) \). In a numerical simulation, we performed the following experiment hundreds of times: We draw \( N_0 = 10,000,000 \) values from a
Gamma distribution with shape parameter $k = 0.1$ (which is attributed to COVID-19\cite{3}). After drawing the values, we compared $\max\{\text{value}\}$ to $\sqrt{N_0}$. At each of the runs, the latter was greater than the former, as required. 

The maximal value of $\frac{\max\{\text{value}\}}{\text{average\{value\}}}$ in all of the runs was 142.88, while $\sqrt{N_0}$ equals to 3162.27.

3.4.2. The behavior of the contributing factors of $R(n)$

Having a formula of the value of $R()$ as a function of $n$, we use it to demonstrate the behavior of the two spreading functions, personal-trait and event-based, whose dynamics drastically differ from each other. Figure 2 depicts the decay of the factors contributing to $R(n)$, classified by their spreading types. In red: $R(n)$; In blue: Personal-trait pure contribution to $R(n)$; In green: Event-based pure contribution to $R(n)$; In yellow: Mutual contribution to $R(n)$. As can be seen, the contribution of the personal-trait spreading drops sharply at early stages of the disease. On the other hand, the contribution of the event-based spreading is affected very little at early stages.

This is because super-spreaders (individuals with high values of $S_p()$ and $I_p()$) are likely to get infected early in the pandemic, and their resultant immunity then decreases the level of personal-trait spreading. In contrast, any reduction of the event-based spreading level is merely proportional to

![Figure 2](image)

**Figure 2.** The over-time reduction in the effective reproduction number, $R(n)$, and its contributing factors as a function of the infection index, $n$, for $R_0 = 3$, $p = 0.5$, and a Gamma distribution with shape $k = 0.1$. Note that $n$ (horizontal-axis) is normalized to percentage.
the decrease in susceptible members of the population, which is linear in \( n \), as it is identically distributed for all individuals.

This difference in the behavior of the two spreading functions has a crucial impact on the effectiveness of different lockdown strategies, and on their consequences, as will be studied in the next section.

4. The effects of control mechanisms: not all lockdowns are born equal

The spreading process can be slowed down by limiting the interaction among the population (network entities). In the case of a pandemics, numerous stay-at-home orders, quarantines, and various other rulings can be enforced. We refer to such restraining methods as lockdowns.

A lockdown will reduce the level of interaction between the individuals and thus “collapse” the effective reproduction number temporally, while it is being enforced. However, in many cases the effects of the lockdown on the post-lockdown spread dynamics, that is, after the lockdown is lifted, remain unclear and these lie at the core of this research.

We show that the long-term effects of lockdowns on the dynamics of the spread are very sensitive to the lockdown type. While some lockdowns affect the disease blocking positively and reduce the HIT – others may generate adverse effects that increase the HIT and may even increase the total number of infected individuals in the long-run, after the lockdown is released. These properties will be established in Theorems 4.3 and 4.4 below.

Note that the results of this section do not require knowledge of the population distribution; rather the section provides general properties of lockdown policies based on the two spreading types.

4.1. Classifying lockdown types

We classify the restrictions into two inherently different types, according to their focus, formally defined as follows:

**Definition 4.1** (Personal-trait (continual) spreading targeted lockdown). For any \( n \) during a personal-trait (continual) spreading targeted lockdown, the personal-trait infectiousness and susceptibility of an individual \( a \) are blocked while the event-based infectiousness and susceptibility remain as is. That is,

\[
I_L(a, n) = p_L \cdot I_p(a) + \bar{p} \cdot I_e(a, n) \quad \text{and} \quad S_L(a, n) = p_L \cdot S_p(a) + \bar{p} \cdot S_e(a, n),
\]

where \( p_L = 0 \) and \( \bar{p} \) remains at its original value.
Definition 4.2 (event-based (occasional) spreading targeted lockdown). For any \( n \) during an event-based (occasional) spreading targeted lockdown, the event-based infectiousness and susceptibility of an individual \( a \) are blocked while the personal-trait infectiousness and susceptibility remain as is. That is,

\[
I_L(a, n) = p \cdot I_p(a) + \tilde{p}_L \cdot I_e(a, i) \quad \text{and} \quad S_L(a, n) = p \cdot S_p(a) + \tilde{p}_L \cdot S_e(a, i),
\]

where \( \tilde{p}_L = 0 \) and \( p \) remains at its original value.

For example, a personal-trait spreading targeted lockdown might include closing or restricting workplaces. An event-based spreading targeted lockdown will include prohibition on cultural gatherings such as sporting events or music concerts.

In order to analyze the effect of each lockdown, we compare two systems: a natural-evolution system, and a lockdown system. The first is subject to “natural” evolution of the spreading process over the population, without any preventive-measures being used, as was analyzed in Section 3. The second is subject to a lockdown. Denote by \( n_b \) and \( n_e \) the steps in the lockdown system at which the lockdown begins and ends, respectively. We couple the natural-evolution system to the lockdown system via \( n \), the infected individual index, and by relating it to \( n_b \) and \( n_e \). Let \( R(n) \) and \( R_L(n) \) be the values of the effective reproduction number for the natural-evolution system and the lockdown system, respectively.

