Approximations of Stochastic Household Models for Comparing Antiviral Allocation Schemes

Michael Lydeamore

Thesis submitted for the degree of

Master of Philosophy

in

Applied Mathematics

at

The University of Adelaide

(Faculty of Engineering, Computer and Mathematical Sciences)

School of Mathematical Sciences

THE UNIVERSITY of ADELAIDE

February 2015

Contents

List of Tables

List of Figures

List of Algorithms

While not explicitly listed, all code used to generate the data throughout this thesis is made available at

`https://github.com/MikeLydeamore/AntiviralAllocationSchemes`.

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: ..... DATE: .....

Acknowledgements

Thank you to everyone for their support throughout this project. Thank you to my supervisors, Dr Joshua Ross, Dr Andrew Black and Prof. Nigel Bean, and also to Judith Lydeamore for her assistance in proof reading this work. Thank you especially to my Mum, Gail, my Dad, Brian, my sister, Kylie, my girlfriend, Alex, and all of my friends and colleagues. This thesis would not have been possible without your support, ideas, suggestions, corrections and guidance. Thank you to everyone who has listened to the concepts, challenged the thoughts and supported the results.

Commonly Used Notation

Rate governing internal infection	β
Rate governing external infection	α
Rate governing progression	σ
Rate governing recovery	γ
Reduction in susceptibility due to antivirals	ρ
Reduction in infectivity due to antivirals	τ
Mean delay until antivirals arrive into household	ζ
Mean effective duration of antivirals	κ
Number of households	N
Household size	k
Mean household size	$\bar{k}$
Household size distribution	$\mathbf{h}$
Size-biased distribution	$\boldsymbol{\pi}$
Amount of available antivirals	M

Abstract

From the first recorded influenza pandemic in 1890, there have been new strains of influenza which have caused pandemics approximately every 30 years including recent events such as the H5N1 Avian 'Flu pandemic and the 2009 H1N1 Swine 'Flu pandemic. Although the 2009 pandemic was mild in nature, if events of the past are any indication then control of future pandemics is of utmost importance.

Vaccination is commonly looked at to help control the spread of a pandemic, however, vaccinations are strain-specific. While developing a new vaccine is possible, the World Health Organisation estimates that this process would take four to five months. This means that vaccination cannot be used to help control the spread of influenza early on in a pandemic. An alternative are *antivirals* which are not strain-specific, meaning that they can potentially be used to help control the spread of influenza early on in a pandemic. Antivirals are, however, not as effective at reducing the spread of disease when compared to vaccination.

In the 2009 Swine 'Flu pandemic, many countries worldwide utilised antiviral medication, with the aim to assist in controlling the spread of influenza. The most common method in which these antivirals were utilised we refer to as dynamic allocation. In dynamic allocation, when the first person in a household experiences influenza-like symptoms, they report to a health professional. Then, a sample is sent for laboratory testing. If the individual is confirmed to have influenza, the entire household is allocated a course of antivirals and every member of the household begins taking them. The potential weakness in this strategy is the delay between becoming infectious and a household receiving antivirals.

We consider an alternative antiviral allocation scheme which we call preallocation. In a preallocation scheme, instead of waiting for antivirals to be delivered after the first confirmed infection, as is the case with dynamic allocation, the antivirals are delivered to households at the beginning of the pandemic. When the first person experiences symptoms, they contact a health professional via a telephone hotline. The professional then decides if it is likely that the individual has influenza. If the individual is likely to have influenza then the entire household starts taking antivirals immediately, just as is the case in dynamic allocation. The advantage of this scheme is that the delay is essentially zero, but there is the potential for the antivirals to be wasted in at least two ways. First, this type of identification of infection is clearly less precise than laboratory testing. Second, it is possible that antivirals will be preallocated to a household who will never experience infection and so those antivirals will essentially be wasted. It is this tradeoff that is the focus of this thesis.

The stochastic households epidemic model which is detailed and developed in this work incorporates the household structure of a general population. This allows us to incorporate the stronger mixing of individuals who share a household compared to individuals in the general population, as well as the fact that antivirals are allocated to an entire household when infection is first detected. To analyse this model, we develop two approximations:
(i) A branching process approximation, and
(ii) a deterministic approximation,
that assist us in calculating quantities associated with a pandemic.

The branching process is very fast to compute, but due to required assumptions in the derivation, it is only able to describe the early stages of the pandemic. The branching process is able to rapidly compute quantities such as the Malthusian parameter, r, and the household reproductive ratio, R_*, but is unable to calculate quantities such as the *final epidemic size*, that is, the total number of people infected over the course of the pandemic.. The deterministic approximation does not allow for as rapid evaluation as the branching process approximation, but is able

to approximately reproduce the entire expected pandemic curve, giving access to quantities such as the expected final epidemic size. Both of these approximations are fast to compute so we can explore a range of parameters and compare the two allocation schemes—dynamic allocation and preallocation.

We show that preallocation of antivirals often leads to a smaller final epidemic size than dynamic allocation for a *severe* pandemic outbreak, while a dynamic allocation scheme often gives a lower Malthusian parameter, r, and household reproductive ratio, R_*. We provide a justification for this behaviour and demonstrate that the results are relatively robust across the parameters controlling the pandemic. We also consider a number of extensions to the deterministic approximation such as the incorrect use of antivirals, a hybrid allocation scheme, and the production of antivirals during the pandemic. Under these extensions, the general behaviour of the two schemes—preallocation yielding a lower final epidemic size but dynamic allocation yielding superior early-time quantities—is unchanged.

Chapter 1

Introduction

Influenza pandemics have occurred all throughout history. The Spanish 'Flu in 1918 was one of the worst influenza pandemics to date and killed an estimated 40 million people, almost 2.5% of the world's population [57]. Every year the risk of an outbreak of a novel strain of influenza is present, so it is of importance that plans are available to facilitate the control of a potential pandemic.

Control measures that can potentially be used to help lower the impact of a pandemic include vaccination, quarantine and isolation of infectious individuals, travel restrictions and antivirals. Each of these control measures has associated advantages and disadvantages. Vaccination is often thought of as the best method to stop the spread of influenza, however vaccinations are strain-specific. This means that vaccination is only effective against particular strains of influenza, and it is the *novel* strains of influenza for which a vaccine does not already exist which cause pandemics. While it is possible to develop a new vaccine for a novel strain of influenza, the World Health Organisation estimates the time to do this to be at least 5 months [75]. For the 2009 Swine 'Flu pandemic, the initial major outbreak of infection happened approximately 2 months into the pandemic, as seen in Figure 1.1. This is before a vaccine would be available, and so should this behaviour occur in future pandemics, vaccination would not be viable to use to help control the initial spread of infection. So, while vaccination is very effective at creating (temporary)

immunity in susceptible individuals when available, it is infeasible to use for the early stages of pandemic control.

Quarantine and isolation of infectious individuals is clearly very effective at preventing the spread of infection, however, correctly identifying those who are infected and isolating them before any new infection occurs is difficult and costly. This difficulty is due to the potential for individuals to be infectious without showing any symptoms, a period of time which is known as the *incubation period* in epidemiology [45]. Identifying infectious people in a timely manner is further complicated because confirmation of illness requires laboratory testing, which adds a noticeable delay to the identification process. While timely identification is important in all control methods, quarantine will have no impact if infection is not identified quickly enough.

Travel restrictions can be very effective at protecting pandemic spread between regions, however, travel restrictions alone are generally unable to stop a pandemic in isolation [24, 26, 27, 42]. In particular, Cooper *et al.* [24] determined that a reduction of at least 95% to travel would be required in order to have any noticeable impact on a pandemic. This, combined with the economics of restricting travel, suggests that travel restrictions are not a viable control scheme in isolation for controlling the spread of infection during a pandemic.

