Theoretical and Numerical Study of SIR with variable susceptibility

by

Connor Wagner

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Applied Mathematics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

We begin with a survey of mathematical epidemic modelling from its inception to present day. We present up-to-date research on the field of variable susceptibility SIR (Susceptible-Infected-Removed system of differential equations model), which takes the classic SIR model and adds another dimension: susceptibility. The susceptible population here is grouped into categories according to their likelihood to contract a disease upon exposure and each of these is governed by differential equations that are identical except for each group's susceptibility coefficient, which appears in a product with the infectivity coefficient β. This model is studied numerically, and the resultant course of epidemic behaviour (cumulative infections) follows the pattern of the Gompertz function of the form $f(t) = Me^{-e^{b-at}}$. This function also fits closely with observed historical epidemic data. The implications of changing factors such as the number of groups and average susceptibility are studied extensively. A few other connections to the literature regarding the variable susceptibility and classic models are also explored numerically. Two proofs are provided that the deterministic model can only generate one course of epidemic behaviour based on a fixed set of initial conditions and parameters, one of which does not rely on continuity of the solution functions. An algorithm is described which, in theory, could retrieve these initial conditions from total or even early epidemic behaviour. A short conclusion and discussion of future directions concludes the thesis.

For Amber, the woman who called me her husband for almost 14 years. You ensured that whatever I accomplished with my life's work would not serve as a counterexample to the age-old claim that all great art comes from pain. You were the fire in which my mettle was tested and you helped me grow into the kind of person that could stand against anything. Thanks to you I do not let fear stop me from following my heart, and I will never again entertain the idea of living on my knees to avoid dying on my feet. You forced me to be better at a time when I had every excuse available to not even try. With gratitude I will dedicate my entire life to returning the favour.

Acknowledgements

I would like to first and foremost acknowledge my supervisor, Dr. Xinwei Yu, for his infinite patience while I worked my way through creating this thesis amidst a number of challenges that slowed my progress and also brought it to a complete halt at times during the 2 years I have been working on this research. I would also like to thank him for his generous financial support throughout this process, which was very helpful and needed as I balanced the demands of my work and my family's needs. Most of all, I would like to thank him for the vision and insight which he possessed and developed, which is the only reason such progress has been made on this thesis. Vision not only to imagine this project out of pieces of existing literature and knowledge, but also to adapt the direction of research as discoveries were made along the way. I would also like to thank him for his constant support, understanding, and encouragement. I would like to thank my parents, Jane and Frank, my brother Cory, and everyone else who provided emotional support and encouragement throughout this process. I would like to thank each of my 9 children for bringing me joy and purpose. I would like to also thank my children for being wonderful and understanding enough that I can spend long hours away on a regular basis and still be greeted upon return with more love and respect than I should hope to expect. I would like to acknowledge the countless colleagues, friends, family, professors and other instructors that helped me learn the mathematics required for such a project. I would also like to thank my defence committee for supplying me with excellent thought-provoking questions. Lastly, I would like to acknowledge the efforts of everyone who has devoted their time and energy to the study of epidemic modelling. Those efforts gave me a starting point to make what I feel is a substantial contribution to the field, which I hope in turn generates further advancement to our ability to manage future outbreaks.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Background

For countless centuries, humankind has applied a scientific approach to a complex, unpredictable, and dangerous world in order to gain advantages from understanding how things work. This has evolved over time from primitive tool-making to discovering a method to create fire and eventually the wheel, and so forth. In the last several centuries, every generation of humans has far more tools, technology, and knowledge than the generation beforehand. Until recently, survival and reproduction was a privilege reserved for the smartest, strongest, and luckiest. Although there are limitations to what humans can control, in times of peace and prosperity, we are safer and live much longer than our ancestors from a few hundred years ago.

One area in which we as a human race have achieved very mixed results and oftentimes very mediocre progress is the eradication of infectious diseases. This is a big problem in modern society that we have every interest in addressing if we wish to avoid the destruction and death caused by large-scale epidemics. The bubonic plague killed more than 25 million people in the 1300s, long before inoculation, mathematical models, and, as evidenced by the severity and reach of the plague, without an adequate and feasible response strategy.

The first systematic formalization of infection and disease-induced death records can be traced almost four hundred years ago to John Graunt [22].

It would not be long before mathematics had caught up. In 1760 Daniel Bernoulli, in front of an audience at the French Royal Academy of Sciences in Paris, illustrated a set of ordinary differential equations (hereafter referred to as ODE) related to normal death rates, but which also model epidemic growth between sets of never-infected people (susceptibles), currently infected people (infectives), and recovered or removed individuals, where the former are considered immune and the latter have died (or in extensions of this model, migrated) [19]. The following ODE were introduced by Bernoulli. Here μ is the death rate for all causes except the disease in question, $S(t)$ is the number of people yet to experience infection, and β is the infectivity rate.

$$\dot{S}(t) = -(\beta + \mu(t))S(t) \tag{1.1}$$

And, with disease mortality rate η, we have an ODE governing the total surviving population $n(t)$:

$$\dot{n}(t) = -\eta\beta S(t) - \mu(t)n(t) \tag{1.2}$$

Note that these ODE are quite different and therefore have different solutions than those of current epidemiological models. These ODE have a solution which can be expressed as the proportion of surviving individuals who still have yet to contract the disease. This proportion is denoted $z(t) \equiv \frac{S(t)}{n(t)}$:

$$z(t) = \frac{e^{-\beta t}}{1 - \eta + \eta e^{-\beta t}} \tag{1.3}$$

Protection by means of inoculation can be easily and accurately incorporated into the model. This event is considered both the first mathematical argument in favour of inoculation and one of the earliest if not the earliest instance of the competing risks problem [11], which lends itself naturally to pandemic management decisions. It was in the same time frame that the *Law of Mass Action* was introduced. This law, which originated in chemistry, states that the rate of a given reaction is determined by the product of the proportions of the reactants. This law, when applied to susceptibles and infectives, would later become an integral part of the most prevalent and useful epidemic

models we have in present day.

It was over a century and a half after this that the (now universally accepted) germ theory of disease was formally introduced, and in 1906 W.H. Hamer proposed that infection rates were reliant on the number (and therefore proportion) of susceptible and infective (and therefore of the recovered/removed) groups [14]. Hamer characterized the rate of new infectives (Δy_t) as the product of the number of susceptibles x_t, the number of infectives y_t, and a constant β representing the infectiousness of the disease itself:

$$\Delta y_t = \beta x_t y_t \tag{1.4}$$

A short time later, in 1926, Kermack and McKendrick, [29], introduced a set of three ordinary differential equations that has served as the basis and origin of the Susceptible-Infected-Removed (SIR) model of epidemic growth which, along with its many versions and variants, has been the major focus of mathematical epidemic modelling ever since. This model allows epidemic growth to be predicted, and scientists and officials can predict, with some accuracy, how the epidemic will behave and ideally make more informed decisions related to mitigating strategies, such as those used during the COVID-19 pandemic.

In general, in a similar fashion to epidemic control, attempts at forecasting epidemics often do not yield acceptable results [28, 10]. There are numerous reasons for this, including but not limited to natural limitations on the complexity of tractable mathematical models, and a poor body of evidence on the effectiveness of virtually all of our available response strategies except possibly vaccination, which was not available during and for some time after the worldwide COVID-19 outbreak of March 2020. It can be argued that the forecasting problem must be resolved in order to even understand what control measures should be implemented. It seems evident that this belief is the driving force behind the recently accelerated research movement in mathematical disease modelling. As a starting point, we return to the model of Kermack and McKendrick, after a brief discussion of its background.

Some time before Kermack and McKendrick's general epidemic model, Ver-

hulst (1838) [4] introduced a simple epidemic. In this model, infectives are considered to remain in that state forever, resulting in the following set of equations:

$$\dot{S} = -\beta x I S \tag{1.5}$$

$$\dot{I} = \beta x I S \tag{1.6}$$

Which, since $I + S = N$, can be solved completely by either of eq. (2.2), eq. (2.3) and this solution is the commonly known logistic function:

$$I(t) = \frac{I(0)}{I(0) + S(0)e^{-\beta t}} \tag{1.7}$$

This model yields long term cumulative infectivity as 100 %, which means the model is overly simplistic. However, more complex variations of the logistic model have been successfully applied to historical COVID-19 data [18]. Models that have finite recovery/removal times apply to the vast majority of real-life cases and will be the focus of the rest of this thesis.

1.2 The Classical SIR model

The SIR model, as introduced by Kermack and McKendrick, uses Hamer's idea that the rate of change of new infections are proportional to the product of the number of infectives and the number of susceptibles. The model predicts differing levels of beginning-to-end epidemic behaviour based on this product, a constant representing the infectiveness (contagiousness) of the disease in question, and another constant representative of the amount of time an infected person remains contagious. In the following equations, βx can loosely be considered the probability of an exposed person to catch the disease (in the variable susceptibility model discussed later, this is refined slightly). To quantify it more accurately would require more intimate knowledge or assumptions about population mixing behaviour, which is incorporated into stochastic variants of the SIR model, the earliest of which can also be attributed to McK-

endrick himself [19]. We take γ to be the reciprocal of the average length of time a person is infected. We take $I(t)$ to be the number of infectives at time t and $S(t)$ to be the number of susceptibles at time t. We take R to be the members of the population that have recovered with full and permanent immunity, died, been placed into permanent isolation, or otherwise have no chance of again becoming susceptible or infective. Notice this variable does not factor into the equations for susceptibles and infectives, and thus can be omitted when finding an expression for $I(t)$ and $S(t)$. This model assumes a closed population with no net birth/death or migration, except possibly death or migration of recovered individuals, which are irrelevant to the calculations:

$$\dot{S} = -\beta x I S \tag{1.8}$$

$$\dot{I} = \beta x I S - \gamma I \tag{1.9}$$

$$\dot{R} = \gamma I \tag{1.10}$$

For ease of calculation, it is common practice (which we follow hereafter with rare exceptions) to normalize the above set of ODE such that the total population $I(t) + S(t) + R(t) = N$ is set to 1.

A natural quantity of interest that stems from eqs. (1.8) to (1.10) is the constant $\mathcal{R}_0 := \frac{\beta x}{\gamma}$, known as the *basic reproduction number*, which is the expected number of new infections resulting from an arbitrary infected individual, excluding both subsequent generations of infections and assuming that all contacts are susceptibles[1] . In some cases calculations use the reciprocal of this number, which is widely referred to as the *relative removal rate* ρ. Kermack and McKendrick affirmed the logical assumption that an epidemic will not grow at all if this relative removal rate is greater than the number of initially susceptible people, (Threshold Theorem) [29].

[1] $\mathcal{R}_0$ could be viewed as the initial reproduction number, when the number of infectives is so small compared to the number of susceptibles that it can be neglected. It can then be assumed with little consequence that an infective at the initial stages of the epidemic would have close to 100% of contacts susceptible to the disease

The system of eqs. (1.8) to (1.10) can be solved for $S(t)$:

$$S(t) = S(0)e^{-\beta x \int_0^t I(\tau)d\tau} \tag{1.11}$$

Similar expressions can also be found for $I(t)$ and $R(t)$. $S(t)$ is usually the desired closed-form function, and it is unfortunate that this solution function for $S(t)$ relies on the function $I(t)$ for all times between 0 and t (and furthermore the converse is also true). Indeed, there are a number of papers that prove semi-analytic, approximate, and parameterized solutions for $S(t)$ (and $I(t)$, $R(t)$ as well), and furthermore many of these also report solutions that involve the other populations of interest or their derivatives, or rely on convenient modifications to the classic SIR model [13, 23, 30, 50]. It seems evident that the most desired solutions for $S(t)$, $I(t)$, and $R(t)$ (these being functions depending only on t) are unknown as of the writing of this thesis. That being said, modern numerical integration methods provide a sufficiently accurate and fine discretization of $S(t)$ that is extremely close to the unknown solution of eq. (1.8). It is easy to see that according to the model and directly from eq. (1.9), if the initial proportion of susceptibles is close to 1, then an appropriate ϵ can be chosen such that $\beta x < \gamma - \epsilon$ which implies $\frac{d}{dt}I(t)\Big|_{t=0} \leq 0$ and the epidemic will not grow. Kermack and McKendrick stated a similar and comparable theorem in 1927 for non-normalized populations where $S(0) + I(0) = N$, with $N \gg 1$ [29]. Kermack and McKendrick here also introduced the *relative removal rate* $\rho := \frac{1}{\mathcal{R}_0}$ and stated that as this number approaches the population size N, total epidemic growth from $t = 0$ to $t = \infty$ is very small, $(\mathcal{O}(N - \rho))$, as would be expected for $\mathcal{R}_0 \ll 1$. This is a logical assumption for which the evidence in the literature is ubiquitous.

The SIR model is a vast improvement upon the overly simplistic and inaccurate exponential growth model, which makes the assumption that the rate of change of new infections is directly proportional to current infections (see eq. (1.12)). This results in the predicted number of susceptibles at the conclusion of the epidemic to be zero, which has not been observed in historical data for the vast majority of diseases observed in sufficiently large populations. Behaviour qualitatively close to exponential growth *does* occur at the beginning

of an epidemic, to some extent in observed real-world data [17], and in the solution to this ODE model, when the number of susceptibles is close to the total population size and for a brief time far exceeds the number of infectives. In this case, eq. (1.9) can be approximated by the following equation, which is the ODE governing exponential growth:

$$\dot{I} \approx (\beta x - \gamma) I \tag{1.12}$$

Since $S(t) \approx 1$ for small t and $I(0)$ close to 0.

This approximation quickly loses accuracy and, as would be predicted by the SIR model, the rate of new infections ceases to grow exponentially and eventually peaks (at the first instance when $\beta x S \leq \gamma$) and then declines until there are virtually no new infections. This qualitative congruence between the real-world data and the epidemic behaviour is encouraging for those seeking more accurate forecasting models and suggests that this model incorporates the most consequential features of the epidemic.

With regards to the model, $\mathcal{R}_0$ is the biggest determinant of epidemic growth, as very small $\mathcal{R}_0$ results in the epidemic declining and dying out as soon as it appears, and large $\mathcal{R}_0$ resulting in an arbitrarily large proportion of cumulative infections over the epidemic's life cycle. This model has been applied a posteriori to real world data on numerous occasions and in many studies, $\mathcal{R}_0$ values have been calculated for a wide range of populations that experienced a COVID-19 outbreak. Many of these results were assimilated in at least two systematic review and meta-analyses published in 2022 [2] [20]. The authors of one study found a wide range (0.4, 12.58) in $\mathcal{R}_0$ [20], while the other [2] found a much smaller range (2.32, 3.69) when looking at large data sets grouped by continent. Statistical tests applied to the latter results yield a high probability that the variation is due to chance and that there is no suggestion that there is a significant difference in true levels of $\mathcal{R}_0$ across continents. As we will see later, even when $\mathcal{R}_0 = 2$ the SIR model suggests 85% long-term cumulative infections. The vast majority of $\mathcal{R}_0$ values in [20] and all of the $\mathcal{R}_0$ values in [2] result in a very high proportion of people experiencing infection over the course of the pandemic. This does happen on occasion in

small communities and to a lesser extent, larger heterogenous populations. By the authors' own admission, higher $\mathcal{R}_0$ values tend to be inflated when looking at preliminary data. Many of the references in this paper do involve preliminary data. The limitations of the SIR model may be illustrated here by the fact that in the early days of the pandemic, very few people have recovered, especially before $\frac{1}{\gamma}$ days have passed. This strengthens the closeness of early epidemic behaviour to exponential growth, as well as exaggerating $\mathcal{R}_0$ values when applied to data. That being said, [2] does not provide evidence against differences between infectivity/susceptibility within populations. It does, however, suggest that average susceptibility is relatively similar when averaging data sets within continents, then comparing those averages. This could be true for a variety of reasons. In regards to [20], the fact that this quantity represents the rate of secondary infections, and that the data in question all applies to the same disease and to groups of humans frequently assumed to have homogenous susceptibility, suggests that expanding the model to incorporate additional dimensions of the population may provide further accuracy in some scenarios. It should be noted that [20] used many small data sets which could have inflated the range of observed $\mathcal{R}_0$ values beyond that caused by other confounding factors.

1.3 Variants of the Classical Deterministic SIR Model

For various reasons, mathematicians have explored different variants of the classical deterministic SIR model. While it is often tempting and sometimes acceptable to ignore additional dimensions such as birth and death rates, there is often an argument for incorporating these and other processes into the model in the hopes that it improves accuracy. For certain categories of outbreaks, additional dimensions must be considered due to the complex and nuanced nature of infectious diseases. We explore many of these here.

We begin with birth and death rates. If we incorporate a constant death rate μ and birth rate Λ we can easily revise eqs. (1.8) to (1.10) to the following

set of ODE:

$$\dot{S} = \Lambda - \mu S - \beta x I S \tag{1.13}$$

$$\dot{I} = \beta x I S - \gamma I - \mu I \tag{1.14}$$

$$\dot{R} = \gamma I - \mu R \tag{1.15}$$

The results of this model are often characterized by the total number of susceptibles remaining after the end of the epidemic. When the epidemic has completely died out, say, at time t_F, then $I(t_F) = I'(t_F) = 0$, and the above equations simplify to

$$\dot{S} = \Lambda - \mu S \tag{1.16}$$

$$\dot{R} = -\mu R \tag{1.17}$$

and the number of recovered individuals experiences exponential decay while the number of susceptible individuals can, at least for finite time intervals beginning at t_F, follow sub-exponential decay, equilibrium, or growth, depending on the birth-rate/death-rate ratio. The assumption here is that at some point in time there will be no new infections and those that were infected will have recovered. If we further assume that $S(t)$ later reaches a critical point (where $S'(t) = 0$), this yields what is known as "Disease-free equilibrium". The disease-free equilibrium for the above sets of equations is

$$S(t) = \frac{\Lambda}{\mu} \tag{1.18}$$

It should be noted that this terminology is usually reserved for outbreaks with significant reach ($\mathcal{R}_0 \geq 1$).