For any \( n \leq n_b \) (i.e. before the lockdown begins in the lockdown system) the two systems are governed by the same spreading process and behave the same. For \( n_b \leq n \leq n_e \), the systems separate since the value of \( p \) (or \( \tilde{p} \)) governing the lockdown system drops, which affects the values of the infectiousness and susceptibility parameters. Consequently, the distribution of the population, \( \rho(\cdot) \), will experience a different progressions between the two systems. For \( n > n_e \) (i.e. following the release of the lockdown in the lockdown system), the two systems are governed again by the same spreading process since the value of \( p \) (or \( \tilde{p} \)) returns to its original value in the lockdown system. Yet, since at the release of the lockdown the distribution of the population differs across the systems, the relation between \( R(n) \) and \( R_L(n) \) is unclear and is investigated next.

We prove the following theorems:

Theorem 4.3 (HIT of personal-trait spreading targeted lockdown). Assume that a personal-trait spreading targeted lockdown was performed for \( n \in [n_b, n_e] \). Then for any \( n \geq n_e \)

\[
R_L(n) \geq R(n).
\]

Consequently, the HIT will increase (in comparison to a natural evolution).
Theorem 4.4 (HIT of event-based spreading targeted lockdown). Assume that an event-based spreading targeted lockdown was performed for \( n \in [n_b, n_c] \). Then for any \( n \geq n_c \)

\[ R_L(n) \leq R(n). \]

Consequently, the HIT will decrease (in comparison to a natural evolution).

The relation between the HIT values is implied directly from the relation between \( R(n) \) and \( R_L(n) \) values by recalling that the parameter \( n \) in \( R(n) \) and \( R_L(n) \) indicate the \( n \)th infected individual.

Due to the generality of Theorems 4.3 and 4.4, we state the following corollary:

**Corollary 4.5.** Any sequence of personal-trait spreading targeted lockdowns will increase the HIT. Similarly, any sequence of event-based spreading targeted lockdowns will decrease the HIT.

The structure of the analysis is as follows: In Subsection 4.2 we establish a general property of stochastic monotonicity preservation between systems which are subject to the same process. In Subsections 4.3 and 4.4 we analyze the stochastic relations between the systems as a result of a lockdown (i.e. establish monotonicity between them at the release of the lockdown). Since at the period following the release of the lockdown, both system are governed by the same process, we use the preservation property established in Subsection 4.2, and using the monotonicity relations based in Subsections 4.3 and 4.4 we conclude the proofs of Theorems 4.3 and 4.4.

### 4.2. Preservation of stochastic monotonicity between systems subject to the same process

We prove preservation of monotone likelihood ratio property between the population distributions of systems subject to the same process (i.e. same values of \( p \) and \( \bar{p} \)). This is established in Lemma 4.7.

**Definition 4.6.** (see\cite{12,18}) The pair of probability density functions \((\rho_1(s,n), \rho_2(s,n))\) possesses the monotone likelihood ratio property (MLRP) if \( \frac{\rho_1(s,n)}{\rho_2(s,n)} \) is non-decreasing in \( s \).

**Lemma 4.7.** Let \( A_1, A_2 \) be two populations of size \( N_0 - n \), going through the same spreading process (i.e., same values of \( p \) and \( \bar{p} \)). Let the personal-trait susceptibility of \( A_1 \) and \( A_2 \), \( S_p(A_1) \) and \( S_p(A_2) \), be distributed according to \( \rho_1(s,n) \) and \( \rho_2(s,n) \), respectively. Assume that the pair \((\rho_1(s,n), \rho_2(s,n))\)
possesses the monotone likelihood ratio property. Then the pair \((\rho_1(s,n+1),\rho_2(s,n+1))\) possesses the monotone likelihood ratio property as well.

In order to prove the lemma, we will use the following three known theorems:

**Theorem 4.8.** (Wolfstetter et al.\([18]\)) Let \(X_1, X_2\), be random variables. Let \(\rho_1(), \rho_2()\) be their pdfs. If the pair \((\rho_1(), \rho_2())\) possesses the monotone likelihood ratio property then \(X_1\) (first order) stochastically dominates \(X_2\) as follows: \(X_1 \succeq_{\text{FSD}} X_2\)\(^{10}.\)[4,18]

**Theorem 4.9** (Hadar and Russell\([4]\)). Consider three independent non-negative random variables \(X_1, X_2\) and \(Y\), and the linear combinations \(aX_1 + bY\) and \(aX_2 + bY\), with \(a > 0, b > 0\). Then:

\[
X_1 \succeq_{\text{FSD}} X_2 \Rightarrow (aX_1 + bY) \succeq_{\text{FSD}} (aX_2 + bY).
\]

**Theorem 4.10.** (Wolfstetter et al.\([18]\)) Let \(r()\) be an injective monotone function. If \(X_1 \succeq_{\text{FSD}} X_2\) then \(E[r(X_1)] \geq E[r(X_2)]\). In particular, \(X_1 \succeq_{\text{FSD}} X_2 \Rightarrow E[X_1] \geq E[X_2]\).