Antivirals, unlike vaccines, are not strain-specific, and so they are potentially effective against many types of influenza. This means that there is no development time, however, antivirals have a generally lower *efficacy*, or effectiveness, when compared to vaccination. The fact there is no development time means that antivirals may prove to be an effective first line of defence during an outbreak. Antivirals are already a part of the Australian Health Management Plan for Pandemic Influenza [22], and were utilised in Australia during the 2009 Swine 'Flu Pandemic [23]. Antivirals are believed to achieve two things: an infectious individual is less likely to transmit infection when contact occurs with a susceptible individual, and an individual who is not infected has a stronger immunity against infection, even when

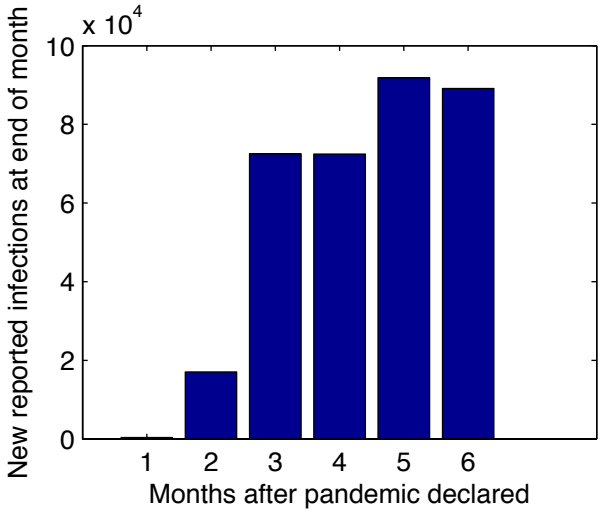


Figure 1.1: Number of new reported cases each month to the World Health Organisation during the 2009 Swine 'Flu Pandemic [76]. After 5 months, countries were no longer obligated to report every case. After 6 months, the data was no longer presented as the estimates were deemed too inaccurate.

contact is made with an infectious individual who is not taking antivirals [36]. The average duration that antivirals are taken for during a pandemic is approximately 5 days, with the potential for a second course if deemed necessary by a physician. One recommended antiviral treatment, Oseltamivir, is estimated in two independent studies to reduce the transmission rate and increase immunity to pandemic influenza by approximately 30%, although this number is likely to be variable [36, 72].

1.1 Antiviral Allocation Schemes

The Australian Health Management Plan for Pandemic Influenza [22] details the protocol by which antivirals will be used during an influenza pandemic. The allocation scheme that is detailed is as follows: after an infectious person is identified in a household that has not yet received antivirals, a course of antivirals will be allocated and the entire household takes antivirals until the course is complete, re-

gardless of the infection status of each individual inside the household. We call this scheme *dynamic allocation*, as antivirals are 'dynamically allocated' throughout the pandemic. The reason for the entire household taking antivirals, regardless of whether each individual is infectious or not is because a noticeable proportion of transmission occurs inside a household, and treatment to susceptible individuals, known as *prophylaxis*, reduces the severity of the outbreak [50]. Under the dynamic antiviral allocation scheme, the members of a household must wait until infection is confirmed, generally requiring laboratory testing, before all the members in the household begin taking antivirals. If this *delay* between becoming infectious and the individuals in the household receiving antivirals is large, then the antivirals will have little impact as the spread of infection will be complete before the antivirals have begun being taken. It has been estimated that the median time until antiviral treatment commenced during the 2009 Swine 'Flu pandemic was approximately three to four days and that the mean infectious period for an individual was approximately two to four days [29], meaning that there were some infectious individuals who did not receive antivirals until they were no longer infectious. At this point, the antivirals would have little to no effect on the pandemic as these infectious individuals had already completed all of their transmission. Susceptible individuals in an infected household would also experience a delay before antiviral treatment begins. As a noticeable proportion of infection is believed to occur within a household, the delay until the infectious individuals receive antivirals means that the susceptible individuals experience the full force of infection from infectious individuals in their own household. If the delay is large then there will be no reduction to the rate of infection at the time when infection is most likely to spread inside a household. One advantage of dynamic allocation, however, is that the antivirals will always be used in a household that is likely to spread infection in the near future.

An alternative antiviral allocation scheme, called *preallocation*, can effectively remove the delay from dynamic allocation but introduces some potential new issues. Under a preallocation scheme, instead of waiting for doctor's confirmation or lab-

oratory results, all antivirals are allocated to households in the population before the pandemic begins. When an individual begins showing symptoms of influenza they are diagnosed but in a potentially less precise way (in comparison to laboratory testing), such as contacting a government phone help-line and talking to an expert. If it is decided that the individual is likely to have influenza then all members of the household begin taking the antivirals just as they would under a dynamic allocation scheme. Under a preallocation scheme, then, the delay between becoming infectious and beginning a course of antivirals is reduced as there is no waiting for a doctor's formal diagnosis or laboratory test results, however, it is possible for people to take their antivirals incorrectly (without being ill). It is also possible to have antivirals that will be essentially wasted as infection may never occur inside a household which has been preallocated antivirals. The preallocation scheme also relies on households using the supply of antivirals correctly. As panic spreads during the pandemic, some households may be likely to take their antivirals without consultation when there is no infection in the household. In this case the household will still receive the benefit of a reduction of susceptibility but will not get the benefit of a reduction to infectivity. While this case is not as bad as allocating antivirals to a household which never experiences infection, it is still undesirable.

The dynamic allocation scheme has a delay until antivirals arrive to a household which has experienced infection, but all antivirals will be used in a situation where infection is likely to spread. Comparatively, the preallocation scheme has the potential for waste, but avoids the delay until antivirals arrive into a household. It is this tradeoff which we aim to investigate in order to determine which antiviral allocation scheme is best.

An investigation into the effectiveness of antivirals was performed with respect to an influenza outbreak in the United States of America [52]. Using a stochastic simulation, it attempted to determine how effective antivirals would be at controlling pandemic influenza. It was determined that a treatment course of four weeks to 80% of those who become infected was almost as effective as vaccinating 80% of the entire

population, however the efficacy of antivirals in this study was substantially higher than other estimates, with a 30% reduction in susceptibility and a 60% reduction in infectivity. A treatment course of four weeks is longer than what is planned for pandemic response in Australia [21], and it is unlikely that a country would be able to stockpile enough antivirals for this scheme [18]. In particular, Carrasco *et al.* note that it is not cost effective for a country to stockpile enough antivirals for more than 20% of the population. However, should there be a sufficient antiviral stockpile then this long treatment scheme may be desirable.

Comparisons between the dynamic antiviral allocation scheme and a preallocation antiviral scheme have been performed previously. Goldstein *et al.* [34] analyse the performance of *Medkits*, which are analogous to antivirals. The model utilised by Goldstein *et al.* has more population structure than the model in our work, including an age-structured population as well as a household structure. The model however is only able to be used to calculate the probability of death over an entire pandemic, and does not reproduce the dynamics of the system, unlike the methods we utilise throughout this work. The focus of their work is on the minimisation of death in the population. This is similar to minimising the *final epidemic size*. However, the models considered in this thesis do not separate death from recovery. Importantly, the preallocation antiviral scheme is shown to lead to a smaller expected number of deaths when compared to a dynamic allocation scheme, however, the required amount of *Medkits* is assumed to be high.

1.2 Pandemic Modelling

The commonly used SIR model divides a population of individuals into three distinct classes – *susceptible*, where an individual is able to catch the disease; *infectious*, where an individual shows symptoms and transmits the infection to other susceptible individuals; and *recovered*, where the individual remains after their infectious period. This introduction of this model is attributed to Kermack and McKendrick [48] in

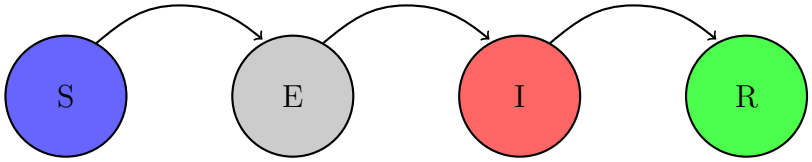


Figure 1.2: The basic SEIR model, showing the stages of an indivual throughout the pandemic.