There is a type of modified SIR model that differentiates between deceased and recovered individuals. This model can be solved semi-analytically due to the simplicity and number of differential equations. This model is known as SIRD (Susceptible-Infectious-Recovered-Deceased). As in the classical SIR model, the differential equations governing susceptible and removed individu-

als involve only those variables and constant terms. It seems clear that the only additional information provided by this model is the discrimination between recovered and deceased individuals, which are classified as $D(t)$:

$$\dot{S} = \Lambda - \mu S - \beta x I S \tag{1.19}$$

$$\dot{I} = \beta x I S - \gamma I - \mu I \tag{1.20}$$

$$\dot{R} = \gamma I \tag{1.21}$$

$$\dot{D} = \mu I \tag{1.22}$$

This model can itself be slightly modified to incorporate vaccination rates $v(t)$ and cumulative vaccinations (proportion of vaccinated individuals, since vaccinations are considered to be permanent) $V(t)$, which in practical terms, should both not be considered to be constant. Consequently, in this model infection and recovery rates are also considered to be a function of time, which results in the following set of differential equations:

$$\dot{S} = -\beta(t) x I S - v(t) S \tag{1.23}$$

$$\dot{I} = \beta(t) x I S - \gamma(t) I \tag{1.24}$$

$$\dot{R} = \gamma(t) I \tag{1.25}$$

$$\dot{V} = v(t) S \tag{1.26}$$

This adds additional complexity to the model and in general it can be analytically intractable, but that issue can be addressed by fixing the following ratios as constant:

$$\frac{\gamma(t)}{\beta(t)} = k \tag{1.27}$$

$$\frac{v(t)}{\beta(t)} = b \tag{1.28}$$

Birth and death processes are omitted from this model in order to avoid additional complexity.

Since immunity is an integral part of epidemic modelling, there are cases where maternally-derived-immunity must be considered and incorporated into the model. For diseases such as measles, babies can be born with maternal antibodies if their mothers have been either vaccinated or have previously become infected and recovered. These babies are often resistant if not immune to the disease for the first few months of their life [32]. The model is similar to the classical SIR model except that maternally immune infants are placed into their own category $M(t)$, and because of this, birth and death rates cannot be ignored. Here δ represents the rate at which this immunity dissipates:

$$\dot{M} = \Lambda - \delta M - \mu M \tag{1.29}$$

021

$$\dot{S} = \delta M - \mu S - \beta x I S \tag{1.30}$$

$$\dot{I} = \beta x I S - \gamma I - \mu I \tag{1.31}$$

$$\dot{R} = \gamma I - \mu R \tag{1.32}$$

Many diseases such as tuberculosis can spread by means of asymptomatic transmission, where a person is infected but has no indication that they are, and those people often spread the disease unknowingly. People who are asymptomatic but are not disease-free and can infect others are classified as *carriers*. There are studies that suggest that this happens with COVID-19 as well [52, 21]. In the corresponding model, individuals can move back and forth from the carrier state to the infectious state at any time, while the state of

recovery is considered permanent.

Many models do not incorporate things such as carrier state, but one compartmental model that has received significant amounts of research attention is the SEIR (Susceptible-Exposed-Infectious-Recovered) model. In this model there is a latency period where exposed individuals carry the disease but are not yet able to transmit infection to other people. Note that this is very different and in some ways opposite to models that incorporate asymptomatic transmission.

This model introduces exposed individuals $E(t)$ and includes the variable α which is the reciprocal of the latency period. For the sake of simplicity the birth rate is usually considered equal to the death rate, but in these cases the death rate (not the birth rate) still appears explicitly in the set of ODE:

$$\dot{S} = \mu(1 - S) - \beta x I S \tag{1.33}$$

$$\dot{E} = \beta x I S - (\mu + \alpha)E \tag{1.34}$$

$$\dot{I} = \alpha E - (\gamma + \mu)E \tag{1.35}$$

$$\dot{R} = \gamma I - \mu R \tag{1.36}$$

There are also models explored in the literature where immunity beyond the infectious period is either nonexistent or at least temporary. These are known (respectively) as the SIS (susceptible-infective-susceptible), and the SIRS (susceptible-infectious-recovered-susceptible) models. These models can also incorporate other considerations such as those mentioned above. For example, the MSEIRS (Maternally immune-susceptible-exposed-infectious-recovered-susceptible) model allows for maternally derived immunity, a latency period between infection and infectiousness, and reliably finite periods of immunity following infection.

Another major consideration that can be incorporated into this model is contact rates. The logical assumption that contact rates between individuals

have a substantial impact on general epidemic behaviour is the basis for including contact rates in mathematical models. This belief is also the primary motivating factor for certain epidemic control measures such as quarantine, lockdown, travel restrictions, and modifications to occupancy limits and businesses' hours of operation. Contact rates are affected by a wide variety of factors including cold weather and school calendars.

Models can incorporate contact rates and still remain relatively simple. Some of these models incorporate a seasonal contact rate where the force of infection is a periodic function that depends entirely on the calendar date and time. Other factors that affect contact rates are ignored for simplicity. For example, consider the following set of ODE with periodic force of infection $\beta(t)$ and constant death rate μ:

$$\dot{S} = \mu(1 - S) - \beta(t)xIS \tag{1.37}$$

$$\dot{I} = \beta(t)xIS - (\gamma + \mu)I \tag{1.38}$$

Where the proportion of recovered individuals $R(t)$ is still considered, but is not needed in order to solve the above set of equations. Models where contacts between individuals are assumed to completely govern epidemic behaviour are best framed as graph theory problems. Epidemic models based on graphs are closely related to the well-researched cops and robbers game [35] and to a greater extent the firefighter problem (see [51]). Models such as these can also incorporate the stochastic nature of transmission. There are a few examples of this in the literature [37]. However, doing this adds complexity and generally results in models that can only be analyzed by computer simulations.

We conclude this section with a discussion of another variant of the classic SIR model, known as the differential susceptibility (DS) model. We first describe a model that serves as an intermediate step.

In this modified SIR model, γ is classified as the rate of recovery incidence, and doubles as the infectivity rate as follows:

$$\gamma := \alpha\beta cI \tag{1.39}$$

Where α represents susceptibility, β represents infectivity, and $c = c(N)$ is a constant describing the number (or proportion of the population) of contacts a person has, and this depends only on the total population size. For a normalized system this amounts to an arbitrary constant describing mixing behaviour. Additionally, δ and ξ are the disease-induced mortality rates for infected and recovered individuals, respectively. In the classic SIR model, infection rates are constant when $\beta xS = \gamma$.

We now compare this with the previously mentioned SIR model incorporating constant birth and death rates:

$$\dot{S} = \Lambda - \mu S - \beta x I S \tag{1.40}$$

$$\dot{I} = \beta x I S - \gamma I - \mu I \tag{1.41}$$

$$\dot{R} = \gamma I - \mu R \tag{1.42}$$

with this model,

$$\dot{S} = \Lambda - \mu S - \gamma S \tag{1.43}$$

$$\dot{I} = \gamma S - (\mu + \gamma + \delta) I \tag{1.44}$$

$$\dot{R} = \gamma I - (\mu + \xi) R \tag{1.45}$$

We see that one difference here is the discrimination and inclusion of both disease-induced and normal (all-other-causes) death rates between and within the groups I, R, in addition to the incorporation of contact rates.

The differential susceptibility model explored in [25] takes the above system of ODE and divides the groups of susceptibles $S(t)$ into groups $S_i(t)$ according to their individual susceptibilities. Groups would, instead of having uniform susceptibility α, would have individual susceptibilities α_i, generating a unique value γ_i for each group which generates this modified system of ODE which

serves as the basis for the DS susceptibility model . The rationale here for different susceptibilities could be attributed to differences in contact rates, immune function, or many other factors, which will be discussed in detail later.

$$\dot{S}_i = \Lambda - \mu S_i - \gamma_i S_i \tag{1.46}$$

$$\dot{I} = \sum_{i=1}^{k} \gamma_i S_i - (\mu + \gamma + \delta) I \tag{1.47}$$

$$\dot{R} = \gamma I - (\mu + \xi) R \tag{1.48}$$

The solutions to this set of ODE can be calculated using matrix algebra, and should be computationally tractable for relatively modest values of k, for much the same reasons that the numerical solutions to the models discussed later are also tractable. Another paper by the same authors further refined this model to discriminate between the mixing behaviour of newly-infected versus, say, recently infected individuals [26].

There is one more variant to the SIR model that has not been explored in this section. This variant also takes into account heterogeneity of susceptibility but is still very different in terms of its solution functions due to nuanced differences between the models. This variant and the original progress on the subject made by the author is the focal point of the remainder of this thesis. We first address the rich body of scientific knowledge of the general and simple stochastic epidemics, which preserve many features of the SIR model.

1.4 Stochastic Modelling

The SIR model introduced in the previous section is also referred to as a deterministic model, as the model generates a single possible function for $S(t)$ based on the initial conditions and the values of βx and γ. That is, the solution to the model does not rely on random chance or any stochastic variables.

Some attention has been devoted to the study of epidemic models in which

susceptibles are given probabilities that lie within the interval (0,1) of making adequate contact with an infective and can thus (at least temporarily) avoid infection. These models are referred to as *stochastic epidemic models* and much research into these models continues to present day [44] [48] [27]. In these models, t is either discretized into reasonably large time steps, or the time steps are considered arbitrarily small (continuous stochastic epidemic) and the general infectivity of the disease can be absorbed into these probabilities easily. The solutions to these problems are sample paths of epidemic growth and are most readily applied to the small populations for which the model is the most useful.

In 1928, Lowel Reed and Wade Hampton Frost gave a series of unpublished lectures on the stochastic model, where these probabilities are the only components of the model, which results in a *chain binomial model*[2]. The mathematics from these lectures was recorded in a paper published in 1952 [1]. Individuals can be classified (both susceptible and infectious) as points on a graph of cardinality $N = I + R + S$, with an edge between two nodes occurring independently with probability $p \in (0, 1)$ where $p = 1 - e^{-\mathcal{R}_0}$ [15]. Any such graph is a version of the Erdos-Renyi graph, as its defining feature is the equal probability of an edge between all possible pairs of nodes. Since contact between susceptibles and infectives is randomized at each step, if one is only interested in cumulative infections and not the epidemic process, it turns out that this problem can be reframed as an equivalent model that makes the simplifying assumption, without further cost to the accuracy of the final epidemic size, that all infections take place at the same time, at the end of the epidemic. This quality raises natural questions about the efficacy of the model, as this simplifying assumption appears to be in defiance of many important known qualities of epidemic spread. It should also be noted that this equivalency does not spare the model from its large computational cost, as its solution is defined recursively. It is only the contacts between individuals that follows the structure of the Erdos-Renyi graph, not necessarily the infectivity at any

[2]These models generate multiple scenarios at each discrete time step based on whether or not each susceptible becomes infected. This process is continued at each subsequent time step and depends on the full history of the epidemic.

time. This is due to the fact that edges represent contact between individuals, where the nodes represent people. It is, however, easy to see that this characterization gives a small amount of credence to the idea that in a population with some stochastic element to contacts, there is no sufficiently large $\mathcal{R}_0$ that guarantees a significant growth in infections, as any finite Erdos-Renyi graph has a strictly positive probability of having no edges at all!

Nevertheless, over the years, the simple, small-scale model of Reed and Frost has given rise to plentiful amounts of research on both the simple and general stochastic epidemic, much of which relies on markov chains. Some defining features of stochastic epidemics include the fact that, due to the probabilistic nature of transmission, there is no sufficiently large $\mathcal{R}_0$ that will guarantee cumulative infections close to 100%. That being said, the distribution of cumulative infections of a given stochastic epidemic is normally distributed with the same mean as the deterministic epidemic as long as the population and the relative removal rate are sufficiently large [7]. Though there are limitations to these models, including a reduction in model usefulness for extended epidemic cycles [19], research on the topic has been accelerating since its inception, possibly fuelled by increasingly powerful computer simulation technology. The interested reader may refer to the 2009 survey on stochastic epidemic models by Britton [15]. We discuss the main results of this survey here.

In the survey, the author mentions the widely accepted and logical view that the probabilistic nature of contact between infectives and susceptibles can make a major difference for small outbreaks and/or small communities, due to the fact that with a small number of initial infectives, chance avoidance of contact between those infected and those currently uninfected can make the difference between epidemic growth and epidemic termination. Small communities are more likely to have smaller outbreaks, at least in terms of total number of initial infectives (as opposed to the initial proportion of infectives), but those small and often close-knit communities would possibly experience fast and widespread epidemic growth if a sufficiently large proportion of a small community becomes infected, and if so, this would generally happen regardless of the effects of migration. Indeed, there have been numerous reports of certain communities that are hit harder by COVID-19 [42, 36], and many

of these are small, close-knit communities [46, 47].

Many of these micro-level outbreaks of COVID-19 within closed or mostly closed communities experience epidemic behaviour that does not parallel the macroscopic infection trends. This may serve as an indication that the infection rate βx may not be the same for different groups of people, whether or not these differences are due to genetic differences (presumably due to natural variations in immune system function and efficacy), or to routine chance differences in contact rates, which could be less likely in closed communities that have little time to impose mitigation measures before the epidemic grows to a point where they would be ineffective.

Heterogeneity of immunity can be incorporated into the SIR model (or any of its variations) by setting i subsets of susceptibles with their own rate of susceptibility βx_i. The stochastic nature of contact rates cannot be incorporated into this parameter in this way, and if model simulations are to produce multiple scenarios based on the same set of initial conditions and assumptions, then they must contain a stochastic element.

This stochastic element is useful due to the fact that it can provide valuable (but oftentimes with unknown confidence) error estimation for deterministic models, presumably if the variability in contact rates estimated by the model is somewhat accurate and the model is somewhat suitable to describe the real-life phenomenon. Epidemic behaviour and duration that cannot be explained by the best available deterministic models could be justified by taking into account the probabilistic nature of transmission, whether or not this is attributed to contact rates or anything else that can impact transmission, such as the probabilistic nature of airborne disease transmission itself, and environmental or social factors that affect immunity and behaviour.

As mentioned, the stochastic model seems most necessary for small populations. The *stochastic general epidemic* mentioned in [15] applies the stochastic element to the length of the infectious period alone. This model describes a Markovian process in which each individual's infectious period is i.i.d and a member of an exponential distribution. In this model, the final size of the epidemic can be defined by a recursive formula, but there is no closed-form analytic solution for the total beginning-to-end behaviour of the general stochas-

tic epidemic, even if we are to relax the requirement that the distribution be exponential.

In [6], which we will revisit later in the thesis, the author provides a recursive formula for the distribution of the total epidemic impact, since a defining quality of the stochastic epidemic is that the same initial conditions generally produce different outcomes due to chance variation, which is integral to the model. Where the impact can be classified as the proportion or number of people who eventually experience infection, or equivalently, (which is the quantity used in each step of calculations), the proportion or number of people who successfully escape infection. As this method (in essence) counts the number of susceptibles (infectives) at each time step, it can be framed as a counting problem with a stochastic element. As seen in many graph theory or counting problems with a stochastic element, these stochastic formulas can easily become computationally intractable. For problems such as this, approximations with or without upper and lower bounds on the quantity in question are seen as an acceptable substitute for closed form solutions or the usual proxy of relatively precise computer-generated solutions. This is seen numerous times in [51], and even these results in the literature from this particular combinatorics problem are too plentiful to list here. In addition, there are further limitations beyond the analytic and computational intractability of viewing epidemic behaviour as a branching process. There are other approximations which rely on other methods. In [34], the authors use fluid and diffusion approximations to estimate results for the usual (general stochastic) Markov chain branching model of epidemic spread.

It is worth mentioning that these recursive formulas can be computed for very small-scale epidemics [19], which are the very epidemics in which the stochastic models are considered the most valuable. That being said the recursive nature of the stochastic processes and the implicit assumption that probabilities of contact, which is a proxy for transmission and usually assumed to be i.i.d, no longer provides a realistic description of epidemic behaviour once a significant proportion of pairwise contacts are between infectives is reached. If epidemic growth is viewed through the lens of contact between susceptibles and infectives, contacts between two infectives do not have an impact on epi-

demic growth, and this becomes significant in the calculations as the number of infectives approaches $\sqrt{N}$. This value is seen countless times in similar problems in the literature [51]. It should be noted here that the stochastic general epidemic relies on actual numbers of susceptibles and infectives, rather than proportions, as the members of the population are treated as individual agents whose infectivity status is used in calculations involving the branching process, and this number N therefore cannot be normalized (which can be done rather easily in the deterministic epidemic) if we are to preserve the essential features of this model.

There are claims in the literature that the true value of $\mathcal{R}_0$ is more accurately estimated using a stochastic model rather than a deterministic model [39]. These claims are supported and echoed in other research articles as well [16]. It is worth mentioning again that $\mathcal{R}_0$ is considered to be the *intial* reproduction number. The fact that, loosely speaking, stochastic models lose their accuracy once the epidemic reaches a certain size, gives cause for doubt that these models provide better estimates for what amounts to average initial transmissibility rates. That being said, it seems possibly unreasonable to ignore the stochastic nature of transmission, which can be interpreted as justification for the unpredictable behaviour of individual epidemics when compared with the predictions stemming from the best available deterministic models. This phenomena brings into question whether or not the complexity of the existing deterministic and stochastic epidemic models is sufficient to accurately predict or explain epidemic behaviour. The countless examples that exist of inadequate epidemic forecasting suggests this may be true. This author would like to reserve judgment on such claims for the time being. These stochastic models and similar models can be applied analogously to variants such as the SIS model [40], [33] [12]. This is worth mentioning as the purpose of these variant models is to provide further accuracy for situations not adequately described by the SIR model. Variable susceptibility SIR, which this author holds the be the most useful model when anticipating epidemic behaviour, is the focal point for the remainder of this thesis. The effects of mitigation measures such as lockdown and quarantine can easily be incorporated into both the deterministic and stochastic version of SIR and most of its variants.