Proof of Lemma 4.7. We start by deriving \(\rho_1(s,n+1)\) using \(\rho_1(s,n)\) (the same can be done for \(\rho_2\)). Recall Eqs. (5), (8), (10), and the analysis in Appendix A; It holds that

\[
\rho_1(s,n+1) = \frac{\rho_1(s,n) \cdot \left(1 - (p \cdot s + \tilde{p} \cdot E[S_c]) \cdot E[\delta_1(n)]\right)}{\int \rho_1(\sigma,n) \cdot \left(1 - (p \cdot \sigma + \tilde{p} \cdot E[S_c]) \cdot E[\delta_1(n)]\right) d\sigma}.
\]

where \(\delta_1(n) = \left[\frac{1}{\sum_{b \in A_1} s(b,n)}\right]\). Note that \(\int \rho_1(\sigma,n) \cdot (1 - (p \cdot \sigma + \tilde{p} \cdot [S_c]) \cdot \delta_1(n) d\sigma\) is identical for all values of \(s\). Denote it by \(c_1\). Repeat this analysis for \(\rho_2(s,n)\) and derive \(c_2\) and \(\delta_2\) (corresponding to \(c_1\) and \(\delta_1\)). Hence,

\[
\frac{\rho_1(s,n+1)}{\rho_2(s,n+1)} = \frac{\rho_1(s,n)}{\rho_2(s,n)} \cdot \frac{c_2}{c_1} \cdot \frac{1 - (p \cdot s + \tilde{p} \cdot E[S_c]) \cdot E[\delta_1(n)]}{1 - (p \cdot s + \tilde{p} \cdot E[S_c]) \cdot E[\delta_2(n)]}.
\]

By Theorem 4.8 the random variables \(S_p(A_1), S_p(A_2)\) which correspond to the densities \(\rho_1(s,n), \rho_2(s,n)\) obey that \(S_p(A_1)\) stochastically dominates \(S_p(A_2)\), since the pair \((\rho_1(\cdot,n), \rho_2(\cdot,n))\) was assumed to possess MLRP. Recall that \(S(a,n) = p \cdot S_p(a) + \tilde{p} \cdot S_c(a,n)\). Using Theorems 4.9 and 4.10 we have that \(E_{b \sim A_1}[S(b,n)] \geq E_{b \sim A_2}[S(b,n)]\). Therefore,

\[
E[\delta_1(n)] = E\left[\frac{1}{\sum_{b \in A_1} S(b,n)}\right] \leq E\left[\frac{1}{\sum_{b \in A_2} S(b,n)}\right] = E[\delta_2(n)].
\]
For any $n_b$, Claim 4.11. The ratio $\rho(s,n_e)/\rho(s,n_b)$ is decreasing in $s$ in the natural-evolution system, and hence the pair $(\rho(s,n_b),\rho(s,n_e))$ possesses MLRP. Following Eq. (22), the pair $(\rho_L(s,n_e),\rho(s,n_e))$ possesses MLRP as well. □

Next, we relate the monotone likelihood ratio of the probability density functions, $\rho()$ and $\rho_L()$, to a relation between their corresponding infection rates, $R()$ and $R_L()$. This is demonstrated in Figures 3 and 4.

Claim 4.12. Let $n > n_e$. If $(\rho_L(s,n),\rho(s,n))$ possesses MLRP then $R_L(n) \geq R(n)$.

Proof of Claim 4.12. The proof assumes approximation of $R(n)$ being exact; see Remark 3.2 that justifies the accuracy of Eq. (11). Recall that after the lockdown is released, for $n > n_e$, both systems have the same values of $p$ and $\bar{p}$. By Eq. (11), $R(n) \approx (N_0 - n) \cdot \int \rho(s,n) \cdot r(s) ds$ where $r(s) =
Since we assumed that $u(s)$ is monotonically non-decreasing in $s$ (i.e. monotonicity of spreading), we have that $r(s)$ is an injective monotone function. Hence, by Theorems 4.8, 4.9 and 4.10 it holds that $R(n) \leq R_L(n)$.
**Proof of Theorem 4.3.** According to Claim 4.11, \((\rho_L(s, n_e), \rho(s, n_e))\) possesses MLRP. Hence, by Lemma 4.7, for any \(n > n_e\), \((\rho_L(s, n), \rho(s, n))\) possesses MLRP. Therefore, by Claim 4.12, \(R(n) \leq R_L(n)\) for any \(n > n_e\), and we conclude the proof.

### 4.4. Event-based lockdown: behavior during the lockdown and proof of Theorem 4.4

We analyze the behavior of \(\rho_L(s, n)\) during the lockdown period, that is \(n_b \leq n \leq n_e\). We establish the MLRP of \(\rho(s, n)\) and \(\rho_L(s, n)\) during this period, and in particular at \(n_e\) (Claim 4.13), and using Lemma 4.7 we prove Theorem 4.4.

**Claim 4.13.** For any \(n_b \leq n \leq n_e\), the pair \((\rho(s, n), \rho_L(s, n))\) under an event-based lockdown possesses the MLRP property.

**Proof of Claim 4.13.** For \(n = n_b\) the claim holds trivially due to equality of the pdfs \((\rho(s, n_b) = \rho_L(s, n_b))\). Next we establish the property for any \(n_b \leq n \leq n_e\), using an induction on \(n\). Let \(n_b \leq n \leq n_e\), and assume that the pair \((\rho(s, n), \rho_L(s, n))\) possesses the MLRP. By Eq. (20), since \(\bar{p} = 0:\)

\[
\frac{\rho(s, n + 1)}{\rho_L(s, n + 1)} = \frac{\rho(s, n)}{\rho_L(s, n)} \cdot \frac{c}{c_L} \cdot \frac{1 - p \cdot s \cdot \mathbb{E}[\delta(n)]}{1 - p \cdot s \cdot \mathbb{E}[\delta_L(n)]}
\]  

where \(c, c_L\) corresponds to \(c_1, c_2\) and \(\delta, \delta_L\) to \(\delta_1, \delta_2\), respectively. By Theorems 4.8 and 4.10, since \((\rho(s, n), \rho_L(s, n))\) possesses the MLRP,