1927, with the now commonly used SIR notation accredited to Hoppensteadt and Waltman [37] in 1970. The SIR model is commonly used to model the spread of influenza through a population [3, 7, 9, 39, 45, 63].

The SEIR model which is utilised throughout this work acts similarly to the SIR model, except that there is an *incubation period* for the disease. That is, when a susceptible individual comes into contact with an infectious person, instead of becoming instantaneously infectious themselves, the individual is considered *exposed* for some period of time and is unable to transmit infection before becoming infectious. This exposed phase would have no impact if any analysis of the model was independent of time, as when an individual is exposed they can neither infect nor be infected and so the individual has no impact on the pandemic [6]. However, for any quantity that depends on time, the exposed phase will impact results. A graphical illustration of the SEIR model can be seen in Figure 1.2. For a disease such as influenza, an incubation period, or exposed phase, appears to occur, with approximations for the incubation time for the 2009 H1N1 Swine 'Flu virus being between 1 and 1.8 days for children and a mean of 4.3 days for adults [51, 73].

1.2.1 Household Structure and Antivirals

The Australian Health Management Plan for Pandemic Influenza [22] states that once there is a single infectious person in a household, antivirals are to be taken by all members of the household, regardless of infection status. This is because it is estimated that 30% of transmission happens inside a household [15, 27]. This dynamic

antiviral allocation scheme also allows for clear definitions, such as a 'household', which helps to prevent delays in using the antivirals due to policy interpretation issues. Because of the Management Plan and because there is likely to be a different infection rate inside a household when compared to the general population, we must take into account the household structure of the population.

We assume that an individual belongs to precisely one household and does not migrate between households. Households can be of varying sizes, but are relatively small when compared to the size of the population. Finally, we assume that the size of each household remains fixed throughout the duration of the pandemic. Note that if all households are of size 1, then the household structure reduces down to the traditional SEIR model without any population structure.

House and Keeling [39] investigated a deterministic SIR pandemic model which incorporates a household structure, as well as a number of potential schemes for antiviral intervention. They determined that treating an entire household with antivirals meant the disease could be eradicated with a lower detection rate of infectious individuals compared to treating only infectious individuals with antivirals. In doing this, some antivirals are taken by individuals who would never contract the disease, but this wastage of antivirals did not stop the eradication of disease.

While the work contained in this thesis is largely focussed on the effort of a single country or area, there has been investigation using models that allow for spread through multiple countries, each of which has a supply of antivirals available [20]. It has been determined that a *cooperative* strategy, that is, a strategy where each country shares their supplies of antivirals, will out-perform the strategy where each country keeps their supplies to themselves with respect to controlling the overall level of infection. It is also noted that should a cooperative strategy be used, the severity of the pandemic decreases, even for mildly effective antivirals.

Work has been performed on the effects of antivirals in a population with household structure, but utilising the SEEIIR model, which is the same as the SEIR model but with an additional exposed and infected phase [15]. In the SEEIIR model, the

average duration an individual spends in the exposed class (which incorporates the two phases) and the infectious class remain unchanged, but the variance of the time spent in both the exposed class and the infectious class is reduced. This reduced variance is believe to be more physically accurate [29], at the cost of a more complex model. Note that the model used by Black *et al.* [15] has an assumption that once antivirals have begun being taken by individuals in a household, they remain being taken for the remainder of the early-growth phase of the pandemic. In Chapter 3, this assumption is investigated by incorporating a finite duration of antivirals.

1.2.2 Quantities of Interest for a Pandemic

There are many quantities that can be used to assess the severity of a pandemic. One of the most common quantities used in pandemic analysis is the basic reproductive ratio, R_0. The basic reproductive ratio is the expected number of secondary infections caused by one infected individual in an otherwise fully susceptible population [3, 63]. It is known that the basic reproductive ratio, R_0, is an *invasion threshold* for a pandemic. That is, the basic reproductive ratio, R_0, must be greater than 1 for there to potentially be a pandemic outbreak [45]. Estimates of the basic reproductive ratio, R_0, for the 2009 Swine 'Flu Pandemic range from 1.06 to 1.4 [29], while the the basic reproductive ratio, R_0, for the Spanish influenza, regarded as one of the worst pandemics in history, is estimated to be between 1.8 and 2 [57]. Closely related to the basic reproductive ratio, R_0, is the household reproductive ratio, R_*, first introduced by Ball *et al.* [11]. The household reproductive ratio is the expected number of households which are infected by all members of a single household in an otherwise fully susceptible population. The household reproductive ratio, R_*, is an invasion threshold for a pandemic in a population with household structure, much like the basic reproductive ratio, R_0, acts as an invasion threshold in a model without household structure. In a household level model, if the basic reproductive ratio, R_0, is greater than 1 there still may not be a pandemic outbreak

[25]. To see this, consider the case when there is no possibility of infecting an individual outside a given household, but infection could still occur inside a household. Then, the household reproductive ratio, R_0, could be larger than 1, but there will be no pandemic outbreak. Comparatively, if the household reproductive ratio, R_*, is greater than 1 then a pandemic outbreak is possible [7]. The household reproductive ratio, R_*, can be used in a similar way to the basic reproductive ratio, R_0, in order to classify the severity of a pandemic and also to determine whether there could be a pandemic outbreak in a population. There exists a variety of potential reproductive numbers that can be used to quantify the severity of an outbreak for a population with household structure. Goldstein *et al.* [35] detail five different household reproductive numbers, all of which are threshold parameters in the same way that R_* is a threshold parameter. Also of note is Theorem 1 of Goldstein *et al.* [35] which tells us that R_* is the largest of the commonly used household reproductive numbers. This means that R_* is a 'worst-case' estimate of severity, relative to other potential household reproductive numbers.

Another quantity that is used to quantify the severity of a pandemic is the *early growth rate*, known as the *Malthusian parameter*, r, which represents the exponential rate at which the expected number of infected people grows per unit time [45]. Obviously, a lower Malthusian parameter, r, implies a less severe pandemic as there are fewer infections happening per unit time. Unlike the basic reproductive ratio, R_0, and the household reproductive ratio, R_*, the Malthusian parameter, r, depends on time. Having the exposed phase in the model compared to the same model without the exposed phase will naturally give a lower Malthusian parameter, r, as people remain in the exposed class for a period of time and are not infecting any other susceptible individuals. The basic reproductive ratio, R_0 and the Malthusian parameter, r, are linked together through what is known as the *generation interval distribution*, which represents the probability of creating a new infection at a particular time point [56, 74].

Methods for calculating the household reproductive ratio, R_*, and the Malthu-

sian parameter, r, using sets of linear differential equations and systems of linear equations already exist [15]. A key assumption in the derivation of these methods is that all infections occur in an otherwise *naive*, or fully susceptible, population. In Chapter 3 we will extend this work in order to incorporate a finite duration for which antivirals are effective, and incorporate a number of other extensions, as well as discussing the effects of antivirals in more detail.

A third quantity used to classify pandemics is the *final epidemic size*, which is the total number of individuals who contracted the disease over the duration of the pandemic [17]. Clearly, the *final epidemic size* cannot be accurately attained while a pandemic is in progress, but final epidemic size can be a useful tool for comparing past pandemics and theoretical scenarios. Final epidemic size can be defined in terms of either the number of individuals infected or in terms of a proportion of a potential population. These forms are equivalent, with the proportion simply being the number of people infected divided by the total population size. There exists equations to calculate the final epidemic size in a population with household structure [6, 60, 66], and also the (mean) final epidemic size in a deterministic setting [46]. However, these methods are not able to be utilised directly with our model, due to features such as the finite amount of antivirals.

The final quantities of interest are the *peak size* and *peak time* of the pandemic, defined to be the maximum number of infected people over the course of the pandemic and the time at which this peak occurs. A lower peak size implies a less severe pandemic, while generally speaking a later peak time is desired, as this allows for more time to attempt control and prepare for the outbreak.