The above are just a few examples of the rich body of research into stochastic epidemic models. The most recent developments borne out of the COVID-19 pandemic explore many facets and modifications of the general stochastic epidemic. In [33], the authors present the classic SIR model with the addition of stochastic uncertainty. While the deterministic SIR model produces one possible value for $S(t)$ for each t, this model allows for variation in $S(t)$ for all $t > 0$. This is accomplished by having an additional stochastic variable which results in a range of possible values for $S'(t)$ at each time step $t > 0$. Note that this is much different than the stochastic epidemic models mentioned in [19] that are based around the stochastic nature of transmission involving individual infectives. These simulations can be run numerically and provide another dimension to the uncertain nature of transmission captured by the model. It is not hard to see how the impact of stochastic uncertainty on final epidemic size can be accounted for or even negated by adjusting average transmission probabilities by the requisite amount. The impact of incorporating stochastic transmission to models that place importance on beginning-to-end epidemic behaviour is likely a more complicated question. Exploring the intricacies of contact behaviour and disease transmission lends itself naturally to game theory [24].

Many current studies on COVID-19 data, both before and after the large-scale rollout of the novel MRNA vaccines, focus on many factors such as contact rates, vaccination rates, duration of infectivity, ethnicity proportions within populations, and others [31], [3] [49]. These factors do not exist in isolation and it is not hard to see how even cultural differences should lead to some differences in, say, contact behaviour and vaccination compliance. This further confounds any results that attempt to provide real evidence that differences in any one of these factors in isolation mitigates or worsens epidemic trajectory, though there are still a large number of studies that claim to provide evidence for the efficacy of MRNA vaccination to reduce transmission (and to a lesser extent, severity of the disease and therefore hospitalization and death rates for infected individuals). This is in spite of the fact that the ability of large-scale vaccination to prevent transmission has been called into question since shortly after these programs were implemented. Indeed, as always with conclusions

based on statistical studies that are inherently incapable of providing useful 100% confidence intervals, those conclusions are subject to revision, which has been happening in all facets of the scientific community for generations. Claims by medical officials in the United States and undoubtedly elsewhere were constantly being revised throughout the course of the pandemic over the last several years and continues to this day.

One aspect that suggests that even the most current epidemic models fail to completely capture the relevant physical details, is the fact that trends in COVID-19 and other airborne disease data suggest a strong seasonal (and therefore periodic) influence on epidemic behaviour. These effects can potentially be ignored for short-lived outbreaks or outbreaks in which a small proportion $p \ll 1$ of the population escapes infection. More research is needed into the idea that observed COVID-19 outbreak behaviour may actually be due to several overlapping outbreaks differentiated by mutations in the virus, for which immunity conferred by infection with previous mutations is thought to be partial or nonexistent. This may be the reason that true vaccine efficacy levels were lower than originally predicted. This explanation was suggested in a talk by Shadwick.

It has been established that stochastic epidemic models produce many differing scenarios stemming from identical sets of initial conditions, that their usefulness and tractability is maximized for small communities, and furthermore that models applied a posteriori to evaluate effectiveness of measures such as vaccine programs cannot be separated from all possible confounding factors. Indeed, under certain and common circumstances, epidemic behaviour observed from beginning to end provides little more information about transition probabilities than the final epidemic size does [9]. It is interesting to consider this combined with the fact that the deterministic epidemic's final size is entirely based on susceptibility or susceptibilities, if you ignore the additional dimension of transmissibility, which is common practice. Since any SIR model with homogenous or heterogenous susceptibility can have $\beta x_{(i)}$ modified to produce any final epidemic size, it could be argued that the deterministic model produces valuable information at a higher benefit-cost ratio. For these reasons, deterministic models are the focal point of the remainder of

this thesis.

Deterministic models such as the SIR model and (to a possibly greater extent) the variable susceptibility SIR model explored by the author in part II both produce solutions that can often be closely fitted to a wealth of previous disease data. Endeavours to find closed-form functions that can be fitted to real data but that are not required to be solutions of any ODE model have produced promising results [17]. Functions known as "Gompertz functions" seem to fit these demands very well, and are introduced in the next chapter. COVID-19 behaviour in relatively closed populations from outbreak to endemicity (which tends to run its course within a year) often follows this trend as well. Despite this, long-term COVID-19 data on large, open populations that spans multiple years may indicate a rare exception to this rule [17]. More investigation is needed into the idea that overlapping outbreaks involving mutations can explain this departure from usual disease trends. If there exists a model that can, with reasonable accuracy, reproduce historical data and also predict future epidemic behaviour, we would be much better equipped as a civilization to eventually understand the effects of the seemingly countless list of natural and artificially manipulated factors which affect or are presumed to affect epidemic growth. Natural factors such as seasonal progression and observed ethnically-based differences in COVID-19 infection trends suggest that immunity, which is affected by such things, may be a factor. These beliefs are the motivation for the remainder of this thesis.

Chapter 2

Numerical Study

2.1 The Gompertz Function and Epidemic Data

Much research continues to this day in an attempt to find an adequate epidemic forecasting model that can properly inform vital decisions regarding pandemic management. Many preliminary models from the early days of COVID-19 resulted in doomsday scenarios that never came to fruition. Analyzing the cost-benefit paradigm of individual decisions, as well as large-scale pandemic management decisions made at the highest level of government and healthcare, is extremely difficult to do when the cost cannot be properly estimated. Taking a careful look at relevant preexisting data is important when creating or refining a mathematical model that is intended to predict analogous future phenomena. In recent times, researchers into epidemic behaviour have discovered that a function, itself discovered approximately 200 years ago, has proven extremely useful in making sense of epidemic behaviour. The Gompertz function:

$$G(t) = Me^{-e^{b-at}}, \tag{2.1}$$

Introduced by its namesake, British actuary Benjamin Gompertz, was presented as a function that can be used to determine (with reasonable accuracy) the proportion (or number within a closed population) of people at given age t who are still alive. Figure 2.1 provides a visual.

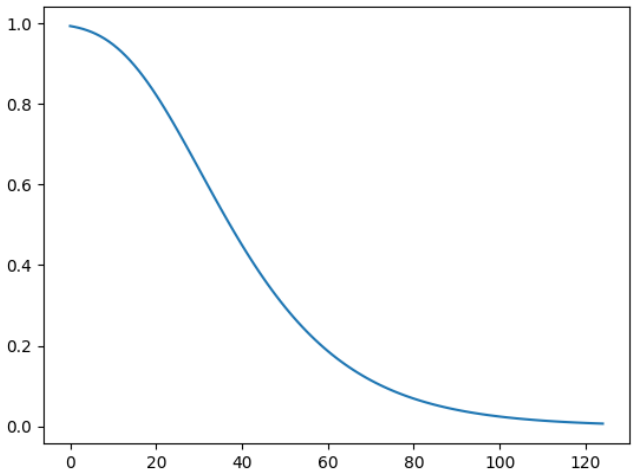


Figure 2.1: $g(t) = 1 - Me^{-e^{b-at}}$ with $M = 1$, $b = 1.61$, $a = 0.0531$. This slightly modified Gompertz function gives a simulated estimate of the surviving proportion of an arbitrary cohort by member age.

It is easy to see how the behaviour of this function seems commensurate with known mortality trends. However, perhaps less obvious is the relationship between this function and the proportion of susceptible (never-infected) individuals within a population experiencing an outbreak. In [17], the Gompertz function is introduced as a *phenomenological* model whereby the function is applied to real-world data [3]. The authors found that the error between the data and the Gompertz curve of best fit[4] was acceptably and in some cases exceptionally low. This implies that the relationship between the function and observed epidemic behaviour can be extracted from historical data and used in an effort to make projections about future epidemic behaviour, as stated by the authors in the paper's introduction. This phenomenological approach makes no claims about the reasons behind the closeness-of-fit, and instead

[3] In the realm of epidemiology, test data will always have to serve as a proxy for true case numbers, the suitability of which depends on numerous factors and can never truly be known. Things such as hospital admissions are accurately documented and the data referenced herein largely applies to metrics such as this as well.

[4] Common statistical methods such as least-squares regression can find the **best possible** fit with no possibility of non-optimality between arbitrary data and a given function such as the Gompertz function.

emphasizes the utility of a function that can be applied to a vast and diverse collection of epidemic data. Indeed, the underlying reasons that the function fits so well for epidemic behaviour remain largely a mystery, and little progress has been made by this author (or elsewhere to his knowledge) on uncovering those reasons.

In [17], the authors apply Gompertz function fits to a variety of diseases including cholera, spanish flu, influenza, and COVID-19, and also to multiple data sets spanning countries and eras, within each disease category. Moreover, fits applied to proxies for infection, such as ICU admissions and deaths also yield encouraging results, which would be expected from infection data if one assumes that death and severe illness rates for newly infected people remain relatively constant. The data for historical diseases was taken over months and years, and close fits with minimal error were produced. The authors claim good "short-to-medium" term predictive power for both historical diseases and COVID-19. The authors stopped short of claiming long-term predictive power due to the fact that the best-fit Gompertz function $g(t)$ for the early days of the studied outbreaks in general has different parameter values than the best-fit function $\tilde{g}(t)$ applied to the same data extended well beyond the peak of the disease.

In addition, the authors discuss an extension of a disease's initial reproduction number $\mathcal{R}_0$ to an effective reproduction number that depends on the time elapsed t since the initial outbreak. This is denoted here as $\mathcal{R}_t$. This parameter again estimates the number of secondary infections produced by one new infection, but in this case it is time specific as this quantity will decrease as the proportion of susceptibles declines as an epidemic progresses ($\mathcal{R}_t$ is a non-increasing function of t). This is a much more refined parameter for describing current and time-specific epidemic behaviour, so much so that the authors found in empirical data, that the time t where the rate of new infections begins decreasing corresponds closely to the time t where $\mathcal{R}_t - 1$ becomes negative.

Many diseases throughout history have essentially disappeared or been eradicated. Measles and smallpox, as well as the bubonic plague are not a cause for concern anymore. After the initial shock of the 2020 advent of

a severe and serious worldwide pandemic had subsided, it is assumed that many people throughout the world held out hope that COVID-19 would be nothing more than a memory in the near future. Four years after the virus began spreading globally, such an event does not seem likely to occur anytime soon. However, due presumably to mutations into weaker viruses, hospitals are no longer overwhelmed by COVID-19-related ICU admissions, testing and mitigation measures have been deemed mostly unnecessary in many parts of the world, and people are resuming their normal lives. What we seem to be observing in present day is a classic example of endemicity, or "endemic behaviour", where we are still getting significant and oftentimes very constant rates of new infection, but without the explosion of infection that is usually seen at the onset of a new outbreak. This effectively corresponds to $\mathcal{R}_t \approx 1$ for large time spans, which is, unfortunately, incompatible with the limiting behaviour of Gompertz functions. In order to maintain the efficacy of the model, the authors allow for piecewise functions that alternate between Gompertz function growth and linear growth. They showed that this was an effective refinement for long-term COVID-19 data, as well as multi-year data for seasonally-affected diseases such as Spanish flu. This alternating pattern has been shown to go through multiple cycles in the course of a few years for a selection from a group of seasonal diseases, of which COVID-19 is a member. A visual for such a piecewise function is shown in Figure 2.2. Gompertz and piecewise Gompertz functions applied to historical data are illustrated in numerous plots in [17], and the invested reader is highly encouraged to view the results and discussion of that paper for more information. In the next section we will see that both the classic and variable susceptibility SIR model have computer generated solutions that are very close (in terms of L_p) to Gompertz functions, which by far are the only reasonable choice when it comes to modelling typical large-scale epidemic behaviour, without resorting to piecewise functions, splines, or similar methods. That being said these models are not without their limitations, as, with regards to historical data, the best fitting Gompertz function for, say, time until 20% of max infections is not necessarily close (in terms of the values of the three constants in the function) to the best fitting Gompertz function for the epidemic's lifetime [17].

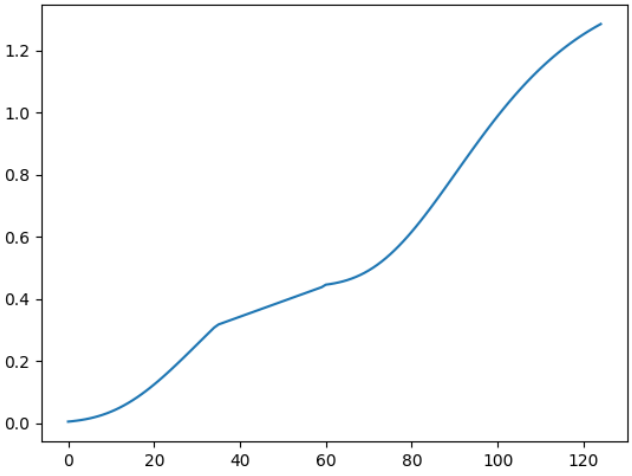


Figure 2.2: Piecewise alternating Gompertz and linear function. The function is linear for $35 \leq t < 60$.

2.2 SIR

Consider again the set of ODE governing the SIR model, which together with its variants dominates the study of epidemic forecasting to this day:

$$\dot{S} = -\beta x I S \tag{2.2}$$

$$\dot{I} = \beta x I S - \gamma I \tag{2.3}$$

We omit $R(t)$ since it is irrelevant to the calculations. Since β and x always occur as a product, the solution to these ODE is based solely on $I(0)$ (since $R(0) = 0$ and $S(0) = 1 - I(0)$), βx, and γ.

Possibly due to the reasons mentioned in Chapter 1, the above set of ODE still has no known fully-analytic, nonparametric solutions depending only on t. Numerical solutions to this set of equations can easily be generated and are found in [45], [39], and elsewhere.

In this thesis, the author utilizes Python's Julia package "DifferentialEquations.jl" [41] to solve this and modified systems of ODE corresponding to the SIR model using a step method. When fitting our results to Gompertz functions, we used Python's SciPy curve_fit package. The original model is also the easiest to simulate. We first run the simulations and discuss the impact of varying the initial conditions and $\mathcal{R}_0$ on the computer-generated results. A short Python program generates plots of $S(t)$ (and $I(t)$ if desired). We set $\beta = 0.325$, $\gamma = 0.13$, consistent with the literature, and vary susceptibility x randomly between 0.5 and 2.5 yielding $1.25 \leq \mathcal{R}_0 \leq 6.25$. We also set $I(0) = 0.05$. A similarly short Python program invokes a different method from its library and finds the curve of best fit. It is trivial to write a program that measure the L_1 error, or any commonly used measure of error.

Two figures appear below. Figures 2.4 and 2.5 show the numerical solution in blue and the Gompertz function of best fit in red. The two curves are indistinguishable in Figure 2.3.

These figures illustrate the fact that the model predicts widespread infection affecting a large majority of the population with an $\mathcal{R}_0$ value of 2.0. The

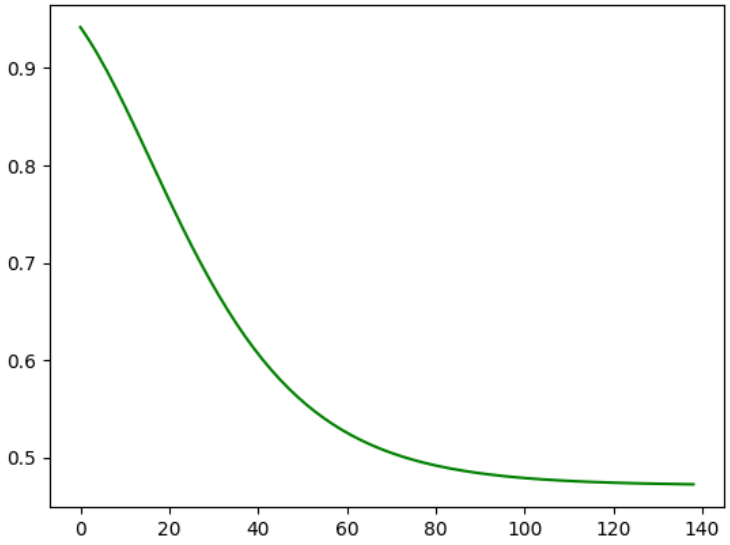


Figure 2.3: Green: $x = 0.531$, $\mathcal{R}_0 = 1.3$, L_2-error: 0.0148, $S_\infty \approx 0.45$

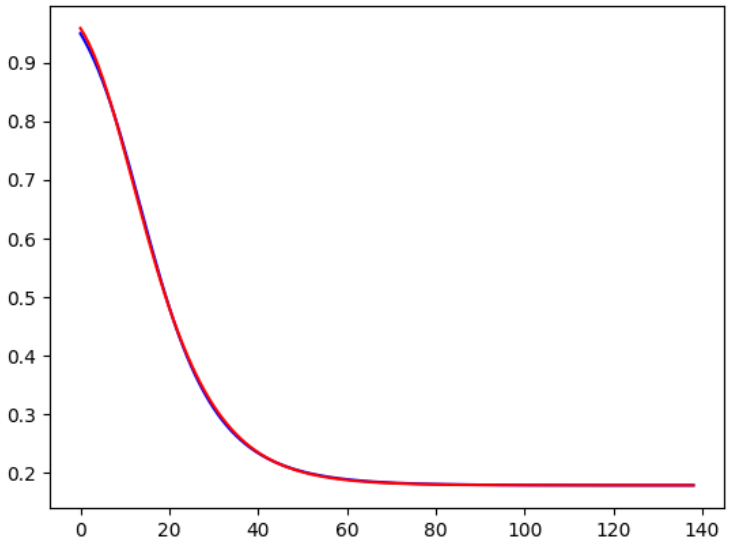


Figure 2.4: Red/Blue: $x = 0.814$, $\mathcal{R}_0 = 2.0$, L_2-error: 0.0313, $S_\infty \approx 0.15$

L_2 error is relatively low, and most of the discrepancy between the computer-generated solution and the Gompertz function fit exists for small t.