![Figure 5](image-url)  

**Figure 5.** The value \(R_L(n)\) under an event-based spreading targeted lockdown (vis-a-vis \(R(n)\) in a natural evolution) for \(p = 0.5\), \(R_0 = 3\) and \(k = 0.1\). The lockdown begins at \(n_b = 5\%\) and ends at \(n_e = 30\%\).
The latter equation implies that $E_{b \in A_n \mid S_L(b, n)} \geq \frac{E_{b \in A_n \mid S(b, n)}}{E_{b \in A_n \mid S_L(b, n)}}$ and this implies that Eq. (23) is non-decreasing in $s$, as it is a product of such non-decreasing ratios. That is, the pair $(\rho(s, n + 1), \rho_L(s, n + 1))$ possesses MLRP.
We state the following claim corresponding to Claim 4.12:

**Claim 4.14.** Let \( n > n_c \). If \( (\rho(s,n), \rho_L(s,n)) \) possesses MLRP then \( R(n) \geq R_L(n) \).

Proof of Claim 4.14. The proof follows similar arguments to those used in the proof of 4.12. \( \square \)

**Proof of Theorem 4.4.** According to Claim 4.13, \( (\rho(s,n_c), \rho_L(s,n_c)) \) possesses MLRP. Hence, by Lemma 4.7, for any \( n > n_c \), \( (\rho(s,n), \rho_L(s,n)) \) possesses MLRP. Therefore, by Claim 4.14, \( R(n) \geq R_L(n) \) for any \( n > n_c \) (demonstrated in Figure 5 and Figure 6), and we conclude the proof. \( \square \)

**Remark 4.15.** Note that the proof of Claim 4.13 was not based on the assumption that \( \bar{p}_L = 0 \), but on the fact \( \bar{p}_L < \bar{p} \). Hence, Theorem 4.4 can be generalized to hold for any \( 0 \leq \bar{p}_L < \bar{p} \), namely for Partial Lockdowns. To demonstrate this, see Figure 7 which depicts the progression of \( R_L \) under a partial lockdown.

5. Numerical evaluations and discussion

5.1. COVID-19: HIT

Our results can be used to inspect the sensitivity of the HIT to the social characteristics of a society, and to deviations in the value of \( p \). Per the analysis in Section 3 and Figure 1, we find that the HIT value is very sensitive to \( p \), the relative weight of the spreading types, and thus different societies may engender significantly different HITs, even if they have the same \( R_0 \), or the same distributions \( S_p, I_p, S_e \) and \( I_e \). In addition, and in accordance with previously reported results,\(^{[2,14]}\) the HIT is influenced by the coefficient of variation (\( CV \)) of the personal-trait spreading distributions, and by the correlation between the infectiousness and susceptibility parameters.

This is demonstrated in Figure 8 which plots the HIT as a function of both the relative weight \( p \), and the shape of the personal spreading distribution (i.e. “heterogeneity level”).

Note that for a fully event-based spreading society (i.e. \( p = 0 \)) the HIT is as high as 66.66%, coinciding with the allegedly “axiomatic” cutpoint of the traditional homogeneous models (where the HIT is equal to \( 1 - 1/R_0^{[1,13]} \)).

In Figures 3 and 4 we evaluate the progression of \( R_L \) under a personal-trait spreading targeted lockdown vis-a-vis a natural evolution for \( p = 0.5 \) and \( p = 0.75 \), respectively. As expected, both figures show that the HIT under the personal-trait lockdown evolution increased in comparison to the natural evolution, consistent with our analysis (Theorem 4.3).

Further, we can see that the relative change of the HIT value (the point where \( R() \) reduces below 1) is more drastic in the case where \( p = 0.75 \).
(rather than when $p = 0.5$). This results from the fact that, somewhat intuitively, restraining the personal-trait infections will have more impact on the stochastic progress of the disease when their relative weight in the infection process is higher.

Similar observations can be made with respect to event-based spreading targeted lockdowns; Figures 5 and 6 provide the comparison between $p = 0.5$ and $p = 0.75$, respectively, for this case.

5.2. Practical implications and considerations

The opposite behavior of the studied lockdowns (as established in Section 4) results from the modeling assumption whereby the personal-trait susceptibility and infectiousness parameters are heterogeneous across the population and static over time, whereas the event-based parameters are homogeneous across the population and dynamically redrawn over time. Thus, the analysis demonstrates that lockdowns which target (namely, limit or prohibit) social/occasional events will suppress the homogeneous part of the spreading and cause a stronger “heterogeneous spread”, which lowers the HIT (as observed in Section 3); in contrast, the personal-trait targeted lockdown will suppresses the heterogeneity and make it behave like an epidemic in an homogeneous population (in which the HIT is relatively large).

5.2.1. Simulation results and lockdown magnitude

We use Monte-Carlo simulations to demonstrate the effects of different lockdown policies on the Herd Immunity Threshold and the magnitude of
the post-lockdown “wave”. In case where the HIT is not reached when the lockdown is released, the subsequent natural increase in interpersonal interactions within the population may then be responsible for causing another “wave” of the pandemic. Our results reveal that the lockdown type may have significant effect on the magnitude of the post-lockdown “wave”.