Comparing key quantities between a homogeneous mixing model and a model with population structure is not a trivial task. A household-structured model has one extra parameter compared to the homogeneous mixing model. To compensate for this, one key quantity must remain fixed in order to fairly compare the two models. House and Keeling [39] assessed three different modelling assumptions in order to match a household-structured model to a homogeneous mixing model. It was

determined that while all three modelling assumptions provided similar pandemic dynamics, choosing just one of these assumptions is not straight-forward. In this work, we do not attempt to compare the household-structured model to a homogeneous mixing model. Instead, we focus on the two different antiviral allocation schemes inside a household-structured model.

1.3 Thesis Outline

In Chapter 2, the technical background required for the rest of the work is presented, as well as an overview of some of the previous pandemic models which have been studied. The *stochastic households epidemic model* is introduced, but it is shown that this model cannot be used in practice, except by simulation which is slow. To gain insight into the stochastic households epidemic model, the first of two approximations, the *branching process* approximation, is introduced. The branching process approximation allows fast computation of some quantities associated with a pandemic, but cannot be used to approximate the long-term behaviour of the pandemic, due to a necessary assumption that the number of households in the population is infinite [9]. Recent work has shown that key quantities associated with a pandemic are calculable using a branching process [13]. However, the branching process discussed in Chapter 2 is not of the form used by Barbour and Reinert, but is instead focussed on fast computation of the early-time dynamics of the pandemic. In Chapter 3, the branching process approximation is extended to incorporate additional realism such as the finite duration of antivirals and having insufficient antivirals for the entire population. Results for these extensions are also discussed, but the shortfalls of the branching process approximation still remain. In this chapter, we also present a method for preallocating antivirals to the population, should the total number of doses not be sufficient for the population size. In Chapter 4, the *deterministic* approximation is introduced. The deterministic approximation can be used to estimate the mean behaviour of both the early

stages and the long-term behaviour of the pandemic, at the cost of a slightly longer computation time when compared to the branching process approximation. There are a number of extensions to the deterministic approximation which are considered: having insufficient antivirals for the entire population; a *hybrid* allocation scheme, that is, a scheme that is a mix of dynamic allocation and preallocation; households using antivirals incorrectly; and the production of antivirals during a pandemic. These extensions allow for the comparison of the two antiviral allocation schemes—dynamic allocation and preallocation—which are discussed in Chapter 5. The comparisons made in Chapter 5 are between pandemics under the two antiviral allocation schemes for both a *severe* pandemic, based on key quantities from the 1918 Spanish Influenza pandemic, and a *mild* pandemic, based on key quantities from the 2009 Swine 'Flu pandemic. Chapter 6 summarises the results and discusses the overall differences between dynamic allocation and preallocation. Also mentioned are some open questions and extensions to the approximations that are presented throughout this thesis.

Chapter 2

Technical Background

In this chapter a background is provided on continuous-time Markov chains as well as an introduction to the *stochastic households epidemic model*. We discuss the issues with using this model in practice, and introduce the first of two approximations—the branching process approximation. The work that has been done previously using this approximation is explored, including an efficient method to evaluate the household reproductive ratio, R_*, and the Malthusian parameter, r. Finally, the choices of parameters that are utilised throughout this thesis are defined and discussed.

2.1 SEIR Model

As discussed in Chapter 1, this work is focussed on the SEIR model. This model allocates each individual in a population into one of four distinct classes: susceptible, exposed, infectious or recovered. Each individual starts susceptible to the disease, then may become exposed if contact is made with an infectious individual. After some period of time, the individual becomes infectious and is able to infect other susceptible individuals. Finally, the individual recovers, and has no further impact on the pandemic. The SEIR model, shown in Figure 1.2, and other similar models such as the SIR model, are commonly modelled as a *Markov chain* [15, 41, 45].

2.2 Continuous-Time Markov Chains

A *continuous-time Markov chain* is utilised throughout this thesis. Consider a process, $X(t)$, which takes values on a finite set of states, S, known as the state space. The process $X(t)$ is a continuous-time Markov chain if,

$$P(X(t+s) = j|X(s) = i, X(u) = k,\ u \leq s) = P(X(t+s) = j|X(s) = i)\ \forall t, u > 0, s > u,$$

that is, the probability distribution of the future states of $X(t)$, conditioned on the past and present states, depends only upon the present state of the process. This is known as the *Markov* property, or the memoryless property. Let T_j be the amount of time until the Markov chain, $X(t)$, is in state $j \in S$, given that it is currently in state $j \in S$. The state j is known as *recurrent* if,

$$P(T_j < \infty) = 1,$$

that is, the Markov chain will almost surely return to state j. A state which is not recurrent is known as *transient*. A recurrent state $j \in S$ is known as *absorbing* if,

$$P(X(t+s) = j|X(s) = j) = 1\ \forall\, t, s > 0.$$

A continuous-time Markov chain is *time-homogeneous* if $P(X(t+s) = j|X(s) = i), \quad i, j \in S,\ s, t \in [0, \infty)$ is independent of s. In this thesis, all Markov chains are time-homogeneous. Let,

$$P_{ij}(t) = P(X(t+s) = j|X(s) = i).$$

The function $P_{ij}(t)$ is the *transition function*, giving the probability of moving from from state $i \in S$ to state $j \in S$ in elapsed time t. The transition function can be expanded as,

$$P_{ij}(t) = P(X(t+s) = j|X(s) = i) = \sum_{k \in S} P(X(t+s) = j, X(t+u) = k|X(s) = i), \tag{2.1}$$

using the law of total probability, where $0 < u < s$. Then, using conditional probability, the right-hand side of Equation (2.1) can be re-written as,

$$\sum_{k\in S} P(X(t+s) = j|X(s+u) = k, X(s) = i)P(X(s+u) = k|X(s) = i), \quad (2.2)$$

which, using the memoryless property and time-homogeneity becomes,

$$\begin{aligned}&\sum_{k\in S} P(X(t+s) = j|X(s+u) = k)P(X(u) = k|X(0) = i),\\ &\quad = \sum_{k\in S} P(X(t-u) = j|X(0) = k)P(X(u) = k|X(0) = i),\end{aligned}$$

that is,

$$P_{ij}(t) = \sum_{k\in S} P_{jk}(u)P_{kj}(t-u)$$

or, in matrix form,

$$\mathbb{P}(t) = \mathbb{P}(u)\mathbb{P}(t-u), \quad (2.3)$$

where $\mathbb{P}(t)$ has elements $P_{ij}(t)$. Equation (2.3) is known as the *Chapman-Kolmogorov* equation. Note that the matrix, $\mathbb{P}(t)$, has time-dependent entries. In order to gain insight into a continuous-time Markov chain, a method to analyse the matrices, $\mathbb{P}(t)$, is required. Define a matrix, Q, with elements,

$$q_{ij} = \lim_{h\to 0^+} \frac{P_{ij}(h)}{h},$$

for $i, j \in S,\ i \neq j$ and,

$$q_{ii} = \lim_{h\to 0^+} \frac{P_{ii}(h) - 1}{h},$$

for $i \in S$. The matrix Q is known as the *infinitesimal generator*, or the Q-matrix of the process $X(t)$. Intuitively, q_{ij} represents the rate at which we enter state j from state i, and $-q_{ii}$ represents the total rate of leaving state $i \in S$. As $P_{ij}(t)$ corresponds to the probability of moving from state i to state j in time t, we have that,

$$\sum_j P_{ij}(t) = 1,$$

for every $i \in S$, $t \in [0, \infty)$. As such,

$$1 - P_{ii}(h) = \sum_{j \neq i} P_{ij}(h),$$

and so,

$$\begin{aligned} \lim_{h \to 0^+} \frac{1 - P_{ii}(h)}{h} &= \lim_{h \to 0^+} \sum_{j \neq i} \frac{P_{ij}(h)}{h} \\ &= \sum_{j \neq i} \lim_{h \to 0^+} \frac{P_{ij}(h)}{h}. \end{aligned} \tag{2.4}$$

The left hand side of Equation (2.4) is precisely $-q_{ii}$, and the term inside the sum on the right hand side is q_{ij}, so Equation (2.4) is,

$$-q_{ii} = \sum_{j \neq i} q_{ij},$$

that is, the diagonal elements of the infinitesimal generator, Q, are the negative of the sum of the off-diagonal elements.