This is consistent with observations about unique behaviour similar to exponential growth at the very onset of epidemic outbreak. Even when calculating error from $t = 0$ onward, expanding this model to handle multiple susceptibilities results in a lower L_2 error. Aside from the fact that $\mathcal{R}_0$ can be manipulated to produce an arbitrarily large or small proportion of total infections, finding a model that fits as closely as possible with the best known closed-form modeling function seems like an important consideration. On an empirical level, one should proceed with caution with a model that assumes homogeneity of natural immunity, which seems to be reserved for the extreme ends of the infectiousness spectrum based on the best available science for diseases prior to COVID-19. Lastly, the highly referenced "best estimate" of $\mathcal{R}_0 = 2.5$ produces the results of Figure 2.5, namely the often-predicted $(1 - S_\infty) \approx 90\%$ cumulative infection rate, generally not seen for individual outbreaks in large populations. Due to the structure of the Gompertz function $g(t) = Me^{-e^{b-at}}$, the fact that it has no known antiderivative, and the qualities of the system of differential equations governing $S(t), I(t)$ and $R(t)$, there seems to be little hope of proving that these Gompertz functions that fit the numerical solutions so closely are indeed the exact solutions of the system of ODE.

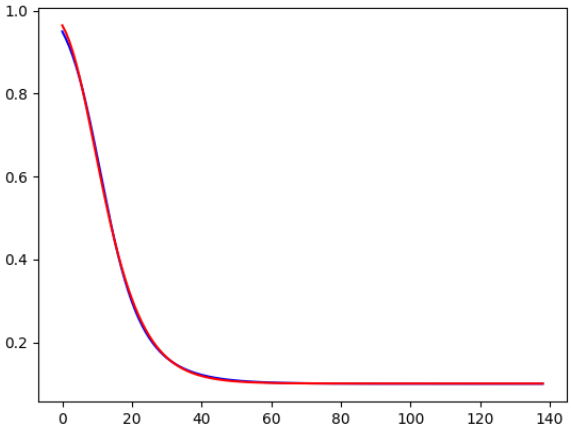


Figure 2.5: Red/Blue: $x = 1$, $\mathcal{R}_0 = 2.5$, L_2-error: 0.0429, $S_\infty \approx 0.1$

2.3 Variable Susceptibility SIR

2.3.1 Background

For many centuries, humankind has been aware of individual variations of immunity. Indeed, the observation that previous exposure and infection to a disease drastically reduces or even negates the chances of re-infection by the same disease can be traced back thousands of years. The idea that immunity can be conferred by vaccination is much older than the first working vaccine, and throughout history, transmission of a variety of illnesses have been thought to be impacted by the qualities of the susceptible hosts themselves. Vitamins, supplements, other health products, lifestyle choices including not only movement and contact behaviour but also eating habits, have all been thought to impact resistance to acquiring diseases and to recovery times. Immunology generally considers two branches of the immune system. The "innate" immune system which provides first-line defence against novel pathogens, and the "adaptive" immune system which is invoked by contact with a pathogen "recognized" by the host's lymphocytes. The adaptive immune system is required for a vaccine to work, while the innate immune system is known to be influenced by factors related to the host's general health [8, 38].

If it is true that there are individual variations in the strength and efficacy of people's innate immune systems, this is a major factor directly related to large-scale epidemic behaviour that cannot be ignored. It can be argued that any useful model must incorporate these discrepancies, as homogenous susceptibility with variable exposure seems unable to account for the wide variety and unpredictability of epidemic behaviours. This argument is only strengthened by the lack of current and accurate epidemic forecasting models. The roots of this expanded model can be traced back to [43], it appears in numerous intermediate stages in a variety of papers, and appears as the model we currently use in [5] and again in [39]. Particularly, in [39], the authors provide numerous insights and parameterized solutions for this more complex, expanded SIR model.

2.3.2 The Model

Expanding the SIR model to incorporate heterogeneity of susceptibility presents a host of advantages, as well as challenges. The basis of the model takes the initial group of susceptibles $S(0)$, and splits them up into k groups represented by proportions $s_i(0)$ such that

$$\sum_{i=1}^{k} s_i(0) = S(0) \tag{2.4}$$

Furthermore,

$$\sum_{i=1}^{k} s_i(t) = S(t) \tag{2.5}$$

Each group s_i has its own distinct susceptibility x_i, and thus we have a modified system of eqs. (2.2) and (2.3) as follows:

$$\dot{s}_i = -\beta x_i s_i I \tag{2.6}$$

$$\dot{I} = \beta \bar{x} I S - \gamma I \tag{2.7}$$

$$S(t) = \sum_{i=1}^{k} s_i(t) \tag{2.8}$$

$$\bar{x}(t) = \frac{(\sum_{i=1}^{k} x_i s_i(t))}{S(t)} \tag{2.9}$$

As before, β can be absorbed into susceptibility, (in this case multiple susceptibilities). This expansion also generates a new function $\bar{x}(t)$ corresponding to average susceptibility. It takes only a moment's reflection to conclude that the relative representation of each susceptibility group changes in the model (and presumably in the real world) as new infections affect more susceptible groups disproportionately. Note that $\bar{x}(t)$ appears explicitly as an independent variable in eq. (2.7).

There is even less hope here of a fully analytic nonparametric solution for

$S(t)$, however, a psuedo-analytic solution for $\bar{x}(t)$ can be attempted. As we will see shortly, Gompertz functions fit even closer for this set of ODE, and if we ignore the error in the numerical approximation to a Gompertz function to represent $S(t)$, we can make the following argument:

$$1 - S(t) = S_\infty e^{-e^{b-at}} \tag{2.10}$$

also, from Equation eqs. (4.3) and (4.4)

$$I(t) = I(0)e^{\int_0^t \beta S x - \gamma \, d\tau} \tag{2.11}$$

$$\dot{S} = -\beta \bar{x} I(0) e^{\int_0^t \beta S x - \gamma \, d\tau} S \tag{2.12}$$

whereas from eq. (4.7) we obtain

$$\dot{S} = -a e^{-e^{b-at}} (\frac{1}{S} - 1) S \tag{2.13}$$

The eqs. (4.9) and (4.10) can be set as equals, and after many steps, which are relegated to the appendices, we find a semi-analytic solution that necessitates numerical computation once again:

$$\bar{x}(t) = \frac{e^{(\gamma - a)t}(\frac{1}{S} - 1)}{C - \beta \frac{e^{-b}}{a} \int_0^t e^{\gamma \tau} \dot{S} \, d\tau} \tag{2.14}$$

And, substituting initial conditions, we have that

$$C = (\frac{1}{S(0)} - 1) \ll 1 \tag{2.15}$$

Here we see a solution for $\bar{x}(t)$ that depends on $S(t)$ and its first derivative. These sorts of solutions abound in the literature for the SIR model and are usually best explored using numerical tools.

As such, we turn to numerical approximations in the hopes that it can provide more information on the usefulness of the model. The Python program used to simulate the classic SIR model can easily be modified to incorporate an arbitrary number of susceptibility groups, and even 10 or 20 susceptibil-

ity groups, which seems adequate to describe even diverse populations with reasonable accuracy, runs in seconds on an average computer. For two susceptibility groups, the system would be set up as such:

$$\dot{s}_1(t) = -\beta x_1 s_1(t) I(t) \tag{2.16}$$

$$\dot{s}_2(t) = -\beta x_2 s_2(t) I(t) \tag{2.17}$$

$$S(t) = s_1(t) + s_2(t) \tag{2.18}$$

$$\bar{x}(t) = \frac{x_1 s_1(t) + x_2 s_2(t)}{S(t)} \tag{2.19}$$

$$\dot{I}(t) = \beta \bar{x}(t) I(t) S(t) - \gamma I(t) \tag{2.20}$$

Where, for example,

$$\beta = 0.325, \gamma = 0.13, x_1 = 0.8, x_2 = 1.6, s_1(0) = 0.45, s_2(0) = 0.5 \tag{2.21}$$

Although the numerical solver demands initial conditions for the first derivatives as well, these can be calculated by the given initial conditions and parameters (as seen above) and inputted into the program.

Python is certainly not the only program that can approximate such a system, and we use a Matlab generated plot to illustrate the solution for 10 susceptibility groups. The scale in this example is non-normalized and represents a population of $N = 810$ with $I(0) = 10$.

In order to gain intuition about the benefits and limitations of the model, it is fitting to begin by making the most modest of transitions from one susceptibility group (classic SIR) to two. We denote the number of susceptibility groups as k.

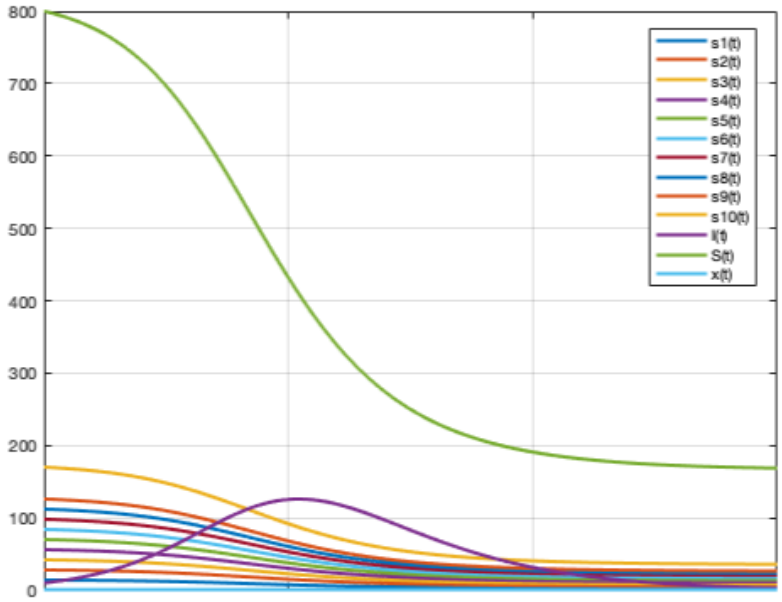
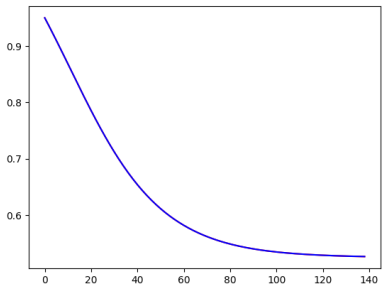


Figure 2.6

2.3.3 $k = 2$

As mentioned before, for the classic SIR model, the solution to these ODE is based entirely on what is essentially one initial condition $I(0)$, and parameters βx, and γ. For two susceptibility groups we consider that now β cannot be joined to x_i anymore but can be neglected based on choices of x_i, and we need four initial conditions $s_1(0), s_2(0), x_1(0), x_2(0)$, and one parameter γ. We first investigate the transition to $k = 2$ by comparing two cases, the first where $x_1 \approx x_2$, but the proportion of the susceptible population in each group are very unequal, and where $s_1(0) \approx s_2(0)$ but x_1 and x_2 differ by a great amount.

We see here that even creating another, very small group, whose groups have very similar susceptibility does not seem to decrease the L_2-error by a significant amount. Neither does increasing the size of that group but keeping the susceptibilities relatively equal. However, setting two equally sized groups with significantly different values for x_i yields a closer approximation with the L_2-error decreasing from the order of 10^{-2} to the order of 10^{-3}. This is



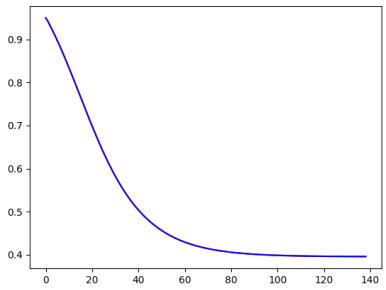


Figure 2.7: Left: $x = [0.500, 0.490]$, $y = [0.940, 0.010]$, $\mathcal{R}_0 \approx 1.25$, L_2-error: $0.0205, S_\infty \approx 0.5$

Figure 2.8: Right: $x = [0.405, 0.850]$, $y = [0.500, 0.451]$, $\mathcal{R}_0 \approx 1.5$, L_2-error: $0.00592, S_\infty \approx 0.4$

in line with the idea that additional and significantly different susceptibility dimensions incorporated into the model brings it slightly closer to emulating the real state of affairs with regards to epidemic behaviour, which unfortunately can never be fully accomplished. It can be argued here that even these results are significant, as it was found in [17] that the Gompertz function provides a great-fitting approximation for numerous types of epidemics, including COVID-19 restricted to the lifespan of single outbreaks. COVID-19 data over the span of years is becoming increasingly readily available and this observed behaviour may be influenced by mutations producing several outbreaks acting concurrently at different phases of their life spans. The existence of distinct mutations is a confounding factor and suggests the need for further expansion. This may be achievable due to the fact that it is not difficult to expand the data-fitting function being set as the sum of multiple Gompertz functions with different constant values.

2.3.4 $k = 5$

We expand now to 5 susceptibility groups, and make observations on the simulations. Intuitively speaking, it seems likely that if we do have heterogenous susceptibility, a model incorporating 5 susceptibility groups should be more realistic than a model with only two. It will be shown however that high

accuracy is maintained but not generally improved upon by increases to the size of k beyond $k = 2$. However, increased accuracy depends on values other than k as well, and depending on the initial spread of susceptibilities and their respective coefficients, we find different levels of error. The following figure shows three examples, the first where $\mathcal{R}_0 \approx 2.5$ and the second where $\mathcal{R}_0 \approx 1.25$. These results seem discouraging as we once again have error on the order of 10^{-2}, however, adjusting for a more medium value of $\mathcal{R}_0 = 1.9$ yields the third figure which possesses a closer approximation. These results indicate no significant improvement in error from $k = 2$ to $k = 5$. Notice that the middle graph with $\mathcal{R}_0 \approx 1.25$ appears to behave in accordance with exponential decay. This may not be surprising as for $\mathcal{R}_0$ close to 1, the epidemic's peak is not long after its inception. Lastly, these figures describe three arbitrary examples with a common distribution of $s_i(0)$ and different distributions of x_i. Varying both initial susceptibility group distributions and susceptibility coefficients randomly allows us to observe both a variety of approximations to Gompertz functions without expanding the scope of the simulations to include modified values of the parameters γ and β. Simulations yielded another example with different distributions of susceptibilities and coefficients from the simulation in the third figure, but that also results in error < 0.01. These conditions were $x_i = [0.688, 0.950, 0.669, 0.352, 0.865]$, $s_i(0) = [0.087, 0.219, 0.169, 0.292, 0.182]$, $\mathcal{R}_0 \approx 2.1$, L_2-error: 0.098, $S_\infty \approx 0.75$. We move on to explore the relationship between the numerical results of our simulations, our qualitative intuition and observations, and the literature, before concluding our discussion of the effects of our manipulations on observed L_2-error.

2.3.5 Numerical and Theoretical Implications of the Model

We again increase the size of k, doubling it this time to 10 distinct susceptibility groups. Before we discuss the impact of this change on numerical accuracy and closeness-of-fit, we take an interlude to discuss the impact of manipulating the variable corresponding to average susceptibility, which for the classic model, when γ and the initial conditions are fixed, corresponds directly to $\mathcal{R}_0$, final

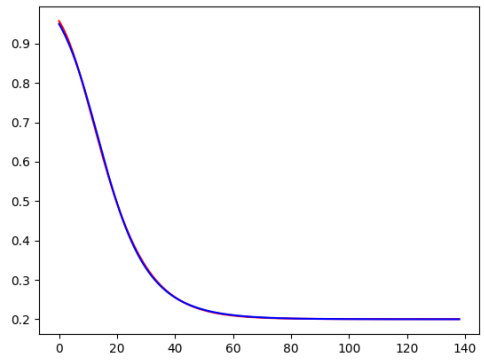


Figure 2.9: $x = [0.915, 0.511, 0.686, 1.031, 0.768]$ $y = [0.186, 0.150, 0.188, 0.266, 0.209]$ $\mathcal{R}_0 \approx 2.5$, L_2-error: 0.026, $S_\infty \approx 0.85$

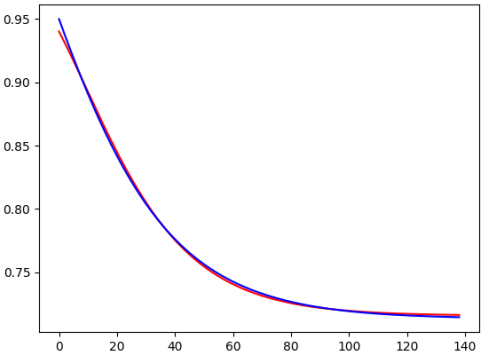


Figure 2.10: $x = [0.515, 0.111, 0.286, 0.631, 0.368]$ $y = [0.186, 0.150, 0.188, 0.266, 0.209]$ $\mathcal{R}_0 \approx 1.25$, L_2-error: 0.023, $S_\infty \approx 0.3$

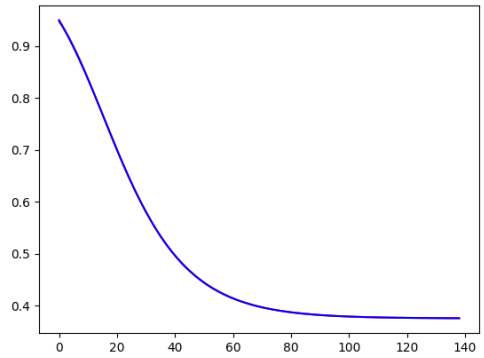


Figure 2.11: $x = [0.686, 0.384, 0.515, 0.773, 0.576]$ $y = [0.186, 0.150, 0.188, 0.266, 0.209]$ $\mathcal{R}_0 \approx 1.9$, L_2-error: 0.0065, $S_\infty \approx 0.65$

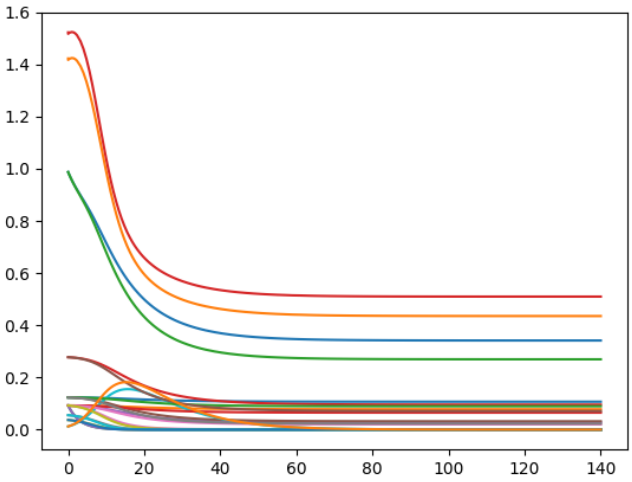


Figure 2.12: Higher susceptibility groups: Red: $x(t)$, Green: $S(t)$, Orange (below): $I(t)$; Lower susceptibility groups: Orange: $x(t)$, Blue: $S(t)$, Teal (below): $I(t)$. All other variables represent $s_i(t)$ for $1 \leq i \leq 20$.

epidemic size, and entire epidemic behaviour over the course of its life span. We also relate these findings to the literature.