To this end we demonstrate that two lockdowns of similar “magnitude” (namely, causing similar temporal reduction in the number of infections during the lockdown) but of a different type, can result in significantly different numbers of infections in the post-lockdown period. In Figure 9 we simulate the full evolution of a pandemic under two different lockdown policies of similar magnitude (grey area), exercised on the same population. That is, we consider a set of individuals and track the disease spread among them. As projected by our results, the event-based lockdown decreased the HIT (relatively to no-lockdown) while the personal-trait lockdown increased the HIT. The latter resulted with a significantly higher second wave, and caused an increase in the total number of infection cases as well.

5.2.2. Applying the results

Our analysis reveals that the post-lockdown effect of each policy should be accounted for, and that it might cause two “looking similar” lockdowns (same temporal reduction in the interactions level) to result with dramatically different post-lockdown evolution, which affects the optimal policy. Yet, the
results should be applied with some caution, as they depend on the model assumptions (which assert that the spreading consists of two types – static heterogeneous one and dynamic homogeneous one). Modeling the human-interaction network using both spreading types seems to be a natural model of reality. Note also that often lockdown policies, or their practical implementation, do not form a fully personal-trait targeted lockdown or a fully event-based targeted lockdown; Rather they set partial limitations on either of the spreading types (e.g. closing non-essential workplaces, possibly combined with prohibiting music concerts). For such policies, one may use the model and the machinery provided in this paper to model the effects of these policies, and to study the evolution of the process under them.

5.2.3. Short-term and long-term Tradeoffs
In real-world epidemics the Herd Immunity is not the sole consideration used to guide lockdowns and other restrictions. In some cases, the goal of a lockdown is to achieve an immediate effect such as to reduce the daily number of new cases to prevent healthcare systems from getting overloaded. Further economic or social considerations, or the attempt to limit the number of infections until a vaccine is developed, may also guide such restrictions. In practice, these short-term objectives must be weighed against the long-term objective of Herd Immunity. Accounting for long term is crucial, since in many cases planned lockdowns may be lifted ahead of time (due to social/economic objectives), yielding repeating waves where the long term is of high importance.

In the event that such short-term considerations are the main factor leading decision makers to conduct a lockdown, our results about the opposite effects of different lockdowns (certain lockdowns increase the HIT while others reduce it) may significantly affect the specific lockdown strategy taken.

To articulate this, consider Figure 9 which demonstrates that two lockdowns (one event-based spreading targeted and the other personal-trait spreading targeted) which are carried for the same amount of time and reduce the temporal infection number of cases to the same level of infections per day, still have completely different long-term effects. Thus, even if a decision maker aims to perform a lockdown toward a short-term objective, the revealed effects of the lockdown type on the HIT allow the decision maker obtaining the short-term objective while selecting the proper lockdown type to optimize for the long-term as well.

The machinery provided in this paper allows one to evaluate (via numerical evaluation) both the short- and long-term effects and optimize accordingly.
6. Concluding remarks and further work

In this work we proposed a new stochastic model which classifies the infectiousness and susceptibility of individuals into two inherently different types: personal-trait (continual) and event-based (occasional) spreading. The classification was motivated by the heterogeneity observed in the ongoing COVID-19 pandemic and other real-life networks.

Consequently, studying the spreading dynamics requires the analysis of a stochastic process consisting of two stochastic functions which affect each other, a static one and a dynamic one, differing drastically in their evolution. We analyzed the stochastic process governing the system and tracked the evolution in the distribution of the population throughout the process. We derived the value of the Herd Immunity Threshold and showed that it is very sensitive to the relative weight of the two spreading types.

Having gained an understanding of the spread behavior in a no-intervention environment (i.e. natural evolution), we further addressed operational aspects and examined various lockdowns strategies. This required the use of additional probabilistic treatments and techniques, in establishing stochastic relations between different lockdown strategies. Our model reveals that the effect of lockdowns on the long term behavior of the pandemics, namely on HIT, is very sensitive to their focus: while event-based-targeted lockdowns reduce the HIT, personal-trait-targeted lockdowns increase HIT and may be counter-productive in the long run.

Note that these results hold for any relative weight of the spreading types in a network, and any spreading distributions, including various heterogeneity levels.

The generality of the model allows it to treat other problems as well. In an ongoing research we utilize it to study the effects of vaccinations and to apply it to computer networks viruses.

A. Probabilistic derivation of Eq. (8)

We derive the probability \( \Pr[x \in A_{i+1} | x \in A_i \land S_p(x) = s] \), using the derivation of \( \Pr[x \not\in A_{i+1} | x \in A_i \land S_p(x) = s] \). Recall that \( S(a,i) = p \cdot S_p(a) + \bar{p} \cdot S_e(a,i) \), where \( S_p(a) \) is pre-determined and \( S_e(a,i) \) is drawn (independently for each individual) from the distribution \( S_e \). Let \( x \in A_i \) such that \( S_p(x) = s \). Let \( \{s^b_e\}_{b \in A_i} \) be a realization of \( \{S_e(b,i)\}_{b \in A_i} \):