The infinitesimal generator, Q, can fully define the Markov chain when paired with a suitable initial condition, $p(0)$. To see this, consider the Chapman-Kolmogorov equation in Equation (2.3),

$$P_{ij}(t + h) = \sum_{k \in S} P_{ik}(h) P_{kj}(t).$$

We have,

$$\begin{aligned} \lim_{h \to 0^+} \frac{P_{ij}(t + h) - P_{ij}(t)}{h} &= \lim_{h \to 0^+} \frac{\left(\sum_{k \in S} P_{ik}(h) P_{kj}(t)\right) - P_{ij}(t)}{h} \\ &= \lim_{h \to 0^+} \frac{\sum_{k \in S} P_{ik}(h) P_{kj}(t)}{h} - \frac{P_{ij}(t)}{h}. \end{aligned} \tag{2.5}$$

Removing the $k = i$ term from the sum in Equation (2.5) yields,

$$\begin{aligned} \lim_{h \to 0^+} \frac{\sum_{k \in S} P_{ik}(h) P_{kj}(t)}{h} - \frac{P_{ij}(t)}{h} &= \lim_{h \to 0^+} \left[\sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) + \frac{P_{ii}(h) P_{ij}(t)}{h} - \frac{P_{ij}(t)}{h} \right] \\ &= \lim_{h \to 0^+} \left[\sum_{k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) - \frac{1 - P_{ii}(h)}{h} P_{ij}(t) \right]. \end{aligned} \tag{2.6}$$

Switching the summation and the limit in Equation (2.6) gives,

$$\begin{aligned}\frac{dP_{ij}(t)}{dt} &= \sum_{k\neq i} q_{ik}P_{kj}(t) + q_{ii}P_{ij}(t) \\ &= \sum_{k\in S} q_{ik}P_{kj}(t),\end{aligned}$$

or, in matrix form,

$$\frac{d\mathbb{P}(t)}{dt} = Q\mathbb{P}(t),$$

which has solution

$$\mathbb{P}(t) = e^{Qt}, \tag{2.7}$$

where e^{Qt} is the *matrix exponential* of the matrix Qt [61]. The distribution of the Markov chain at time t, that is, the probability of being in each state in the state space at time t is,

$$p(t) = p(0)\mathbb{P}(t) = p(0)e^{Qt},$$

for some known initial distribution, $p(0)$.

2.3 Matrix Exponential

The matrix exponential for a matrix, A, is defined as,

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!}A^k. \tag{2.8}$$

The series in Equation (2.8) converges for any matrix A, and so the matrix exponential is well defined. In this work, matrix exponentials are numerically evaluated using two methods. Evaluation of $\mathbf{v}e^{Ax}$ for some $\mathbf{v}$ and x is performed using EXPOKIT for MATLAB [70], which is regarded as more efficient and numerically stable when compared to the MATLAB in-built function `expm` [59], while e^A is evaluated using the MATLAB in-built function `expm`.

Let Q be the generator matrix of the Markov chain. A property which will be useful in Section 3.4 is,

$$\int_0^{\infty} e^{-t(rI-Q)}\,dt = (rI - Q)^{-1}, \tag{2.9}$$

for $r \geq 0$. A property of the matrix exponential is, if D is invertible,

$$\int e^{Dt}\, dt = D^{-1}e^{Dt} + C. \tag{2.10}$$

Let $D = -(rI - Q)$. The matrix D is irreducible, and has off-diagonal elements being non-negative, and so the matrix D is *ML* in the sense of Seneta [68]. Consider the definite integral,

$$\int_0^\infty e^{-t(rI-Q)}\, dt = \lim_{b\to\infty} \int_0^b e^{-t(rI-Q)}\, dt, \tag{2.11}$$

which, using Equation (2.10) gives,

$$\lim_{b\to\infty} \int_0^b e^{-t(rI-Q)}\, dt = -(rI-Q)^{-1} \lim_{b\to\infty} e^{-t(rI-Q)}\Big|_{t=0}^{b}. \tag{2.12}$$

At $t = 0$, $e^{-t(rI-Q)} = I$, the identity matrix, so Equation (2.12) can be rewritten as,

$$\lim_{b\to\infty} e^{-t(rI-Q)}\Big|_{t=0}^{b} = -(rI-Q)^{-1}\left(\lim_{b\to\infty} e^{-b(rI-Q)} - I\right). \tag{2.13}$$

As the matrix D is *ML*, Theorem 2.7 of Senata [68] says,

$$e^{Db} = e^{-b(rI-Q)} = e^{-b\tau}\mathbf{w}\mathbf{v}' + O(e^{-b\tau'}),$$

where τ is the dominant eigenvalue of $D = -(rI-Q)$ with left and right eigenvectors, $\mathbf{w}$ and $\mathbf{v}$ respectively, and $\tau' < \tau$. The dominant eigenvalue, τ, is non-positive provided r is non-negative, and so it follows that,

$$\lim_{b\to\infty} e^{-b(rI-Q)} = \lim_{b\to\infty} e^{b\tau}\mathbf{w}\mathbf{v}' + O(e^{b\tau'}) = 0. \tag{2.14}$$

Substituting Equation (2.14) into Equation (2.13), we see that,

$$\int_0^\infty e^{-t(rI-Q)}\, dt = (rI-Q)^{-1}, \tag{2.15}$$

for $r \geq 0$.

2.4 Path Integrals of Markov Chains

We follow the work of Ross *et al.* [66] who use a Markov chain model and then derive equations for the household reproductive ratio, R_*, and the Malthusian parameter, r, for a pandemic. Let $X(t)$ be an absorbing continuous-time Markov chain which has a finite state space S. Consider partitioning the state space, S, into two distinct subsets such that $S = A \cup C$, where A is the set of absorbing states and C is the set of transient states. Consider a function $f : S \to [0, \infty)$ with the property $f(k) = 0$ for all $k \in A$. The function f represents a per unit-time reward associated with each state k. Consider the path integral,

$$\Gamma = \int_0^\infty f(X(t))\, dt, \tag{2.16}$$

which represents the total reward obtained over the lifetime of the process. As the Markov process, $X(t)$, is a random variable, it follows that Γ is also a random variable.

Pollett and Stefanov [65] demonstrate that the conditional expectation of Γ can be found using a system of linear equations: Let $\mathbb{E}\left[\Gamma | X(0) = j\right] = \nu_j, j \in S$; then, ν_j is the minimal non-negative solution to

$$\sum_{j \in C} q_{ij}\nu_j + f(i) = 0, \quad \forall\, i \in C. \tag{2.17}$$

Equation (2.17) is a system of linear equations, and can be solved using MATLAB's 'backslash' operator.

Similarly, consider the function

$$\bar{\Gamma}(c) = \int_0^c f(X(t))\, dt. \tag{2.18}$$

Equation (2.18) represents the total reward of a process over a finite time, $c \in \mathbb{R}^+$. The conditional expectation of $\bar{\Gamma}(c)$ can be found, given any initial state [15]. Let $\mathbb{E}\left[\bar{\Gamma}(t) | X(0) = j\right] = \xi_j(t)\ \forall c \in C$. Then, $\xi_j(t)$ satisfies

$$\frac{d\xi_j(t)}{dt} = \sum_{l \in C} q_{jl}\xi_l(t) + f(j), \quad \forall\, l \in C, \tag{2.19}$$

with initial condition $\xi(0) = \xi_0$, which represents the initial 'reward'. We take $\xi_0 = 0$, as at time 0 no reward has been earned. We numerically solve the system of ordinary differential equations in Equation (2.19), giving the expectation of Equation (2.18) from any initial state. This system of linear differential equations can be solved efficiently using EXPOKIT for MATLAB, by using the function `phiv` [70]. This method is generally faster than the Runge-Kutta methods utilised by MATLAB's inbuilt differential equation solvers such as `ode45`.