We begin with a computer generated plot of two parallel setups for $k = 10$ in the figure. All parameters and initial conditions are the same except that the set of susceptibilities (which we separate by "cohort") is governed by the following equation:

$$x_{i+10} = x_i + 0.1, \{1 \leq i \leq 10\} \tag{2.22}$$

Where the two cohorts are separated accordingly ($1 \leq i \leq 10$ and $11 \leq i \leq 20$).

According to this simulation, though the difference in $\bar{x}(0)$ for the two cohorts is exactly 0.1, this difference is not maintained to the end of the epidemic and the difference for limiting values of t is closer to 0.8. This is an interesting result, and suggests that with higher average susceptibilities (lower mean immune health for population or community), the most susceptible groups are actually more disproportionately affected than the less susceptible groups. This is an interesting result, given:

$$\frac{x}{y} > \frac{x+\delta}{y+\delta}, \{0 < y < x, \delta > 0\} \tag{2.23}$$

After running numerous simulations in this vein, the specifics of which are not recorded here, the following qualitative observations were made:

- Increasing the value of $Min\{x_i\}$ results in a lower final proportion of $S(t)$. Also, values of $S(t)$ approach global minimum faster (in less time) as well and we reach endemicity in a lower amount of time. This corresponds to a Gompertz function with slightly larger values of a and S_∞

- Leaving minimum susceptibility low, even while increasing maximum susceptibility results in a plot where near-maximum infection occurs later, maximum infection is a smaller proportion of the population, and corresponds to smaller values of the constant a.

- The individual functions $s_i(t)$, as shown in the above figure, can also be fit extremely closely to Gompertz functions, which is to be expected as they can be viewed as individual microcosms of homogenous susceptibility within a heterogenous population.

- Finally, letting the initial values of x_k and s_k be an even and large spread about the mean produces the same sort of plot of $S(t)$ as an initial distribution where the vast majority of the population has either low or average susceptibility and x_k is large for large k. This suggests that diversity of susceptibility/immune strength is protective, even if average susceptibility is fixed.

The final observation above suggests that high variability in susceptibility is inherently protective against a high proportion of people eventually experiencing infection, even when controlling for average susceptibility. This is an interesting idea that can be explored numerically by devising an appropriate collection of simulations. The question of whether or not this is a logical necessity based on the mathematics of the classic and variable susceptibility models is an important one, and fortunately this question was explored nearly 30 years

ago and was answered in the affirmative, at least when directly comparing the homogenous and heterogenous models.

In [5], the authors found a closed-form expression for $R(t)$ based on the variable susceptibility SIR model. We introduce it using that author's terminology:

$$X(t) = \sum_{i=1}^{k} s_i(0) e^{-\frac{\alpha_i Z(t)}{\gamma}} \tag{2.24}$$

Where $X(t)$ is the number of susceptible individuals, N is the number of initially susceptible individuals, $Z(t)$ is the number of recovered individuals, γ represents the same quantity as our notation does, and α_i is the infectivity rate, which, according to page 5 of this paper, is equal to $\beta x_i s_i$ This directly translates to:

$$S(t) = \sum_{i=1}^{k} s_i(0) e^{-\frac{\beta}{\gamma} x_i(0) s_i(0) R(t)} \tag{2.25}$$

Compare with our function assumed to closely represent $S(t)$:

$$S(t) = 1 - S_\infty e^{-e^{b-at}} \tag{2.26}$$

Perhaps a challenge in 1985, numerical explorations of these findings are easy and give us insight into the effects of expanding the role of variable susceptibility in the model. We first explore the degree to which this finding agrees with ours.

$R(t)$ also produces very close Gompertz function fits, as expected, and presents as a Gompertz function of the form $R(t) = Me^{-e^{b-at}}$, shown in the figure.

We apply a Gompertz function fit to the semi-numerical approximation to eq. (2.28) which gives Figure 2.13. The DAE solver produces a numerical solution for $R(t)$, and this can be used to calculate a numerical solution to this equation using obtained values of $R(t)$ for each time t. Since Ball's theorem is sound, given the model from which it arises, it was important to check whether or not the solution to $S(t)$ provided by the DAE solver was consistent with

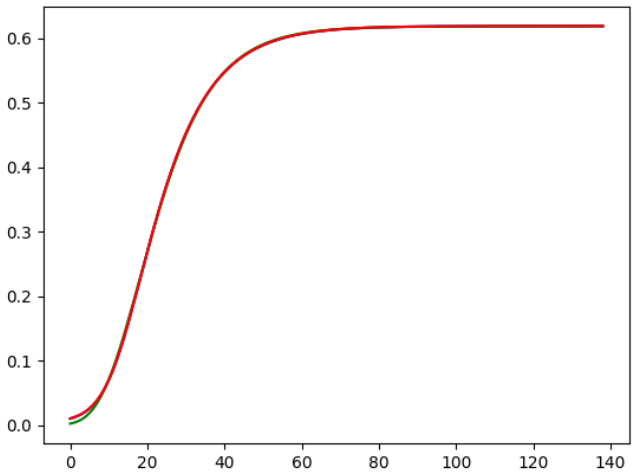


Figure 2.13: R(t), Numerical result in green with Gompertz function curve-of-best-fit in red.

the theorem's representation of $S(t)$. The two functions are plotted in Figure 2.14, and Figure 2.15 illustrates a Gompertz curve-of-best-fit applied to the semi-numerical approximation to eq. (2.28).

It is interesting to note that the L_2-error is higher in the second figure, which could be due to the fact that there are two rounds of simulations required, one to provide the numerical solution for $R(t)$ and another when Ball's $S(t)$ is calculated numerically from those results. We will now, in a similar fashion, explore a related theorem from the same paper. Ball's formula for $S(t)$ for the variable susceptibility model, adjusted to use our notation and using $\hat{S}(t)$ to distinguish the heterogenous susceptibility $S(t)$ from its homogenous counterpart is as follows:

$$\hat{S}(t) = \sum_{i=1}^{k} s_i(0) e^{-\frac{\beta}{\gamma} x_i(0) s_i(0) R(t)} \tag{2.27}$$

and a homogenous susceptibility analogue:

$$S(t) = S(0) e^{-\frac{\beta x}{\gamma} R(t)} \tag{2.28}$$

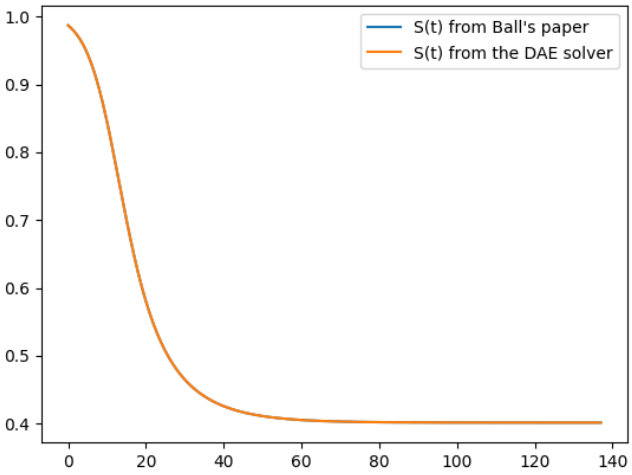


Figure 2.14: Both plots are indistinguishable. L_2-error ≈ 0.00086

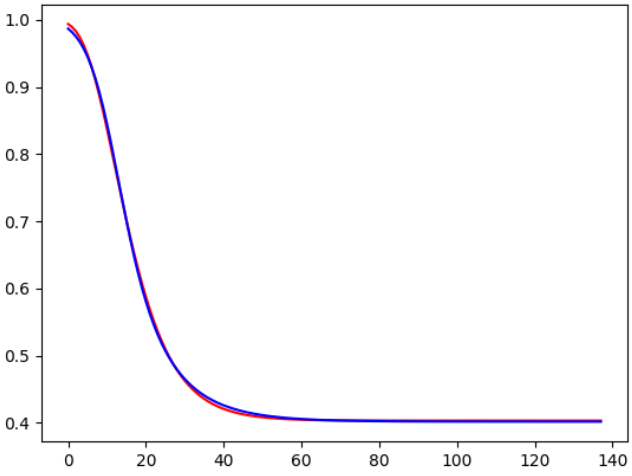


Figure 2.15: Ball's function in blue (same as above) and the Gompertz function approximation in red, L_2-error ≈ 0.037.

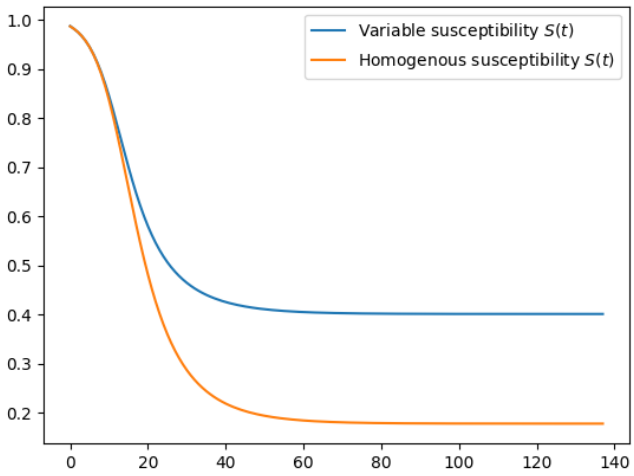


Figure 2.16

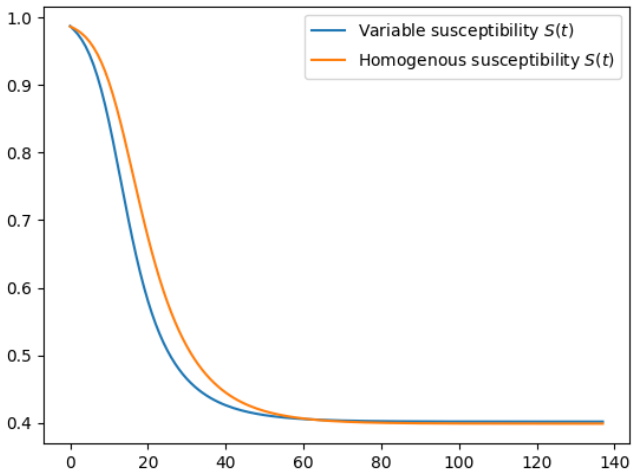
and we also have the following theorem proved by Ball himself:

Theorem 2.1 *For β, γ, $S(0)$, $I(0)$, $R(0)$ fixed, and for $\beta x = \mathcal{R}_0 = \hat{R}_0 = \sum_{i=1}^{k} x_i(0)s_i(0)$, and $\hat{S}(t)$, $S(t)$ defined as above:*

$S(t) \leq \hat{S}(t)$ for all times t.

Comparing them numerically yields the expected results, which are shown in the above figure. It should be noted here that the spread of heterogenous susceptibilities was large.

This presents a perfect opportunity to examine and compare the plots of the classic and variable susceptibility SIR numerically in another fashion. In the previous figure, the marked difference between final epidemic size compared between models is illustrated. To investigate the usefulness of each model, adjusting βx for the classic SIR to closely match the final epidemic size produced by the variable susceptibility model provides us with some insight. As could perhaps be deduced from the first figure, artificially adjusting $\mathcal{R}_0$ for the homogenous susceptibility SIR results in slower early epidemic growth in the classic model, which, with the same initial proportion of infectives as its ana-



logue, corresponds to a more exaggerated departure from the near-exponential growth than that observed in real-life outbreaks.

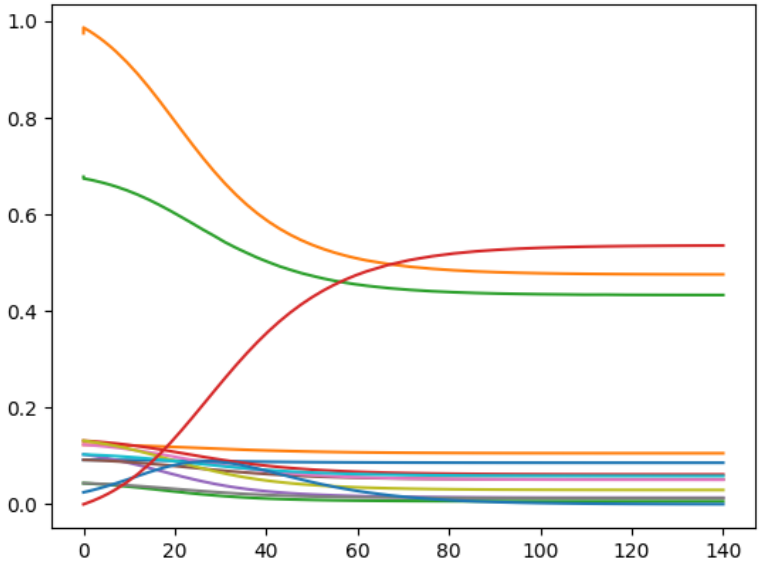
After some reflection, it seems self-evident that these results generalize when comparing small variation in susceptibility versus large variation with the same average value with perhaps also the same skew. Since final size is decreased when increasing from 0 variance, and $\mathcal{R}_0$ can be manipulated to yield any final epidemic size, it seems conclusive that a larger spread with the same average and skew would result in a lower final proportion of total infections. It could also be easy to determine whether there are any conditions required on the skewness of the susceptibility distributions being compared. A comparable claim related to the second observation would state that larger uniform spreads of susceptibility coefficients would result in plots more closely resembling exponential growth in the very early stages of the epidemic than those produced by smaller uniform spreads.

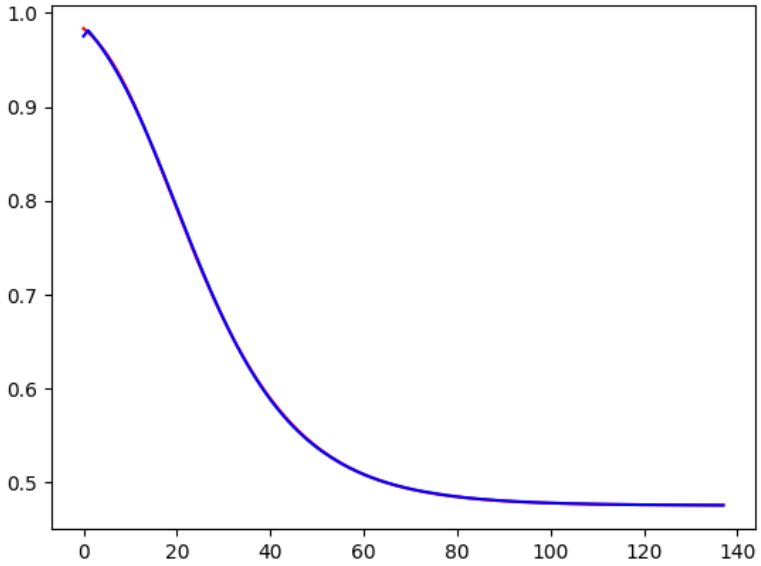
We return again to investigating the error for increased values of k, and note that even the L_2-error between $S(t)$ determined by using Ball's analytic formula and the Gompertz function of best fit was close to 10^{-3}, which suggests that this value may be a natural ceiling related to this author's computing methods and choices of numerical schemes, as well as to current universal limitations on computational tools.

2.3.6 $k = 10$ and Additional Observations

We expand yet again to $k = 10$, and make observations.

Increasing k from 5 to 10, unfortunately, does not provide us with an increased accuracy beyond that observed previously. The following figures show the numerical solution for all variables, with $\bar{x}(t)$ in pea green, and $s_i(0)$ for all $1 \leq i \leq 10$ located in the bottom quarter of the figure followed by a plot of $S(t)$ and its best-fit Gompertz function, which visually seems to fit very closely. The L_2-error here is on the order of 10^{-2}, but ignoring only the first data point and calculating from $t \geq 1$ gives us error on the order of 10^{-3}. This indicates a very close fit and perhaps also indicates close to the best-possible result for numerical simulations ran in this fashion. The fact that the number of steps the computer must perform grows quickly as a function of the number of variables may be a contributing factor to the lack of substantial improvement for increasingly large k. Whether or not it is feasible to further improve the accuracy does not dictate whether or not this endeavour has any utility, and for the time being we may decide to be satisfied with the level of error produced. As seen at the end of the previous subsection, and hopefully in the next section, the value of the variable susceptibility model is not limited to an improvement in closeness-of-fit to an appropriately chosen Gompertz function. In the final chapter, we examine some theoretical aspects of the model and display some exciting results.


1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
120
140


1.0
0.9
0.8
0.7
0.6
0.5
0
20
40
60
80
100
120
140

Chapter 3

Theoretical Study of the Model

3.1 Theoretical Considerations

Consider once again the model for variable susceptibility SIR with $k = 2$, which is in many respects the simplest example of the expanded model, and is central to our explanation of the findings.

$$\dot{s}_1(t) = -\beta x_1 s_1(t) I(t) \tag{3.1}$$

$$\dot{s}_2(t) = -\beta x_2 s_2(t) I(t) \tag{3.2}$$

$$S(t) = s_1(t) + s_2(t) \tag{3.3}$$

$$\bar{x}(t) = \frac{x_1 s_1(t) + x_2 s_2(t)}{S(t)} \tag{3.4}$$

$$\dot{I}(t) = \beta \bar{x}(t) I(t) S(t) - \gamma I(t) \tag{3.5}$$

To this point in time we are not aware of any closed form functions that perfectly represent solutions to this set of differential and algebraic equations. Whether that changes, and whatever the functions are, they are indeed continuous, which comes from general well-known ODE theory. Note here that

all of the functions are functions of a single variable, as susceptibility and recovery times are parameters, and the representation of groups are the initial conditions of the model. Since there are a limited number of types of discontinuity for single-variable functions, it can easily be deduced from the equations and the natural qualities of the dependent variables, that the equations and their derivatives are continuous. We could deduce smoothness from the equations themselves and the induction method. The functions that we fit to the numerical solutions are clearly C^∞ and are all monotone except for $I(t)$.