\[
\Pr[x \not\in A_{i+1} | x \in A_i \land S_p(x) = s \land \forall b \in A_i : S_e(b,i) = s^b_e] = (p \cdot s + \bar{p} \cdot s^e) \cdot \left( \frac{1}{\sum_{b \in A_i} S(b,i)} \right) = (p \cdot s + \bar{p} \cdot s^e) \cdot \left( \frac{1}{\sum_{b \in A_i} p \cdot S_p(b) + \sum_{b \in A_i} \bar{p} \cdot s^e} \right).
\]
Hence, by un-conditioning we get:

\[
\Pr[x \notin A_{i+1} | x \in A_i \land S_p(x) = s] = \sum_{\{s_i\} \in (\text{Supp}(S_i))^{A_i}} \left( \prod_{b \in A_i} \Pr[S_e(b, i) = s'_e] \cdot (p \cdot s + \bar{p} \cdot s'_e) \left( \frac{1}{\sum_{b \in A_i} p \cdot S_p(b) + \sum_{b \in A_i} \bar{p} \cdot S_e(b, i)} \right) \right).
\]

This expression can be separated to two. The first is (recall the definition of \(\delta(i)\) given in Eq. (9), \(\delta(i) = \frac{1}{\sum_{b \in A} s(b)}\)):

\[
(p \cdot s) \cdot \mathbb{E}\left[ \frac{1}{\sum_{b \in A_i} p \cdot S_p(b) + \sum_{b \in A_i} \bar{p} \cdot S_e(b, i)} \right] = p \cdot s \cdot \mathbb{E}[\delta(i)].
\]

as the expected value of the denominator is independent of \(s\). The second is:

\[
\mathbb{E}\left[ \frac{\bar{p} \cdot S_e(x, i)}{\sum_{b \in A_i} p \cdot S_p(b) + \sum_{b \in A_i} \bar{p} \cdot S_e(b, i)} \right].
\]

Denote the random variable expressed by this fraction by \(\delta'(i)\). Note that the \(S_e(x, i)\) in the numerator is not independent of the \(\sum_{b \in A_i} q \cdot S_e(b, i)\) expression in the denominator, as it is included in it. In order to approximate the expected value of \(\delta'(i)\), we may use the fact that \(|A_i|\) is large, and assume that \(S_e(x, i)\) and \(\sum_{b \in A_i} S_e(b, i)\) are approximately independent. Hence, \(\mathbb{E}[\delta'(i)] \approx \bar{p} \cdot \mathbb{E}[S_e] \cdot \mathbb{E}[\delta(i)]\), where \(S_e\) is a random variable drawn from \(S_e\), and

\[
\Pr[x \in A_{i+1} | x \in A_i \land S_p(x) = s] \approx 1 - (p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \cdot \mathbb{E}[\delta(i)].
\]  

(24)

Hence, \(\delta(n, i) = \Pr[x \in A_n | x \in A_0 \land S_p(x) = s] \approx \prod_{i=0}^{n-1} (1 - (p \cdot s + \bar{p} \cdot \mathbb{E}[S_e]) \cdot \mathbb{E}[\delta(i)]).
\]

**B. Derivation of Eqs. (11) and (12)**

**Derivation of Eq. (11)**
The expected value of the effective reproduction number combines for each individual \(a\) the probability that it will be the \(n\)th infected, and the expected number of secondary cases \(a\) will cause. That is,

\[
R(n) = \mathbb{E}_a \left[ \Pr[a \text{ is the } n \text{th infected}] \cdot (\text{infectiousness of } a) \cdot \left( \sum_{b \in A_n, b \neq a} \text{ (susceptibility of } b) \right) \right]
\]

where the expectation is taken over all individuals and all possible scenarios of infection.

Let \(A_n\) be a realization of the susceptible population at step \(n\). The value of the effective reproduction number under the realization \(A_n\), denoted by \(R_{A_n}(n)\), is given by:

\[
R_{A_n}(n) = \mathbb{E} \left[ \sum_{a \in A_n} \sum_{b \in A_n, b \neq a} \frac{S(a, n)}{S(b, n)} \cdot I(a, n + 1) \cdot \sum_{b \in A_n} S(b, n + 1) \right] = \mathbb{E} \left[ \frac{\sum_{b \in A_n, b \neq a} S(b, n + 1)}{\sum_{b \in A_n} S(b, n)} \cdot \sum_{a \in A_n} S(a, n) \cdot I(a, n + 1) \right].
\]

(25)

where the expectation is over the values of \(S\) and \(I\). This can be approximated by:\n
12
\[ R_{A_n}(n) \approx \mathbb{E}\left[ \sum_{a \in A_n} S(a, n) \cdot I(a, n + 1) \right]. \]  

(26)

By taking expectation over \( A_n \) of \( R_{A_n} \) (Eq. (26)) it holds that:

\[ R(n) \approx (N_0 - n) \cdot \int \rho(s, n) \cdot (p \cdot s + \tilde{p} \cdot \mathbb{E}[S_s]) \left( p \cdot \varphi(s) + \tilde{p} \cdot \mathbb{E}[L]\right) ds. \]  

(27)

**Derivation of Eq. (12)**

The analysis follows the methodology used in the approximation provided in the proof of Claim I in Oz et al.\[14\] Recall Eq. (6): \( \Pr[a \in A_n] = \prod_{i=0}^{n-1} \left( 1 - \frac{N(a, i)}{\sum_{b \in A_i} S(b, i)} \right) \). Taking natural log,

\[ \log (\Pr[a \in A_n]) = \sum_{i=0}^{n-1} \log \left( 1 - \frac{S(a, i)}{\sum_{b \in A_i} S(b, i)} \right). \]  