Throughout this thesis, expressions of the form,

$$\mathbb{E}\left[\int_0^b f(X(t))\,dt\right], \tag{2.20}$$

represent the *vector* of conditional expectations, with the jth component of Equation (2.20) being,

$$\mathbb{E}\left[\int_0^b f(X(t))\,dt \,\middle|\, X(0) = j\right].$$

This slightly unorthodox notation allows for a clean representation of path integrals from any initial state. In particular, if $p(0)$ is the initial distribution of the process $X(t)$, then,

$$p(0)\mathbb{E}\left[\int_0^b f(X(t))\,dt\right],$$

represents the expected value of the path integral until time b. Often, we assume a fixed initial state, meaning that,

$$p(0) = \mathbf{u}_j,$$

where $\mathbf{u}_j$ is a vector that has value 1 in the jth element, and 0 otherwise. Then,

$$\mathbf{u}_j\mathbb{E}\left[\int_0^b f(X(t))\,dt\right] = \mathbb{E}\left[\int_0^b f(X(t))\,dt \,\middle|\, X(0) = j\right].$$

2.5 Stochastic Households Model

The effects of influenza on a population of individuals with household structure can be modelled as a Markov chain, $X(t)$. Consider a population which consists of

N households each of size k, giving a total population size of Nk. We say that a household in the population is in state $Z_j = (s, e, i)$ if the household has s susceptible individuals, e exposed individuals and i infectious individuals. Note that as the household size is fixed, the number of recovered individuals, $r = k - (s + e + i)$. There are three possible events which can occur inside a single household in the population: infection, progression and recovery, which are detailed in Table 2.1. Note that throughout this work we assume *frequency-dependent transmission*, that is,

$$\beta_k = \begin{cases} \frac{\beta}{k-1} & \text{if } k > 1, \\ 0 & \text{if } k = 1 \end{cases},$$

for some governing rate of (internal) infection, β. We assume frequency-dependent transmission as the probability of being infected on contact with an infectious individual is (relatively) independent of population density inside a household [15, 45, 66]. Considering the infection rate more precisely, the probability of the contact between a susceptible individual and an infectious individual is $\frac{i}{k-1}$, and there are s such susceptible individuals with infection being successfully transmitted at rate β, and so the total internal infection rate is $\frac{\beta}{k-1}si = \beta_k si$ [66]. Should $k = 1$, then there is no possible internal infection and so we set $\beta_1 = 0$.

As the population has household structure, we are able to incorporate *two levels of mixing*. This allows the stochastic households model to have a different internal infection rate and external infection rate, which means that the model can incorporate the estimated increase in infection transmission inside a household compared to infection transmission in the general population [15, 27]. Denote the governing rate at which an individual infects a member of the population not in their own household by α. The total *force of infection* out of a single household in state (s, e, i) is αi, and so the total force of infection into the general population is,

$$\frac{\alpha}{Nk} \sum_{j=1}^{N} I(Z_j),$$

where $I(Z_j) = i$. Note here that we have again assumed frequency-dependent transmission of infection, just as for infection inside a household.

Using these transmission rates, the Markov chain for a single household, $X_k(t)$, can be constructed. The generator for $X_k(t)$, denoted $Q^{(k)}$, has elements,

$$\begin{aligned} q^{(k)}_{(s,e,i),(s-1,e+1,i)} &= \beta_k si + \frac{\alpha}{Nk} s \sum_{j=1}^{N} I(Z_j) \\ q^{(k)}_{(s,e,i),(s,e-1,i+1)} &= \sigma e \\ q^{(k)}_{(s,e,i),(s,e,i-1)} &= \gamma i \\ q^{(k)}_{(s,e,i),(s,e,i)} &= - \sum_{j \in S \setminus \{(s,e,i)\}} q^{(k)}_{(s,e,i),j}. \end{aligned} \tag{2.21}$$

The state space for the Markov chain representing a single household, $X_k(t)$, is,

$$S_k = \{(s,e,i) | s + e + i \leq k \; s,e,i \geq 0\}.$$

To determine the number of possible states for a household, consider allocating the k individuals in a household into three distinct classes, with classes able to have zero members. Hence, the number of possible states for a household is

$$|S_k| = \binom{k+3}{3} = \frac{1}{6}\frac{(k+3)!}{k!} = \frac{1}{6}(k+1)(k+2)(k+3).$$

In this representation of the stochastic households model, we have assumed that individuals inside the household are indistinguishable, but each household is able to be individually identified. Because of this, the process $\mathbf{Z}(t)$, contains the status of the N unique households, each with $|S_k|$ possible states, and so the size of the state space for the stochastic households model is,

$$|S| = |S_k|^N = \left(\frac{1}{6}(k+1)(k+2)(k+3)\right)^N.$$

For any reasonable N, then, the size of the state space for the stochastic population model is impractically large. The stochastic households model can still be simulated however, by using Monte-Carlo simulation.

Event	Transition	Rate
Infection	$(s, e, i) \to (s-1, e+1, i)$	$\beta_k si$
Progression	$(s, e, i) \to (s, e-1, i+1)$	σe
Recovery	$(s, e, i) \to (s, e, i-1)$	γi

Table 2.1: Possible events inside a single household during a pandemic.

2.5.1 Monte Carlo Simulation

While the state space for the stochastic households model means that the model cannot be analytically evaluated, simulation of the Markov chain is still possible using Monte Carlo methods, commonly known as the *Gillespie Algorithm* [32]. This is summarised in Algorithm 1.

Each simulation returns a single *realisation* or *sample path* from the stochastic households model, sampled at a set of defined time points. When comparing different pandemic scenarios, the interest is in the general behaviour of the model rather than the specifics of each realisation. Because of this, a large number of these realisations are obtained, and the corresponding mean, $\mathbb{E}[X(t)]$, is calculated and used to calculate quantities associated with the pandemic.

While the stochastic households model is detailed, Monte Carlo Simulation takes approximately 30 minutes to produce one pandemic curve using the *severe* parameters in Table 2.3, using an Intel Xeon 2.6 GHz processor. This means that simulation of the stochastic households model is impractical for comparisons across many pandemic scenarios. We investigate two approximations to this stochastic household model—a branching process approximation and later, a deterministic approximation—which are faster to compute and are shown to accurately reproduce the key quantities associated with a pandemic.

Initialize state, $\mathbf{Z}(0)$, containing the initial configuration of each household $Z_j(t) = (s, e, i) \in S,\ j = 1, \ldots, N$. Set $c = 1$, $t = 0$;

Initialize vector of time points to sample the process, $\hat{\mathbf{t}}$, with mth element $\hat{t}_m$.

Let the number of exposed individuals in the current state $\mathbf{Z}(t)$ be,

$$\hat{E}(\mathbf{Z}(t)) = \sum_{Z_j(t)=(s,e,i)} e,$$

and let the number of infectious individuals in the state $\mathbf{Z}(t)$ be,

$$\hat{I}(\mathbf{Z}(t)) = \sum_{Z_j(t)=(s,e,i)} i.$$

while $\left(\hat{E}(\mathbf{Z}(t)) + \hat{I}(\mathbf{Z}(t))\right) > 0$ **do**

Calculate rate of each event, $a_l, l = 1, \ldots, A$, where A is the number of possible events. Set $a_0 = \sum_{l=1}^{A} a_l$;

Generate two random variables, $r_1, r_2 \sim U[0, 1]$;

Set $\Delta = (1/a_0)\ln(1/r_1)$;

while $t + \Delta > \hat{t}_c$ **do**

Record $I_c = \sum_{Z_j(t)} i$, the number of infectious individuals at time t;

Set $c = c + 1$;

end

Find $\mu \in \{1, \ldots, A\}$ such that,

$$\sum_{j=1}^{\mu-1} a_j < r_2 a_0 \leq \sum_{j=1}^{\mu} a_j.$$

This is equivalent to choosing the jth reaction with probability a_j/a_0;

Update state $\mathbf{Z}(t)$ according to event type μ, set $t = t + \Delta$;

end

Output: The number of infectious individuals, $\mathbf{I}$, with jth component I_j, at each sampled time point in $\hat{\mathbf{t}}$.