Perhaps a more interesting question, one which is potentially related to the question of whether or not the model fits the phenomena it intends to describe, is whether or not distinct and fundamentally different sets of initial conditions, values of the parameters, and number of groups could generate 100% identical solutions for $S(t)$. We begin with a precise description of the conditions required for our proof in the negative.

Consider first $\mathcal{R}_0$, which is equal to $\frac{\beta\bar{x}(0)}{\gamma}$ for our model. This means that it is equal to the product of average initial susceptibility and average recovery time. Adjusting only γ, say, in the positive direction would necessarily result in lower virus reproductivity and result in a different function representing $S(t)$ throughout and especially at endemicity, which would be reached sooner as well. Adjusting β accordingly, which is the simplest method of recovering the cumulative total cases before adjusting γ, would certainly have the same net-zero result in the classic SIR model. It is possible to extend these results to arbitrary values of k:

Since $\mathcal{R}_0$ must remain constant in any SIR model, that means that an equivalent set of DAE can be generated by multiplying both β and γ by the same constant factor which we call m.

We then have:

$$\dot{s}_1(t) = -m\beta x_1 s_1(t) I(t) \tag{3.6}$$

$$\dot{s}_2(t) = -m\beta x_2 s_2(t) I(t) \tag{3.7}$$

$$S(t) = s_1(t) + s_2(t) \tag{3.8}$$

$$\bar{x}(t) = \frac{x_1 s_1(t) + x_2 s_2(t)}{S(t)} \tag{3.9}$$

$$\dot{I}(t) = m\Big(\beta x_1 s_1(t) + \beta x_2 s_2(t)\Big) I(t) - m\gamma I(t) \tag{3.10}$$

By replacing the last equation with an alternative version based on equations (3.8) and (3.9), we see that the solution to eqs. (3.6), (3.7) and (3.10) do not rely on eqs. (3.8) and (3.9) but are the same functions as the solution to the full set of DAE. eq. (3.10) follows from eqs. (3.8) and (3.9) as:

$$\bar{x}(t)S(t) = \Big(\frac{x_1 s_1(t) + x_2 s_2(t)}{S(t)}\Big) S(t) = x_1 s_1(t) + x_2 s_2(t) \tag{3.11}$$

Multiplying each equation through by m should have little consequence, as the initial conditions are fixed and each of the first derivatives is simply being scaled by a constant factor. Since these are ordinary derivatives this can be characterized as a substitution:

$$\frac{df_i}{dt} \to \frac{1}{m}\frac{df_i}{dt} \tag{3.12}$$

for each function f_i. This can also be equivalently characterized by the substitution

$$t \to mt \tag{3.13}$$

We state here that for the purposes of proving a one-to-one correspondence, we ignore the fact that this system is subject to what amounts to arbitrary horizontal stretches.

With these assumptions, we shall, in the next section, present a proof for an arbitrary fixed k, and note that the proof does not rely on any of the groups $s_i(0)$ being non-empty, and thus could be used as evidence in favour of a claim that any smooth solution to the above set of DAE is uniquely determined by the complete set of initial conditions and parameters. We have from eqs. (3.6)

and (3.7), the following representation of $s_i(t)$ for arbitrary i.

$$s_i(t) = s_i(0)e^{-\beta x_i \int_0^t I(\tau)\, d\tau} \tag{3.14}$$

From

$$\dot{s}_1(t) = -\beta x_1 s_1(t) I(t) \tag{3.15}$$

Which means

$$S(t) = \sum_{i=1}^{k} s_i(0)e^{-\beta x_i \int_0^t I(\tau)\, d\tau} \tag{3.16}$$

3.2 Uniqueness

We now provide a proof of the injectivity of the following mapping:

$$\eta : (\beta, s_i(0), x_i(0)) \to S(t) \tag{3.17}$$

Consider now there being two representations of $S(t)$ that are identical (they agree pointwise for all t). Let us denote them as $S(t)$ and $\tilde{S}(t)$. We extended this notation to the individual susceptibility groups $\tilde{s}_i(t)$, order them monotonically by $s_i(0)$ without loss of generality, and assume that at least one of them differs from its analogue. We then arrive at a contradiction which concludes the proof. This proof relies on the smoothness of all of the functions, but there is an alternative proof in the same vein, involving iterated integration, which relies only on the integrability of these functions. Consider the following characterization:

$$\tilde{S}(t) = \sum_{i=1}^{k} \tilde{s}_i(0)e^{-\beta \tilde{x}_i \int_0^t I(\tau)\, d\tau} = \sum_{i=1}^{k} s_i(0)e^{-\beta x_i \int_0^t I(\tau)\, d\tau} = S(t) \tag{3.18}$$

We make the following substitutions and relabel:

$$n := n(t) := \int_0^t I(\tau)\, d\tau, \tag{3.19}$$

$$C_i = e^{-\beta x_i}, D_i = s_i(0) \tag{3.20}$$

$$\sum_{i=1}^{k} \tilde{D}_i \tilde{C}_i^n = \sum_{i=1}^{k} D_i C_i^n \tag{3.21}$$

Consider the following function:

$$f(n) := \sum_{i=1}^{k} \tilde{D}_i \tilde{C}_i^n - \sum_{i=1}^{k} D_i C_i^n = \tilde{S}(t) - S(t) \tag{3.22}$$

whose smoothness can be concluded from the smoothness of $S(t), \tilde{S}(t)$. This, coupled with the fact that $f(n) \equiv 0$ implies

$$f'(n) = f''(n) = f'''(n) = ... = 0 \tag{3.23}$$

for all n.

Suppose for the sake of contradiction that there is at least one $\tilde{C}_i \neq C_i$, and further assume that $\tilde{C}_1$ is unique and has the largest value of all. For the remainder of the proof, we make assumptions without loss of generality and without making further explicit mention of such.

Let

$$k_i = \frac{\log \tilde{C}_i}{\log C_i} \text{ so that } \tilde{C}_i = C_i^{k_i}, \text{ and let } k_1 > \max(1, \max_{i>1}(k_i)). \tag{3.24}$$

We assume also that $k_1 > \frac{1}{k_i}$ for all i.

We know that

$$f'(n) = \sum_{i=1}^{k} \tilde{D}_i \tilde{C}_i^n \log \tilde{C}_i - \sum_{i=1}^{k} D_i C_i^n \log C_i \tag{3.25}$$

and in general

$$f^{(j)}(n) = \sum_{i=1}^{k} \tilde{D}_i \tilde{C}_i^n (\log \tilde{C}_i)^j - \sum_{i=1}^{k} D_i C_i^n (\log C_i)^j \tag{3.26}$$

Now, we take the jth derivative of n and take the limit as n approaches 0 from the right.

Which means that

$$\lim_{n \to 0^+} f^{(j)}(n) = \sum_{i=1}^{k} \tilde{D}_i (\log \tilde{C}_i)^j - \sum_{i=1}^{k} D_i (\log C_i)^j = 0 \tag{3.27}$$

However, we can rewrite the above equation as follows:

$$\lim_{n \to 0^+} f^{(j)}(n) = \sum_{i=1}^{k} \tilde{D}_i k_i^j (\log C_i)^j - \sum_{i=1}^{k} D_i (\log C_i)^j = 0 \tag{3.28}$$

Now recall that k_1 is the largest of all the k_i's and their reciprocals.

We can rewrite the above equation as follows:

$$\tilde{D}_1 k_1^j (\log C_1)^j = D_1 (\log C_1)^j - \sum_{2 \le i \le k} \tilde{D}_i k_i^j (\log C_i)^j + \sum_{2 \le i \le k} D_i (\log C_i)^j \tag{3.29}$$

However, since we assumed that $k_1 > \max\{k_i, \frac{1}{k_i}\}$ for $i \neq 1$, we know that the base term on the left hand side of eq. (3.29) is larger than any of the base terms on the right hand side. The number of terms on the right hand side is equal to $2k - 1$.

We can rewrite eq. (3.29) as

$$a^j = \sum_{2 \le i \le 2k} c_i \alpha_i^j a^j \tag{3.30}$$

By letting

$$a = k_1 (\log C_1), \alpha_i = \frac{k_i (\log C_i)}{a} \text{ for } 2 \le i \le k, \alpha_i = \frac{(\log C_{i-k})}{a} \text{ for } k{+}1 \le i \le 2k \tag{3.31}$$

and,

$$c_i \in \{\frac{\pm\tilde{D}_i}{\tilde{D}_1}, \frac{\pm D_{i-k}}{\tilde{D}_1}\} \tag{3.32}$$

Rearranging again, we have

$$1 = \sum_{2\leq i\leq 2k} c_i \alpha_i^j \tag{3.33}$$

and by assumption

$$\alpha_i < 1, 2 \leq i \leq 2k \tag{3.34}$$

but $\lim_{n\to\infty} cx^n = 0, x < 1$, which means the above equation is false when j is large, representing $f^{(j)}(n)$ for n close to 0.

This means that our assumption that $\tilde{C}_1 > \max\{C_i\}$ cannot be true, which means without loss of generality that $\tilde{C}_1 = C_1$ and therefore $k_1 = 1$.

Further, we can rewrite eq. (3.29) as

$$\tilde{D}_1 k_1^j (\log C_1)^j - D_1 (\log C_1)^j = -\sum_{2\leq i\leq k} \tilde{D}_i k_i^j (\log C_i)^j + \sum_{2\leq i\leq k} D_i (\log C_i)^j \tag{3.35}$$

and we know that $k_1 = 1$, so we have

$$(\tilde{D}_1 - D_1)(\log C_1)^j = -\sum_{2\leq i\leq k} \tilde{D}_i (\log C_i)^j + \sum_{2\leq i\leq k} D_i (\log C_i)^j \tag{3.36}$$

then,

$$(\tilde{D}_1 - D_1) = \sum_{2\leq i\leq k} (D_i - \tilde{D}_i)\Big(\frac{(\log C_i)}{(\log C_1)}\Big)^j p^j \tag{3.37}$$

Where $p \in \{1, k_i\}$, and thus $p \leq 1$

Also note that $\frac{(\log C_i)}{(\log C_1)} < 1$

This means that as j approaches infinity, the right hand side approaches 0, and since the left hand side is a constant and this equation must hold for all j, then $\tilde{D}_1 = D_1$

We now rewrite eq. (3.38) after setting the left hand side as 0 and repeat the argument for the remaining indices:

$$\sum_{2\leq i\leq k} \tilde{D}_i \tilde{C}_i^n = \sum_{2\leq i\leq k} D_i C_i^n \tag{3.38}$$

In conclusion, the function $\eta : (\beta, s_k(0), x_k(0)) \to S(t)$ is one-to-one.

We now illustrate an alternative proof that does not rely on the smoothness of the solution functions. Consider again the same setup as before:

$$\sum_{i=0}^{k} \tilde{s}_i(0) e^{-\beta \tilde{x}_i \int_0^t I(\tau)\, d\tau} = \sum_{i=0}^{k} s_i(0) e^{-\beta x_i \int_0^t I(\tau)\, d\tau} \tag{3.39}$$

Which we relabel:

$$\sum_{i=0}^{k} \tilde{D}_i \tilde{C}_i^n = \sum_{i=0}^{k} D_i C_i^n \tag{3.40}$$

We can relax our requirements and claim only that $I(t)$ is almost surely continuous, which means that

$$n(t) = \int_0^t I(\tau) d\tau \tag{3.41}$$

is a continuous function, which means that

$$\int_0^{n(T)} \big(\sum_{i=0}^{k} \tilde{D}_i \tilde{C}_i^n - \sum_{i=0}^{k} D_i C_i^n \big) dn \equiv 0 \tag{3.42}$$

For time T being the exit phase of the epidemic.

The last equality must be true as the integrand itself is identically equal to 0.

However the left hand side of eq. (3.42) is also equal to

$$\big(\sum_{i=0}^{k} \frac{\tilde{D}_i}{\log \tilde{C}_i} \tilde{C}_i^n - \sum_{i=0}^{k} \frac{D_i}{\log C_i} C_i^n \big) \equiv 0 \tag{3.43}$$

Which by the same logic means that

$$\Big(\sum_{i=0}^{k} \frac{\tilde{D}_i}{(\log \tilde{C}_i)^2}\tilde{C}_i^n - \sum_{i=0}^{k} \frac{D_i}{(\log C_i)^2}C_i^n\Big) = \int_0^{n(T)} \Big(\sum_{i=0}^{k} \frac{\tilde{D}_i}{\log \tilde{C}_i}\tilde{C}_i^n - \sum_{i=0}^{k} \frac{D_i}{\log C_i}C_i^n\Big) \equiv 0 \tag{3.44}$$

and in general, after m integrations we have:

$$\Big(\sum_{i=0}^{k} \frac{\tilde{D}_i}{(\log \tilde{C}_i)^m}\tilde{C}_i^n - \sum_{i=0}^{k} \frac{D_i}{(\log C_i)^m}C_i^n\Big) \equiv 0 \tag{3.45}$$

for all m.

However, we again assume that there is an extreme nonzero value, this time a minimum value, say, $(\log \tilde{C}_1) := \tilde{k}_1$ which is smaller than $(\log C_1)$ and smaller than all other $(\log \tilde{C}_i)$.

We can rewrite eq. (3.45) as follows:

$$\Big(\frac{\tilde{D}_1}{(\tilde{k}_1)^m}\tilde{C}_1^m + \sum_{i=2}^{k} \frac{\tilde{D}_i}{(\log \tilde{C}_i)^m}\tilde{C}_i^n - \sum_{i=1}^{k} \frac{D_i}{(\log C_i)^m}C_i^n\Big) \equiv 0 \tag{3.46}$$

and by assumption, $\frac{\tilde{k}_1}{\log C_i} < 1$ and $\frac{\tilde{k}_1}{\log \tilde{C}_i} < 1$ for $i \neq 1$

Allowing m to increase in magnitude (taking the limit as m approaches infinity) and multiplying through by $(\tilde{k}_1)^m$ we obtain:

$$\tilde{D}_1\tilde{C}_1^n \equiv 0, 0 \leq n \leq n(T) \tag{3.47}$$

This implies that $C_1 = 0$, which violates our assumptions. Therefore all C_i and $\tilde{C}_i$'s are equivalent, up to a reordering of the indices. Since this must hold for all n in an interval which includes 0, we know that the same is true for the D_i's and our proof is complete. As in the original proof, the existence of an exact correspondence between C_i and $\tilde{C}_i$ for all i comes from repeating the process i times.

These results seem consequential in the search for a more accurate method of predicting epidemic behaviour. To the extent to which historical epidemic data follows the unknown solution to the mathematical model, one can at least make the assumption that any given historical course of epidemic behaviour is

a result of a small range of possible susceptibility landscapes, with $\mathcal{R}_0 = \beta\bar{x}(0)$ being completely determined by the proportion of infectives at the end of the epidemic's life up to the level accuracy and precision of this proportion.

It is unfortunate that to this point in time, efforts to justify Gompertz functions as a class of exact solutions to the variable susceptibility SIR model have fallen short. Perhaps even more unfortunate is that we are not aware whether this justification is possible or not. What is clear from the findings here is that a well-suited Gompertz function is often very close to the solution to the theoretical model, and bridging this gap may be the key to predicting long-term epidemic behaviour from initial data. The combination of this closeness along with the uniqueness of initial conditions and parameters required for an exact fit gives case for hope. In the final section of this chapter, we explore an algorithm that can be used to retrieve the initial conditions for susceptibilities and their relative representation in groups. We illustrate the algorithm using $k = 2$ for simplicity and discuss its possibilities as well as its limitations.

3.3 The Algorithm

We now examine an algorithm, that, under certain conditions, could retrieve the initial susceptibility scheme of susceptibilities as well as the representation of each group. We hypothesize that the fit to the historical data by the Gompertz function is an adequate approximation to the exact shape and scale of the data which it fits and likewise that the result achieved by the algorithm is an adequate approximation to the true spread of susceptibilities, which otherwise seems hopeless to measure. Without marrying the algorithm to tangible data we claim that any numerical method that can solve the algorithm can retrieve the same initial conditions that generate the data, up to the grand result of the numerical error combined with the error in approximating the unknown solution functions with Gompertz functions, which presumably would be a small range in possible values of βx_i and $s_i(0)$ for each $1 \leq i \leq k$.

We discuss the algorithm and its possibilities, followed by its current limitations, after introducing a claim which forms the basis of our argument. We also ignore the discrepancy between the numerical solutions to the system of DAE and the Gompertz function fits and assume their equality to form the basis of the algorithm. Python's symbolic math package SymPy provides ease of calculation.