(28)

We use the bound

\[ -x - x^2 \leq \log (1 - x) \leq -x, \quad \forall x \in [0, 1/2] \]

and have that:

\[
\log (\Pr[a \in A_n]) = \sum_{i=0}^{n-1} \left( -\frac{S(a, i)}{\sum_{b \in A_i} S(b, i)} - O\left( \frac{\left( S(a, i) \right)^2}{\sum_{b \in A_i} S(b, i)} \right) \right)
\]

\[ = -\sum_{i=0}^{n-1} \frac{S(a, i)}{\sum_{b \in A_i} S(b, i)} - O\left( n \cdot \left( \max_{b \in A} S(b, i) \right)^2 \right) \]

\[ = -\sum_{i=0}^{n-1} \frac{S(a, i)}{\sum_{b \in A_i} S(b, i)} - O\left( n \cdot \frac{\max_{b \in A} S(b, i)^2}{\mathbb{E}[S(b, i)]} \right). \]  

(29)

which is similar to the expression derived in Oz et al.\[14\] Like there, we use the assumption that \( \sqrt{N_0} \gg O\left( \max_{i \in [\mathbb{N}]} S(a, i) \right) \), and since the number of steps, \( n \), is bounded by \( N_0 \), we have the following approximation (similar to there):

\[ \log (\Pr[a \in A_n]) \approx -\sum_{i=0}^{n-1} \frac{S(a, i)}{\sum_{b \in A_i} S(b, i)}. \]  

(30)

Note that the error in approximation (30) is the right term of (29) which is very small relatively to (30).

According to Definition 1, \( S(a, i) \) is constructed as: \( S(a, i) = p \cdot S_p(a) + \tilde{p} \cdot S_\Sigma(a, i) \). and (Eq. (8)) \( \delta(i) = \frac{1}{\sum_{b \in A} S(b, i)} \). Following the analysis provided in Appendix A and by Eq. (30) we have that

\[ \log \Pr[a \in A_n] \approx -(p \cdot S_p(a) + \tilde{p} \cdot \mathbb{E}[S_a]) \cdot \mathbb{E}[\Delta(n)] \]

where \( \Delta(n) = \sum_{i=0}^{n-1} \delta(i) \).

### C. Technical details of the Numerical simulations

In our simulations we consider a set of individuals and follow the disease spread among them. At the beginning of the process, the entire population is Susceptible except for a small number of "Infected" individuals. Each infected individual \( a \) has an incubation period, which starts at the exposure moment of \( a \) and lasts for \( t_0 \) days, until \( a \) becomes infectious.

Afterwards, a is infectious for a period of \( t_1 \) days (Infectious period). During its infectious period, a will infect a susceptible individual b with probability \( I(a) \times S(b) \), where the exposure moment of b is spread uniformly over the infectious period of a.

\( I(a) \) and \( S(b) \) are calculated according to Eqs. (1) and (2). We assign to each individual, in the beginning of the process, personal-trait susceptibility and infectiousness parameters, \( S_p(a) \) and \( I_p(a) \). These values accompany the individuals throughout the entire simulation and remain at the same values.

Additionally, at each time period, we draw the event-based susceptibility and infectiousness parameters, \( S_e(a) \) and \( I_e(a) \) of each of the infectious and susceptible individuals. We maintain a queue of the currently infected individuals.

The selection of \( t_0 \) and \( t_1 \) does not affect the analysis results (i.e. it does not bias the effect of the lockdown policy on the HIT). This was verified by repeating the simulations over several \( t_0 \) and \( t_1 \) values.

We let the disease spread until there are no new infections and all the individuals either remain susceptible or become recovered.

**Acknowledgments**

This research was supported in part by the Israel Science Foundation 10.13039/501100003977 (grant No. 2482/21) and by the Blavatnik Family Foundation 10.13039/100011643.

**Disclosure statement**

The authors declare they have no conflicting financial interests.

**Notes**

1. Where connections between individuals form a network.
2. In this work we adopt the SIR model, under which one can not be re-infected. The SIR model will be further described in Section 2.
3. Gamma distribution with shape parameter \( k \approx 0.1 \), and \( R_0 \approx 3^{[3]} \)
4. Our analysis focuses on reducing the number of infection cases prior to reaching Herd Immunity. Note that lockdowns may have other positive short-term effects (e.g., reducing overloads of the healthcare system) which are not at the focus of this study; a short discussion on this subject is provided in Section 5.
5. We model stochastic number of infections according to general analysis of arbitrary distributions, this can accommodate modeling of arbitrary lengths of the infectious period. The simulations whose results are provided in Section 3, Figure 1 and Section 5, Figure 9 have been repeated with several different infectious period lengths and verified that these lengths do not affect the results of our analysis.
6. If an individual goes to many weddings or music concerts (i.e. significantly more than the average), then this will be modeled by assigning the individual a higher personal-trait spreading parameter.
7. Practically – these values can be produced by sampling a population. Examples of how they were estimated can be found in\(^{[3]}\) and\(^{[11]}\). If one wants to attribute them to a specific theoretical distribution, say Gamma (\( \ldots \)) one may draw each of the values from that distribution and obtain \( S_p \) and \( I_p \) that approximate the theoretical Gamma very well.
8. Note that alternatively to using $p$ and $\tilde{p}$, one could use $p$ and an arbitrary $0 \leq q \leq 1$. That however will not increase the modeling power since the selection of $p$, $S_p, S_e$ (and $T_p, T_e$) provides equivalent modeling power.