Algorithm 1: Simulation Algorithm for the stochastic households model.

2.6 Branching Process Approximation

The first approximation to the stochastic households model is a *branching process* approximation. A branching process is a stochastic process in which an individual has a number of offspring over their lifetime, and each offspring goes on to have their own independent, and usually identically distributed lifetime. The offspring lifetimes need not be identically distributed in general, but making this assumption makes the model more tractable.

More precisely, consider an individual who lives for a period of time which is exponentially distributed, known as their lifetime. Over this lifetime, a number of offspring are created, the number of which is governed by the so-called *offspring distribution*. Each offspring lives for a period of time which is independent and identically distributed to that of the parent. Each offspring also creates a number of new offspring, usually controlled by the same offspring distribution as the parent, and so the process continues. Clearly with the assumption of each parent having the same offspring distribution, the total number of offspring produced is unbounded. If the mean number of offspring is greater than one, then the branching process is known as *supercritical* while if the mean number of offspring is less than one, then the branching process is known as *subcritical*, and if the mean number of offspring is exactly one, then the branching process is known as *critical*. In this work, only *supercritical* branching processes are considered, as the other two cases often produce 'trivial' processes—that is, branching processes where the process becomes extinct quickly. In our context, the process becoming extinct means that the pandemic does not take off inside the population, and so control schemes for the pandemic are not relevant.

It is possible to think of the early stages of a pandemic as a branching process. A household has infectious individuals for some amount of time, and over that time some number of other individuals in the population will become infected. Those new individuals will be infectious for an independent and identically distributed amount

of time and will again infect some number of other individuals in the population. For a finite population, however, the number of new infections created in the general population by all members of a single household cannot have the same distribution through time simply because eventually there will be insufficient susceptible people to infect. As such, an infinite population size must be assumed, and so a branching process approximation is valid only for the early stages of a pandemic. The results of Ball and Donnelly [9] demonstrates that a branching process approximation converges almost surely to the stochastic households model of the pandemic as the population size becomes large. In particular, during a *minor outbreak*, that is, a pandemic with household reproductive ratio, R_*, less than or equal to one, the pandemic behaves like a branching process when the population size, N, is large, while during a *major outbreak*, that is, a pandemic with household reproductive ratio, R_*, being strictly greater than one, the branching process is valid until approximately $\sqrt{N}$ members of the population have been infected [10]. Quantities associated with a pandemic from the branching process can be calculated using *path integral* methods.

2.6.1 Calculating Key Quantities from the Branching Process Approximation

There have been studies which give methods to calculate the early-time quantities of a pandemic using the path integral of a Markov chain [66, 15]. These methods are detailed here for our purposes. Let the Markov chain for a household of size k be $X_k(t)$. Consider the function $I(X_k(t)) = i$ where $X_k(t) = (s, e, i)$ is the state of the Markov chain at time t. The function $I(X_k(t))$ gives the number of infectious individuals in a household at time t. The set of absorbing states for $X_k(t)$ is,

$$A = \{(s, 0, 0) | s \leq k\}.$$

For all $i \in A$, $I(i) = 0$ as there are no infectious people in each absorbing state of the Markov chain and so $I(\cdot)$ satisfies the requirements of the reward function, $f(X(t))$, in Equations (2.17) and (2.19).

Recall from Section 2.5 that the governing rate of external infection is α. From a single household, the force of infection at time t is $\alpha I(X_k(t))$. As external infections are transmitted into a naive household due to the assumption of an infinite population size, external infection events follow a *Poisson process* with (time-dependent) rate, $\alpha I(X_k(t))$. The household reproductive ratio, R_*, is the expected number of secondary households infected, that is,

$$R_* = p(0)\mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\, dt\right]. \tag{2.22}$$

We can evaluate this integral conditioned on any initial state using Equation (2.17).

To determine the Malthusian parameter, r, we follow the workings of Wallinga and Lipsitch [74]. The Euler-Lotka equation for modelling population growth has the form

$$b(t) = \int_0^\infty b(t-a)n(a)\, da, \tag{2.23}$$

where $b(t)$ here represents the total expected number of new infections into fully susceptible households in time period t and $n(a)$ the average rate at which new (external) infections occur from a household at "age" a. Here, the 'age' of a household refers to the time since the first infection event inside the household. As external infections occur according to a Poisson process, the number of new infections into fully susceptible households is growing at an exponential rate per unit time. That is,

$$b(t) = b(t-a)e^{ra}.$$

Substituting this into Equation (2.23) yields,

$$b(t) = \int_0^\infty b(t)e^{-ra}n(a)\, da,$$

or

$$1 = \int_0^\infty e^{-ra}n(a)\, da. \tag{2.24}$$

Also, as $n(a)$ represents the average rate at which external infections occur, it follows that $n(a) = \mathbb{E}\left[\alpha I(X_k(a))\right]$. Substituting this yields the equation for the Malthusian parameter, r:

$$1 = \mathbb{E}\left[\int_0^\infty \alpha e^{-rt} I(X_k(t))\, dt\right]. \tag{2.25}$$

Note here that we have moved the expectation outside the integral. This is only to allow for more efficient numerical evaluation. This equation has been derived previously by Ball [7] and utilised in pandemic analysis [66, 15].

As presented in Norris [61], Equation (2.25) is a path integral with *exponential discounting* by a factor, $r > 0$. This is equivalent to using the original process and adding an additional rate, r, from each $i \in C$ to the absorbing class A. Using this modified process, $\bar{X}_k^r(t)$ for a given value of r, Equation (2.25) becomes

$$1 = \mathbb{E}\left[\int_0^\infty \alpha I(\bar{X}_k^r(t))\, dt\right]. \tag{2.26}$$

The right hand side of this expression can be evaluated for a given value of r using the system of linear equations in Equation (2.17). The value of r for which equality holds in Equation (2.26) can be determined by using a numerical root-finding algorithm such as MATLAB's `fzero` function.

Both Equations (2.22) and (2.25) will be extended throughout this work in order to incorporate more complexity and realism into the model.

2.7 Incorporating Heterogeneous Household Sizes

Until now we have considered only a population made up entirely of households of a single, fixed size, but it is possible to relax this assumption; including heterogeneous household sizes in the stochastic households model is straightforward. Simply initialize the state, $\mathbf{Z}(0)$, to contain the required number of households of each size, then simulate the Markov chain as in Section 2.5.1.

To incorporate heterogeneous household sizes in the branching process approximation, denote the probability that a randomly selected household in a population is of size k by h_k. Clearly, $\sum_k h_k = 1$. We utilise the *size-biased distribution*, $\boldsymbol{\pi}$, which

represents the probability that a randomly selected individual in the population belongs to a household of size k. The size-biased distribution, $\boldsymbol{\pi}$, allows incorporation of the fact that transmission occurs between random individuals in the population. The jth element of the size-biased distribution is,

$$\pi_j = \frac{jh_j}{\sum_k kh_k}.$$

When considering a population with heterogeneous household sizes, the probability of contact of an infectious individual with a susceptible individual who belongs to a household of size k is π_k. The expected number of secondary infections from such a household is,

$$\pi_k \mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\, dt\right].$$

Hence,

$$R_* = \sum_k p_k(0)\pi_k \mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\, dt\right] \tag{2.27}$$

where $X_k(t)$ is the Markov chain for a household of size k and $p_k(0) = \mathbf{u}_{(k-1,1,0)}$. Similarly the Malthusian parameter r, can be found using

$$\sum_k p_k(0)\pi_k \mathbb{E}\left[\int_0^\infty \alpha e^{-rt} I(X_k(t))\, dt\right] = 1. \tag{2.28}$$

Each individual path integral in both Equations (2.27) and (2.28) can be evaluated using Equations (2.17) and (2.19) respectively [15].