Theorem 3.1 *If $I(0)$, $S(0), X(0)$ are known, then so are $I'(0), I''(0), ...$ $S'(0), S''(0), ...$ $X'(0), X''(0), ...$*

Proof 3.1 *Consider the following subset of equations that generate $S(t)$:*

$$\dot{s}_k = -\beta x_k s_k I \tag{3.48}$$

$$\dot{I} = \beta X I S - \gamma I \tag{3.49}$$

We also assume

$$S(0) = 1 - S_\infty e^{-e^b} = \sum_{i=0}^{k} s_i(0) := D_0 \tag{3.50}$$

$$S'(0) = -aS_\infty e^{-e^b} = \sum_{i=0}^{k} \beta x_i s_i(0) I(0) := D_1 \tag{3.51}$$

$$S''(0) = a^2(1-e^b)S_\infty e^{-e^b} = \sum_{i=0}^{k} -\beta x_i s_i(0) I'(0) + (\beta x_i)^2 I^2(0) := D2 \tag{3.52}$$

$$... \tag{3.53}$$

and

$$\bar{x}(t) = \sum_{i=0}^{k} x_i s_i(t) \tag{3.54}$$

which means

$$\frac{d}{dt}\bar{x}(t) = \sum_{i=0}^{k} x_i \frac{d}{dt} s_i(t) = \sum_{i=0}^{k} -(\beta x_i)^2 s_i(t) I(t) \tag{3.55}$$

but,

$$I(0) = 1 - S(0) = 1 - D_0 \tag{3.56}$$

and, based on equation eq. (3.49)

$$I'(0) = (\beta X(0) S(0) - \gamma) I(0) \tag{3.57}$$

Where each term in the above equation is known, including $X(0)$ *as*

$$X(0) = \sum_{i=0}^{k} -\beta x_i s_i(0) = \frac{D_1}{I(0)} \tag{3.58}$$

$$I''(0) = \beta(X'(0)S(0)I(0) + X(0)S'(0)I(0) + X(0)S(0)I'(0)) - \gamma I'(t) \tag{3.59}$$

In general:

$$I^{(j)}(0) = F(I(0), S(0), X(0), I'(0), S'(0), X'(0), ...I^{(j-1)}(0), S^{(j-1)}(0), X^{(j-1)}(0)) \tag{3.60}$$

$$S^{(j)}(0) = F(I(0), S(0), X(0), I'(0), S'(0), X'(0), ...I^{(j-1)}(0), S^{(j-1)}(0), X^{(j-1)}(0)) \tag{3.61}$$

$$X^{(j)}(0) = F(I(0), S(0), X(0), I'(0), S'(0), X'(0), ...I^{(j-1)}(0), S^{(j-1)}(0), X^{(j-1)}(0)) \tag{3.62}$$

and since $I(0), X(0), S(0)$ are all known, all of their derivatives are known by the principle of strong induction. □

Theorem 3.2 *$S(0), S'(0), S''(0), ...$ generates a system of nonlinear equations with $2k$ variables with an invertible Jacobian*

Proof 3.2 *Recall:*

$$S(0) = 1 - S_\infty e^{-e^b} = \sum_{i=0}^{k} s_i(0) := D_0 := E_0 \tag{3.63}$$

$$S'(0) = -aS_\infty e^{-e^b} = \sum_{i=0}^{k} -\beta x_i s_i(0) I(0) := D_1 := E1 \tag{3.64}$$

$$S''(0) = -ae^b S_\infty e^{-e^b} = \sum_{i=0}^{k} -\beta x_i s_i(0) I'(0) + (\beta x_i)^2 I^2(0) := D_2 \tag{3.65}$$

Let

$$\sum_{i=0}^{k} (\beta x_i)^2 I^2(0) = -\frac{I'(0)}{I(0)} D_1 + D_2 := E_2 \tag{3.66}$$

and in general, we have an infinite system of equations, one for each derivative of $S(t)$ evaluated at 0. We can rewrite these equations as:

$f_1(s_1(0), ...s_k(0), \beta x_1, ...\beta x_k)$,.. $f_{2k}(s_1(0), ...s_k(0), \beta x_1, ...\beta x_k)$ *Where each* f_i *is equal to* $E_i - E_i = 0$

We have the following vector F *representing the set of equations* f*:*

$$\begin{bmatrix} s_1(0) + ... + s_k(0) - E_0 \\ s_1(0)\beta x_1 I(0) + ... + s_k(0)\beta x_k I(0) - E_1 \\ s_1(0)(\beta x_1)^2 I^2(0) + ... + s_k(0)(\beta x_k)^2 I^2(0) - E_2 \\ . \\ . \\ . \\ s_1(0)(\beta x_1)^{2k} I^{2k}(0) + ... + s_k(0)(\beta x_k)^{2k} I^{2k}(0) - E_{2k} \end{bmatrix}$$

and its Jacobian:

$$\begin{bmatrix} 1 & ... & 1 & 0 & ... & 0 \\ \beta x_1 I(0) & ... & \beta x_k I(0) & s_1(0) & ... & s_k(0) \\ (\beta x_1 I(0))^2 & ... & (\beta x_k I(0))^2 & 2\beta x_1 I(0) s_1(0) & ... & 2\beta x_k I(0) s_k(0) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ (\beta x_1 I(0))^{2k} & ... & (\beta x_k I(0))^{2k} & 2k(\beta x_1 I(0))^{2k-1} s_1(0) & ... & 2k(\beta x_k I(0))^{2k-1} s_k(0) \end{bmatrix}$$

This system can generally be solved numerically using newton's method or other appropriate numerical methods. One should take note that two identical values of x_i and x_j with $i \neq j$ which could yield an uninvertible Jacobian would call for a regrouping. This is due to the fact that susceptibility groups s_i, s_j with the same susceptibility x_i, x_j should be considered parts of a larger susceptibility group. We conclude this section with a short example of the algorithm for $k = 2$. We take a trivially modified version of the above system of equations by replacing E_2 with D_2.

We now show the following system of equations extrapolated from our system of DAEs evaluated at $t = 0$:

$$\begin{bmatrix} s_1(0) + ... + s_k(0) - D_0 \\ -s_1(0)\beta x_1 I(0) - s_2(0)\beta x_2 I(0) - D_1 \\ s_1(0)(\beta x_1)^2 I^2(0) - s_1(0)\beta x_1 I(0) I'(0) + s_2(0)(\beta x_2)^2 I^2(0) - s_2(0)\beta x_2 I(0) I'(0) - D_2 \\ -s_1(0)(-\beta x_1)^3 I^3(0) - ... \ ... - s_2(0)\beta x_2 I(0) I'(0) I'''(0) - D_3. \end{bmatrix}$$

For testing purposes, the following initial conditions produced a plot of $S(t)$. Note that the initial conditions take the form of the above representation, which can be calculated from the given values of $s_i(0), \beta x_i(0)$. It is of no consequence that they are in their usual form when inputted into the DAE solver.

$D_0 = S(0) = 0.95, D_1 = -0.00926192516980103, D_2 = -0.03352400468827454, D_3 = 0.03257776930443156$

These values, when applied to a numerical solver, can produce a vector which we denote p that lists the variables in the following order for $k = 2$: $(s_1(0), s_2(0), \beta x_1, \beta x_2)$ The values corresponding to the above values of D_i are:

$(s_1(0), s_2(0), \beta x_1, \beta x_2) = (0.5828, 0.3672, 0.1463, 0.2764)$

Unfortunately when applying the algorithm to values of D_i generated only by $S(0)$ and $I(0)$, where their derivatives are calculated using the techniques illustrated earlier in the section, we get very different values for the above vector, which are not listed here, and yield the following results:

$(s_1(0), s_2(0), \beta x_1, \beta x_2) = (-4.806 * 10^{-11}, 0.9500, -2.1957 * 10^5, 0.1950)$

This is unfortunate as the algorithm puts the entire population into one susceptibility group. It seems that this problem is caused by the discrepancy between the vector produced by evaluating the derivatives of the Gompertz functions modelling and representing $I(t)$ and $S(t)$ and the vector produced by the true values used to generate the numerical data. At this time the result of this discrepancy is a solution which (effectively) does not take into account the existence of more than one susceptibility group. Finding a way to bridge this gap could be the key to producing an effective forecasting algorithm that necessitates only early data, if such a thing is possible.

For now we present psuedosolutions based on a modification to the calculation techniques that uses both known and calculated values. We calculate the

initial value of vectors based on finite difference coefficients for $I(0), I'(0)...$ and therefore $S(0), S'(0)...$ by combining them with the original values of the initial susceptibility conditions in the equations listed in the vector. The known values

$0.95, -0.00926192516980103, -0.03352400468827454, 0.03257776930443156$

versus the psuedosolutions' input values:

$0.9500000000, -0.009333572087, -0.03177985912, 0.005279331335$

Produces the two sets of initial conditions, the first of which was used to generate the data representing $(s_1(0), s_2(0), \beta x_1 \beta x_2)$

$0.582815832789, 0.367184167210, 0.146250000005, 0.276250000008$

$0.58281583278, 0.367184167210, 0.146250000005, 0.276250000008$

The fit is exact to 10 decimal places.

Attempts at solving these equations for higher values of k and/or without resorting to using initial values the algorithm was designed to find would be contingent on finding a pattern and a way of reconciling the difference between the true initial susceptibility scheme and that generated by the data. Alas, the place of most disagreement between the solution to the DAE and the Gompertz function fit is at $t = 0$.

3.4 Future Directions

Though these attempts at finding a nonparametric closed-form solution to the variable susceptibility SIR model fell short, this author hopes that these discoveries lay the groundwork for more precise and useful epidemic models. The variable susceptibility model adds another dimension to the previous best-available model, and the dimension of heterogenous susceptibility (or alternatively, immunity) seems like it would be an integral factor in the true-to-life dynamics of individual outbreaks, the intricacies of which we can never fully decipher. Indeed, universal homogenous susceptibility seems unable to account for the wide variety in global epidemic behaviour, even when adjusting for countless other factors such as vaccines, contact behaviour, and population density. Immunity has, in recent centuries, been thought to be a complex, dynamic process influenced by a wide range of factors. If susceptibility, as well

as transmissivity (which can generally be absorbed into one category in the model) is the driving force behind the large observed variance in epidemic outcome, then these results give us promise for the future.

The literature provides evidence that epidemic data follows Gompertz function behaviour closely, as do the numerical results of the model. Though early-stage epidemic behaviour is the most different from Gompertz function behaviour, an obvious key to predicting future outbreaks would be a pattern in the relationship between the early-stage behaviour of an epidemic and the grand total of epidemic behaviour from outbreak to extinction or into endemicity. This could be explored numerically if nothing else.

The algorithm provided would work in an ideal world where models captured phenomenon completely. Due to the complex nature of epidemics and the natural world in general, this is too much to hope for. If the relationship between the Gompertz function fits and the initial conditions and early behaviour that generated it can be decoded, then early behaviour could be enough to predict the course of the epidemic, which is always the goal. In this author's and many other's opinion, there is much room for improvement in our best available models. It is incumbent on us as a species to work together to come to an understanding of the true nature of things, while many expected and unexpected stumbling blocks line our paths. We will never comprehend the butterfly effect of good people working together to conquer the challenges to our livelihoods and survival, but nobody can deny what kind of a different world we would have without that. It is this author's sincere hope that the results herein provide a starting point for a new approach to epidemic modelling. Much disagreement has resulted from differing opinions on pandemic management in recent years. Division is only necessary until there is a resolution, and with the stakes of an epidemic outbreak, we must pay the piper and put in the work to find ways to predict the paths of these diseases, or have the correct response be a mere afterthought after the crisis has had its lasting effects.

Chapter 4

Appendix

4.1 Appendix A: Characterization of $\bar{x}(t)$

Here we provide detailed steps leading to the following characterization of $\bar{x}(t)$:

$$\bar{x}(t) = \frac{e^{(\gamma-a)t}(\frac{1}{S} - 1)}{C - \beta\frac{e^{-b}}{a}\int_0^t e^{\gamma\tau}\dot{S}\,d\tau} \tag{4.1}$$

$$C = (\frac{1}{S(0)} - 1) \ll 1 \tag{4.2}$$

Recall

$$\dot{s}_i = -\beta x_i s_i I \tag{4.3}$$

$$\dot{I} = \beta\bar{x}IS - \gamma I \tag{4.4}$$

$$S(t) = \sum_{i=1}^{k} s_i(t) \tag{4.5}$$

$$\bar{x}(t) = \frac{(\sum_{i=1}^{k} x_i s_i(t))}{(\sum_{i=1}^{k} s_i(t))} \tag{4.6}$$

Given this system, and the assumption that the solution is closely approximated by the equation $S(t) = 1 - S_\infty e^{e^{b-at}}$, we combine the two and proceed as follows.

$$1 - S(t) = S_\infty e^{-e^{b-at}} \tag{4.7}$$

And also, from Equation 4.4

$$I(t) = I(0)e^{\int_0^t \beta S x - \gamma \, d\tau} \tag{4.8}$$

and from this and Equation 4.3

$$\dot{S} = -\beta \bar{x} I_0 e^{\int_0^t \beta S \bar{x} - \gamma \, d\tau} S \tag{4.9}$$

Whereas from Equation 4.7 we obtain

$$\dot{S} = -a e^{-e^{b-at}} (\frac{1}{S} - 1) S \tag{4.10}$$

Therefore

$$a e^{-e^{b-at}} (\frac{1}{S} - 1) S = \beta \bar{x} I_0 e^{\int_0^t \beta S \bar{x} - \gamma \, d\tau} S \tag{4.11}$$

Which implies

$$a e^{b-at} (\frac{1}{S} - 1) = \beta \bar{x} I_0 e^{\int_0^t \beta S \bar{x} - \gamma \, d\tau} \tag{4.12}$$

Taking the natural log of both sides we obtain

$$\log a - at + b + \log(\frac{1}{S} - 1) = \log \beta I_0 + \log \bar{x} + \int_0^t \beta S \bar{x} - \gamma \, d\tau \tag{4.13}$$

Taking the derivative with respect to t of both sides

$$-a + \frac{1}{\frac{1}{S} - 1} \frac{-\dot{S}}{S^2} = \frac{\dot{\bar{x}}}{\bar{x}} + \beta \bar{x} S - \gamma \tag{4.14}$$

or

$$-a - \frac{\dot{S}}{S - S^2} = \frac{\dot{\bar{x}}}{\bar{x}} + \beta \bar{x} S - \gamma \tag{4.15}$$

Rearranging

$$\frac{\dot{\bar{x}}}{\bar{x}} = (\gamma - a) - \frac{\dot{S}}{S - S^2} - \beta \bar{x} S \tag{4.16}$$

$$\frac{\dot{\bar{x}}}{\bar{x}^2} = \frac{1}{\bar{x}} \left((\gamma - a) - \frac{\dot{S}}{S - S^2} \right) - \beta S \tag{4.17}$$

Let $y = \frac{1}{\bar{x}} \rightarrow \dot{y} = -\frac{\dot{\bar{x}}}{\bar{x}^2}$

$$-\dot{y} = y\left((\gamma - a) - \frac{\dot{S}}{S - S^2}\right) - \beta S \tag{4.18}$$

$$\dot{y} + y\left((\gamma - a) - \frac{\dot{S}}{S - S^2}\right) = \beta S \tag{4.19}$$

This is a linear first order ODE with $\beta S = q(t)$, $\gamma - a - \frac{\dot{S}}{S-S^2} = p(t)$, with integrating factor $\mu(t)$ as follows.

$$\mu(t) = e^{\int_0^t (\gamma - a) - \frac{\dot{S}}{S-S^2}, d\tau} \tag{4.20}$$

$$\mu(t) = e^{(\gamma-a)t + \log(\frac{1}{S} - 1)} \tag{4.21}$$

$$= (\frac{1}{S} - 1)e^{(\gamma-a)t} \tag{4.22}$$

we obtain that

$$\mu(t)y - C = \beta \int_0^t e^{(\gamma-a)\tau}(\frac{1}{S} - 1)S d\tau \tag{4.23}$$

$$\mu(t)y - C = -\beta\frac{e^{-b}}{a} \int_0^t -ae^{(\gamma\tau)}e^{(-a\tau+b)}(1 - S)d\tau \tag{4.24}$$

Note from Equation 4.7 we have

$$\dot{S} = (1 - S)e^{b-at} \tag{4.25}$$

Therefore,

$$\mu(t)y = -\beta\frac{e^{-b}}{a} \int_0^t e^{(\gamma\tau)}\dot{S}d\tau + C \tag{4.26}$$

$$\mu(t)\frac{1}{\bar{x}} = -\beta\frac{e^{-b}}{a} \int_0^t e^{(\gamma\tau)}\dot{S}d\tau + C \tag{4.27}$$

Finally

$$\bar{x}(t) = \frac{e^{(\gamma-a)t}(\frac{1}{S} - 1)}{C - \beta\frac{e^{-b}}{a}\int_0^t e^{\gamma\tau}\dot{S}\, d\tau} \tag{4.28}$$

And, substituting initial conditions, which produces numerous simplifications at $t = 0$, we have that

$$C = (\frac{1}{S(0)} - 1) \ll 1 \tag{4.29}$$

4.2 Appendix B: Computer Codes

All of the calculations that required numerical solutions to our systems of DAE utilized Python's Julia package "DifferentialEquations.jl" [41] to solve this and modified systems of ODE corresponding to the SIR model using a step method. When fitting our results to Gompertz functions, we used Python's SciPy curve fit package. The package has a default method (forward-difference, backward-difference, netwon, etc.) corresponding to each type of equation or system, but can effortlessly be coded to be solved by different numerical methods, and these generally produce very similar results. Logic dictates that the default methods are considered to be best for each corresponding type of system.