9. For discrete random variables one should treat this notation ($\int$) as the corresponding sum.

10. Recall that $X_1$ (first order) stochastically dominates $X_2$ if $\Pr[X_1 > z] \geq \Pr[X_2 > z]$ for all $z$.

11. It should be noted that in many cases, economic or social considerations forced lockdown restrictions to be lifted as soon as the number of infections dropped below some containable level, and before the HIT is reached. Thus, studying the effects of on the post-lockdown spread are of high importance.

12. In the approximation we assumed approximate equality of $\sum_{b \in A_n} S(b, n)$ and $\sum_{b \in A_n} S(b, n + 1)$ since $|A_n|$ is large and $S_e(b, n)$ and $S_e(b, n + 1)$ are drawn i.i.d from the same distribution. Additionally, we added the infected individual $a$ to the numerator summation. Hence, the approximation is valid as long as no individual infects a significant fraction of the population.

**ORCID**

Jhonatan Tavori [http://orcid.org/0000-0001-7111-8960]

**References**

[1] Anderson, R. M.; May, R. M. *Infectious Diseases of Humans: dynamics and Control*; Oxford University Press: Oxford, 1992.

[2] 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, 846–849. DOI: 10.1126/science.abc6810.

[3] Endo, A.; Abbott, S.; Kucharski, A. J.; Funk, S. Estimating the Overdispersion in COVID-19 Transmission Using Outbreak Sizes outside China. *Wellcome Open Res.* 2020, 5, 67. DOI: 10.12688/wellcomeopenres.15842.3.

[4] Hadar, J.; Russell, W. R. Stochastic Dominance and Diversification. *J. Econ. Theory* 1971, 3, 3, 288–305. DOI: 10.1016/0022-0531(71)90024-X.

[5] He, Z.; Cai, Z.; Wang, X. Modeling propagation dynamics and developing optimized countermeasures for rumor spreading in online social networks. In *IEEE 35th Int. Conf. on Dist. Computing Systems*, 2015, 205–214.

[6] Hethcote, H. W. The Mathematics of Infectious Diseases. *SIAM Rev.* 2000, 42, 599–653. DOI: 10.1137/S0036144500371907.

[7] Jacquet, P.; Mans, B.; Rodolakis, G. Information Propagation Speed in Mobile and Delay Tolerant Networks. *IEEE Trans. Inform. Theory* 2010, 56, 5001–5015. DOI: 10.1109/TIT.2010.2059830.

[8] Khouzani, M. H. R.; Sarkar, S.; Altman, E. Saddle-Point Strategies in Malware Attack. *IEEE J. Select. Areas Commun.* 2011, 30, 31–43. DOI: 10.1109/JSAC.2012.120104.

[9] Khouzani, M. H. R.; Altman, E.; Sarkar, S. Optimal Quarantining of Wireless Malware through Reception Gain Control. *IEEE Trans. Automat. Contr.* 2011, 57, 49–61. DOI: 10.1109/TAC.2011.2150350.

[10] Lloyd, A. L.; May, R. M. How Viruses Spread among Computers and People. *Science* 2001, 292, 1316–1317. DOI: 10.1126/science.1061076.
[11] Lloyd-Smith, J. O.; Schreiber, S. J.; Kopp, P. E.; Getz, W. M. Superspreading and the Effect of Individual Variation on Disease Emergence. *Nature* 2005, 438, 355–359. DOI: 10.1038/nature04153.

[12] Müller, A.; Stoyan, D. *Comparison Methods for Stochastic Models and Risks*; Wiley: New York, 2002.

[13] Newman, M. E. J. The Structure and Function of Complex Networks. *SIAM Rev* 2003, 45, 167–256. DOI: 10.1137/S003614450342480.

[14] Oz, Y.; Rubinstein, I.; Safra, M. Heterogeneity and Superspreading Effect on Herd Immunity. *J. Stat. Mech: Theory Exp.* 2021, 3, 33405.

[15] Pastor-Satorras, R.; Castellano, C.; Van Mieghem, P.; Vespignani, A. Epidemic Processes in Complex Networks. *Rev. Mod. Phys.* 2015, 87, 925–979. DOI: 10.1103/RevModPhys.87.925.

[16] Rock, K.; Brand, S.; Moir, J.; Keeling, M. J. Dynamics of Infectious Diseases. *Rep Prog Phys* 2014, 77, 26602. DOI: 10.1088/0034-4885/77/2/026602.

[17] Tavori, J.; Levy, H. Continual versus Occasional Spreading in Networks: Modeling Spreading Thresholds in Epidemic Processes. *SIGMETRICS Perform. Eval. Rev.* 2022, 49, 9–11. DOI: 10.1145/3512798.3512803.

[18] Wolfstetter, E.; Dulleck, U.; Inderst, R.; Kuhbier, P.; Lands-Berger, M. *Stochastic Dominance: theory and Applications*; Citeseer: New Jersey, 1996.

[19] Zhang, X.; Giovanni Neglia, G.; Kurose, K.; Towsley, D. Performance Modeling of Epidemic Routing. *Comput. Netw.* 2007, 51, 2867–2891. DOI: 10.1016/j.comnet.2006.11.028.