For contrast, we display the code required to run the solver for a system of ODE for $k = 1$, versus a system of DAE for $k = 5$, which has at least one algebraic (non-differential) equation, which is always the case for $k \geq 2$. The last part of code displayed here includes an invocation of SciPy curve fit which finds best-possible fits from a given arbitrary function to data points along with code that calculates the L_2 error and plots the solution along with its curve-fitted approximation.

```
#k=1

b=1.5
#For now the program fixes "b"
B = 0.325
N = 810
G = 0.13
S=0

x1=1*random.uniform(low=0.5, high=2.5, size=None)

y1=1
```

```
def f(u,p,t):
    x, y = u
    return [-B*x1*x*y, B*x1*x*y-G*y]

u0 = [800/N, 10/N]
tspan = (0.0, 140.0)
prob = de.ODEProblem(f, u0, tspan)
sol = de.solve(prob)

#k=5

x1=[0.09153548390291504, 0.21150608966123877, 2.68604797033284, 1.0305976694578645,

y1=[0.18632752, 0.25007464, 0.08836813, 0.2664609 , 0.20876882]


for i in range(0,5):
        y1[i]=0.95*y1[i]


#Randomizes vectors x and s independently each time the program runs.


B = 0.325
N = 810
G = 0.13


#Convenient values of Beta, N, and Gamma. The system is normalized

def f(du,u,p,t):
  resid1 = - B*x1[0]*u[5]*u[0] - du[0]
  resid2 = - B*x1[1]*u[5]*u[1] - du[1]
  resid3 = - B*x1[2]*u[5]*u[2] - du[2]
  resid4 = - B*x1[3]*u[5]*u[3] - du[3]
  resid5 = - B*x1[4]*u[5]*u[4] - du[4]
  resid6 = B*u[6]*u[7]*u[5]-G*u[5] - du[5]
  resid7= u[1]+u[2]+u[3]+u[4]+u[0]-u[6]
  resid8= (x1[0]*u[0]+x1[1]*u[1]+x1[2]*u[2]+x1[3]*u[3]+x1[4]*u[4])/u[6]-u[7]
```

```
  resid9= G*u[5] - du[8]

  return [resid1, resid2, resid3, resid4, resid5, resid6, resid7, resid8, resid9]

#Initial conditions for R(t) have changed to be in line with the physical intuition

u0 = [y1[0], y1[1], y1[2], y1[3], y1[4], 10/N, 800/N, 0.7, 0.00]
du0 = [-y1[0]/N, -y1[1]/N, -y1[2]/N, -y1[3]/N, -y1[4]/N, 2/N, -0.16, -1/140, 0.00]
tspan = (0.0,140.0)
differential_vars = [True,True,True,True,True,True,False,False,True]
prob = de.DAEProblem(f,du0,u0,tspan,differential_vars=differential_vars)
sol = de.solve(prob)

u10 = [sol.u[i][6] for i in range(0,len(sol.u))]

plt.plot(sol.t, u10)
plt.show()

def func(z, a1, d, Q):
        return Q*np.exp(-np.exp(d-c*z))


xdata = np.linspace(0, 138, 139)
h=[]
for i in range(0, 139):
        h.append(0)

for i in range(0,139):
        h[i]=sol(i)[8]

y = func(xdata, a, b, S)
ydata= h
#plt.plot(xdata, ydata, 'b-', label='data')
popt, pcov = curve_fit(func, xdata, ydata)
popt

Lone=0
for i in range(0, 138):
        Lone+=(func(xdata, *popt)[i]-h[i])**2
```

```
Ltwo=math.sqrt(Lone)


h=[]
for i in range(0, 139):
        h.append(0)

for i in range(0,139):
        h[i]=sol(i)[8]









plt.plot(xdata, func(xdata, *popt), 'g')
plt.plot(xdata, h, 'r')
plt.savefig('*****')
plt.show()
```

We also here display the code used to generate the $2k \times 2k$ matrix from the algorithm. By using Python's "SymPy" symbolic math package, matrices of arbitrary size can be generated and solving attempts can be made using newton's method with a fairly short, simple python program. It is this author's opinion that the code can be condensed and streamlined even further if necessary for large k. This example is for $k = 5$.

```
s1, s2, s3, s4, s5  = symbols ('s1 s2 s3 s4 s5 ', cls=Function)

k1, k2, k3, k4, k5 = symbols('k1 k2 k3 k4 k5')

r1, r2, r3, r4, r5 = symbols('r1 r2 r3 r4 r5')

g = symbols('g', cls=Function)

s1=r1*sympy.exp(-k1*g(t))

s2=r2*sympy.exp(-k2*g(t))

s3=r3*sympy.exp(-k3*g(t))

s4=r4*sympy.exp(-k4*g(t))

s5=r5*sympy.exp(-k5*g(t))

s  = symbols ('s', cls=Function)

s=s1+s2+s3+s4+s5

f0, f1, f2, f3, f4, f5, f6, f7, f8, f9 =

symbols ('f0 f1 f2 f3 f4 f5 f6 f7 f8 f9', cls=Function)

f0=s

f1=f0.diff(t)
```

```
f2=f1.diff(t)

f3=f2.diff(t)

f4=f3.diff(t)

f5=f4.diff(t)

f6=f5.diff(t)

f7=f6.diff(t)

f8=f7.diff(t)

f9=f8.diff(t)

E=[0,0,0,0,0,0,0,0,0,0]

x1, x2, x3, x4, x5, x6, x7, x8, x9, x10 = symbols ('x1 x2 x3 x4 x5 x6 x7 x8 x9 x10')



print("TEST")



E[0]=f0.subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6),
(k2,x7), (k3,x8), (k4,x9), (k5,x10), ((Derivative(g(t), t)), C[0]), (g(0),0)])

E[1]=f1.subs(Derivative(g(t), t), C[0])

E[1]=E[1].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6),
(k2,x7), (k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[2]=f2.subs([(Derivative(g(t), (t, 2)), C[1]), (Derivative(g(t), t), C[0])])


E[2]= E[2].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6),
```

```
(k2,x7), (k3,x8), (k4,x9), (k5,x10), (g(0),0)])

E[3]=f3.subs([(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[3]=E[3].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6),
(k2,x7), (k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[4]=f4.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]), (Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[4]=E[4].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5),
 (k1,x6), (k2,x7), (k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[5]=f5.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]),(Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[5]=E[5].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6), (k2,x7)
(k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[6]=f6.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]), (Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[6]=E[6].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6), (k2,x7)
(k4,x9), (k5,x10),  (g(0),0)])

E[7]=f7.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
```

```
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]), (Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[7]=E[7].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6), (k2,x7)
 (k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[8]=f8.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]),(Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[8]=E[8].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6), (k2,x7)
(k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

E[9]=f9.subs([(Derivative(g(t), (t, 11)), C[10]),(Derivative(g(t), (t, 10)), C[9]),
(Derivative(g(t), (t, 9)), C[8]),(Derivative(g(t), (t, 8)), C[7]),
(Derivative(g(t), (t, 7)), C[6]),(Derivative(g(t), (t, 6)), C[5]),
(Derivative(g(t), (t, 5)), C[4]), (Derivative(g(t), (t, 4)), C[3]),
(Derivative(g(t), (t, 3)), C[2]), (Derivative(g(t), (t, 2)), C[1]),
(Derivative(g(t), t), C[0])])

E[9]=E[9].subs([(t,0), (r1,x1), (r2,x2), (r3,x3), (r4,x4), (r5,x5), (k1,x6), (k2,x7)
(k3,x8), (k4,x9), (k5,x10),  (g(0),0)])

for  i in range(0,10):
        E[i]=E[i]-D[i]




Y = Matrix([x1,  x2,x3,x4,x5,x6,x7,x8,x9,x10])
F = Matrix(E)
```

```
Ja=F.jacobian(Y)
F=function('F')(x1, x2, x3, x4, x5, x6, x7, x8, x9, x10)
#H=hessian(F, (x1, x2, x3, x4, x5, x6, x7, x8, x9, x10))
```

```
z=[0.03, 0.08, 0.15, 0.25, 0.25, 0.02, 0.45, 0.55,0.025,0.205]
p=z
G=F.subs([(x1,z[0]),(x2,z[1]),(x3,z[2]),(x4,z[3]),(x5,z[4]),(x6,z[5]),
(x7,z[6]),(x8,z[7]),(x9,z[8]),(x10,z[9])])
J=Ja.subs([(x1,z[0]),(x2,z[1]),(x3,z[2]),(x4,z[3]),(x5,z[4]),(x6,z[5]),
(x7,z[6]),(x8,z[7]),(x9,z[8]),(x10,z[9])])
```

```
K=J.inv()
L=[0,0,0,0,0,0,0,0,0,0]
```

```
for i in range(0,100):
        print("Running at: " +str(i))
        J=Ja.subs([(x1,p[0]),(x2,p[1]),(x3,p[2]),(x4,p[3]),(x5,p[4]),(x6,p[5]),
        (x7,p[6]),(x8,p[7]),(x9,p[8]),(x10,p[9])])
        G=F.subs([(x1,p[0]),(x2,p[1]),(x3,p[2]),(x4,p[3]),(x5,p[4]),(x6,p[5]),
        (x7,p[6]),(x8,p[7]),(x9,p[8]),(x10,p[9])])
        K=J.inv()
        L=K*G
        for j in range(0,10):
                p[j]=p[j]-L[j]
```

Bibliography

[1] Helen Abbey. An examination of the reed-frost theory of epidemics. *Human biology*, 24 3:201–33, 1952.

[2] Tanvir Ahammed, Aniqua Anjum, Mohammad Meshbahur Rahman, Najmul Haider, Richard Kock, and Md Jamal Uddin. Estimation of novel coronavirus (covid-19) reproduction number and case fatality rate: A systematic review and meta-analysis. *Health Science Reports*, 4(2):e274, 2021.

[3] Ebraheem Alzahrani and Muhammad Altaf Khan. Mathematical modeling and analysis of COVID-19 infection: application to the Kingdom of Saudi Arabia data. *J. Math.*, pages Art. ID 6623005, 20, 2023.

[4] Nicolas Bacaër. *Verhulst and the logistic equation (1838)*, pages 35–39. Springer London, London, 2011.

[5] Frank Ball. Deterministic and stochastic epidemics with several kinds of susceptibles. *Advances in Applied Probability*, 17(1):1–22, 1985.

[6] Frank Ball. A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models. *Advances in Applied Probability*, 18(2):289–310, 1986.

[7] ANDREW D. BARBOUR. The duration of the closed stochastic epidemic. *Biometrika*, 62(2):477–482, 08 1975.

[8] Moises E Bauer, Cristina M Moriguchi Jeckel, and Clarice Luz. The role of stress factors during aging of the immune system. *Annals of the New York Academy of Sciences*, 1153(1):139–152, 2009.

[9] Niels Becker. The uses of epidemic models. *Biometrics*, 35(1):295–305, 1979.

[10] Andrea L. Bertozzi, Elisa Franco, George Mohler, Martin B. Short, and Daniel Sledge. The challenges of modeling and forecasting the spread of covid-19. *Proceedings of the National Academy of Sciences*, 117(29):16732–16738, July 2020.

[11] Z.W. Birnbaum. Origins and fundamental concepts of the competing-risks problem. *Environment International*, 1(6):303–308, 1978.

[12] Ágnes Bodó. SIS epidemic propagation on hypergraphs. *Alkalmaz. Mat. Lapok*, 36(2):189–195, 2019.

[13] Martin Bohner, Sabrina Streipert, and Delfim F.M. Torres. Exact solution to a dynamic sir model. *Nonlinear Analysis: Hybrid Systems*, 32:228–238, May 2019.

[14] Fred Brauer. Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling*, 2(2):113–127, May 2017. Publisher Copyright: © 2017 KeAi Communications Co., Ltd.

[15] Tom Britton. Stochastic epidemic models: a survey, 2009.

[16] Marcos A. Capistrán, J. Andrés Christen, and Jorge X. Velasco-Hernández. Towards uncertainty quantification and inference in the stochastic SIR epidemic model. *Math. Biosci.*, 240(2):250–259, 2012.

[17] Ana Cascon and William F. Shadwick. Predicting the course of covid-19 and other epidemic and endemic disease. *medRxiv*, 2021.

[18] Ding-Geng Chen, Xinguang Chen, and Jenny Chen. Reconstructing and forecasting the covid-19 epidemic in the united states using a 5-parameter logistic growth model. *Global Health Research and Policy*, 5, 12 2020.

[19] D. J. Daley and J. Gani. *Epidemic Modelling: An Introduction*. Cambridge Studies in Mathematical Biology. Cambridge University Press, 1999.

[20] Bibha Dhungel, Md. Shafiur Rahman, Md. Mahfuzur Rahman, Aliza K. C. Bhandari, Phuong Mai Le, Nushrat Alam Biva, and Stuart Gilmour. Reliability of early estimates of the basic reproduction number of covid-19: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, 19(18), 2022.

[21] Shasha Gao, Mingwang Shen, Xueying Wang, Jin Wang, Maia Martcheva, and Libin Rong. A multi-strain model with asymptomatic transmission:

application to COVID-19 in the US. *J. Theoret. Biol.*, 565:Paper No. 111468, 14, 2023.

[22] J. Graunt. *Natural and political observations mentioned in a following index, and made upon the bills of mortality.* John Martyn and James Allestry, London, 1662.

[23] Tiberiu Harko, Francisco S.N. Lobo, and M.K. Mak. Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates. *Applied Mathematics and Computation*, 236:184–194, 2014.

[24] Yunhan Huang and Quanyan Zhu. Game-theoretic frameworks for epidemic spreading and human decision-making: a review. *Dyn. Games Appl.*, 12(1):7–48, 2022.

[25] James Hyman and Jia Li. Differential susceptibility epidemic models. *Journal of mathematical biology*, 50:626–44, 07 2005.

[26] James Hyman and Jia Li. Infection-age structured epidemic models with behavior change or treatment. *Journal of biological dynamics*, 1:109–31, 01 2007.

[27] Wahab A. Iddrisu, Inusah Iddrisu, and Abdul-Karim Iddrisu. Modeling cholera epidemiology using stochastic differential equations. *J. Appl. Math.*, pages Art. ID 7232395, 17, 2023.

[28] John P.A. Ioannidis, Sally Cripps, and Martin A. Tanner. Forecasting for COVID-19 has failed. *International Journal of Forecasting*, 38(2):423–438, 2022.

[29] William Ogilvie Kermack and À. G. Mckendrick. A contribution to the mathematical theory of epidemics. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 115:700–721, 1927.

[30] M Kröger and R Schlickeiser. Analytical solution of the sir-model for the temporal evolution of epidemics. part a: time-independent reproduction factor. *Journal of Physics A: Mathematical and Theoretical*, 53(50):505601, nov 2020.

[31] Alexanderia Lacy, Morganne Igoe, Praachi Das, Trevor Farthing, Alun L. Lloyd, Cristina Lanzas, Agricola Odoi, and Suzanne Lenhart. Modeling impact of vaccination on COVID-19 dynamics in St. Louis. *J. Biol. Dyn.*, 17(1):2287084, 25, 2023.

[32] E. Leuridan, M. Sabbe, and P. Van Damme. Measles outbreak in europe: Susceptibility of infants too young to be immunized. *Vaccine*, 30(41):5905–5913, 2012.

[33] Zhiming Li and Zhidong Teng. Analysis of uncertain SIS epidemic model with nonlinear incidence and demography. *Fuzzy Optim. Decis. Mak.*, 18(4):475–491, 2019.

[34] Xin Liu, Anuj Mubayi, Dominik Reinhold, and Liu Zhu. Approximation methods for analyzing multiscale stochastic vector-borne epidemic models. *Math. Biosci.*, 309:42–65, 2019.

[35] Shravan Luckraz. A survey on the relationship between the game of cops and robbers and other game representations. *Dyn. Games Appl.*, 9(2):506–520, 2019.

[36] Shruti Magesh, Daniel John, Wei Tse Li, Yuxiang Li, Aidan Mattingly-app, Sharad Jain, Eric Y. Chang, and Weg M. Ongkeko. Disparities in COVID-19 Outcomes by Race, Ethnicity, and Socioeconomic Status: A Systematic Review and Meta-analysis. *JAMA Network Open*, 4(11):e2134147–e2134147, 11 2021.

[37] Krzysztof Michalak and Joshua Knowles. Simheuristics for the multiobjective nondeterministic firefighter problem in a time-constrained setting. pages 248–265, 04 2016.

[38] Harvey Moldofsky. Sleep and the immune system. *International Journal of Immunopharmacology*, 17(8):649–654, 1995. Psychoimmune Networks in Health and Disease: Proceedings of the Third International Scientific Meeting of the Australian Behavioral Immunology Group.

[39] Jonas Neipel, Jonathan Bauermann, Stefano Bo, Tyler Harmon, and Frank Jülicher. Power-law population heterogeneity governs epidemic waves. *PLOS ONE*, 15(10):1–27, 10 2020.

[40] Nhu N. Nguyen and George Yin. Stochastic partial differential equation SIS epidemic models: modeling and analysis. *Commun. Stoch. Anal.*, 13(3):Art. 8, 22, 2019.

[41] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1), 2017.

[42] William T. Riley and George A. Mensah. Perspective: Social determinants of health and implementation research: Lessons from the covi d-19 pandemic. *Ethnicity & Disease*, 31(1):pp. 5–8, 2021.

[43] S. Rushton and A. J. Mautner. The deterministic model of a simple epidemic for more than one community. *Biometrika*, 42(1-2):126–136, 06 1955.

[44] Zhenfeng Shi and Daqing Jiang. Stochastic modeling of SIS epidemics with logarithmic Ornstein-Uhlenbeck process and generalized nonlinear incidence. *Math. Biosci.*, 365:Paper No. 109083, 15, 2023.

[45] Syafruddin Side, Melisa Utami, Sukarna Sukarna, and Muh Pratama. Numerical solution of sir model for transmission of tuberculosis by rungekutta method. *Journal of Physics: Conference Series*, 1040:012021, 06 2018.

[46] Courtney R. Smith, Charmaine Enns, Dan Cutfeet, Shannon Alfred, Nicole James, Jennifer Lindbeck, and Shannon Russell. Covid-19 in a remote first nations community in british columbia, canada: an outbreak report. *Canadian Medical Association Open Access Journal*, 9(4):E1073–E1079, 2021.

[47] Yoshito Takasaki, Christian Abizaid, and Oliver Coomes. Covid-19 contagion across remote communities in tropical forests. *Scientific Reports*, 12:20727, 12 2022.

[48] Thomas Torku, Abdul Khaliq, and Fathalla Rihan. SEINN: a deep learning algorithm for the stochastic epidemic model. *Math. Biosci. Eng.*, 20(9):16330–16361, 2023.

[49] Rob Trangucci, Yang Chen, and Jon Zelner. Modeling racial/ethnic differences in COVID-19 incidence with covariates subject to nonrandom missingness. *Ann. Appl. Stat.*, 17(4):2723–2758, 2023.

[50] Nikolay K. Vitanov and Kaloyan N. Vitanov. Epidemic waves and exact solutions of a sequence of nonlinear differential equations connected to the sir model of epidemics. *Entropy*, 25(3), 2023.

[51] Connor Wagner. A new survey on the firefighter problem, *Master's thesis, University of Victoria*, June 2021.

[52] Chuanqing Xu, Zonghao Zhang, Xiaotong Huang, Kedeng Cheng, Songbai Guo, Xiaojing Wang, Maoxing Liu, and Xiaoling Liu. A study on the

transmission dynamics of COVID-19 considering the impact of asymptomatic infection. *J. Biol. Dyn.*, 17(1):Paper No. 2244980, 22, 2023.