Note that if we take a single household size k, then $\pi_k = 1$ while all other elements of $\boldsymbol{\pi}$ are zero, and we arrive back at Equations (2.22) and (2.25).

2.8 Modelling Antiviral Intervention

Recall that the Australia Health Management Plan for Pandemic Influenza states that antivirals will be allocated to an entire household as soon as there is one confirmed infectious individual in the household. The branching process approximation can be extended to incorporate the effects of antivirals. The dynamic allocation

scheme, introduced in Section 1.1, has been incorporated into the branching process approximation previously [15].

Recall that antivirals have two effects on individuals—a reduction in susceptibility, denoted $\rho \in (0,1)$, and a reduction in infectivity, denoted $\tau \in (0,1)$. When under the effects of antivirals, the rate of infection inside a household, β_k, is reduced by both ρ and τ, as all members of the household are taking antivirals. The rate of infection to other households in the population, α, is reduced by τ, as the infectivity of individuals inside the household is reduced while the individuals are taking antivirals.

There are two possible statuses of antivirals inside a household, a: those being when a household has not yet received antivirals, $a = 0$; and where individuals inside a household are currently taking antivirals, $a = 1$. Let the state $(s, e, i, 0)$ represent a household which has not yet received antivirals and state $(s, e, i, 1)$ represent a household that is currently taking antivirals. At this point, we assume that antivirals will have an essentially infinite effective duration. This assumption is relaxed in Chapter 3. The time between an individual becoming infectious and receiving a course of antivirals, known as the *delay*, must also be modelled. The delay until antivirals arrive is assumed to be exponentially distributed with mean $\zeta > 0$. Because of the exponential distribution, antivirals arriving into a household can be modelled as a transition within the Markov chain. Let the Markov chain which models a household during a pandemic where antivirals may be allocated be $X_k(t)$. The transitions of this process are described in Table 2.2. The size of the state space for the Markov chain, $X_k(t)$, is

$$|S_k| = \frac{1}{3}(k+1)(k+2)(k+3),$$

which has doubled compared to the case without antivirals.

An alternative way to view the generator matrix for $X_k(t)$ is to construct two matrices, Q_1 and Q_2, where Q_1 is the generator matrix with $\tau = \rho = 0$, and Q_2 the generator matrix with τ and ρ potentially non-zero. Finally, construct a 'linking'

Event	Transition	Rate
Infection	$(s, e, i, a) \to (s-1, e+1, i, a)$	$(1-a\tau)(1-a\rho)\beta_k si$
Progression	$(s, e, i, a) \to (s, e-1, i+1, a)$	σe
Recovery	$(s, e, i, a) \to (s, e, i-1, a)$	γi
Antivirals begin	$(s, e, i, 0) \to (s, e, i, 1)$	$\frac{1}{\zeta}$ for $s + e \neq k$

Table 2.2: Possible events inside a single household during a pandemic with antivirals. Here, $a = 0$ corresponds to antivirals not being active and $a = 1$ to antivirals currently being active within a household.

matrix L which encodes the transitions from state $(s, e, i, 0)$ to $(s, e, i, 1)$ at rate $1/\zeta$. Then the generator matrix, Q, for the Markov chain $X_k(t)$ has the block form

$$Q = \begin{bmatrix} Q_1 & L \\ 0 & Q_2 \end{bmatrix}.$$

This generator matrix, Q, can be used in Equations (2.22) and (2.25), taking

$$f(X_k(t)) = \begin{cases} \alpha I(X_k(t)), & \text{when } a = 0, \\ (1-\tau)\alpha I(X_k(t)) & \text{when } a = 1, \end{cases}$$

where $I(X_k(t)) = i$ is the number of infectious individuals in a household at time t. The first case above represents a household which is not taking antivirals and so receives none of the reductions in infectiousness. The second case refers to a household which is currently under the effects of antivirals and so the rate of infection from this household to individuals in any other household is reduced. The effectiveness of antivirals is clearly dependent on the delay [15, 71]. Should the delay between the first infectious case and the antivirals arriving into the household be large, then the transmission of infection occurs without any intervention, and so the antivirals will have little impact.

2.8.1 Incorporating Constant Time Delay until Antivirals Arrive

Having the delay until antivirals arrive into a household being of constant duration, instead of exponentially distributed, has been analysed previously [15, 71]. In a Markov chain, the holding time in each state is exponential, and so antiviral arrival can no longer be modelled as a transition in the Markov chain. Instead, consider splitting the pandemic into three intervals:

1. $[0, T_1)$—the interval from the time at which the first individual is exposed until the first individual becomes infectious in a household,

2. $[T_1, T_1 + T_a)$—the interval of time between the first infectious case and the time when antivirals arrive into the household, and

3. $[T_1 + T_a, \infty)$—the remainder of the pandemic.

Consider having two generator matrices, Q_1 and Q_2, which describe the dynamics of the pandemic in a household that has not yet received antivirals, and one that has, respectively. Then, from time $t = 0$ until the time at which antivirals arrive into the household, $t = T_1 + T_a$, the dynamics evolve according to Q_1. After time $t = T_1 + T_a$, the dynamics evolve according to Q_2 and continue in this way for the remainder of the pandemic.

Consider first the expected number of secondary households which are infected. We consider a single path integral from Equation (2.27),

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right], \tag{2.29}$$

with $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$. As there is no infection in the time interval $[0, T_1)$, this interval can be removed from Equation (2.22), and the initial condition changed to

$\bar{p}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$. As such, Equation (2.29) is

$$\begin{aligned} p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right] &= \bar{p}(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(t))\,dt + \int_{T_a}^\infty (1-\tau)\alpha I(X_k(t))\,dt\right] \\ &= \bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(t))\,dt\right] \\ &\quad + \bar{p}_k(T_a)\mathbb{E}\left[\int_0^\infty (1-\tau)\alpha I(X_k(t))\,dt\right] \end{aligned} \tag{2.30}$$

using the Markov property and that $\bar{p}_k(T_a) = \bar{p}(0)e^{Q_1 T_a}$ is the distribution of the Markov chain at time T_a, restricted to the set of transient states. As T_a is a constant, the first term of Equation (2.30) is a path integral over a constant time and can be evaluated using Equation (2.19), and the second is a path integral over infinite time and so can be evaluated using Equation (2.17).

Now, consider the Malthusian parameter, r. Again, we expand a single path integral from the summation in Equation (2.28), that being,

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))e^{-rt}I(X_k(t))\,dt\right] = 1,$$

with $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$. Here the first time interval cannot be ignored as was the case when considering the household reproductive ratio, R_*, due to the time dependence of the Malthusian parameter, r. We divide the expectation up into three terms,

$$\begin{aligned} p_k(0)\mathbb{E}\Bigg[&\int_0^{T_1} \alpha I(X_k(t))e^{-rt}\,dt + \int_{T_1}^{T_1+T_a} \alpha I(X_k(t))e^{-rt}\,dt + \ldots \\ &\int_{T_1+T_a}^\infty (1-\tau)\alpha I(X_k(t))e^{-rt}\,dt\Bigg] = 1 \end{aligned} \tag{2.31}$$

The first term here is zero as there are no new infections in the time interval $[0, T_1)$. Let $s = t - T_1$, and $ds = dt$. Substituting this time-shift into Equation (2.31) yields,

$$p_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s+T_1))e^{-r(s+T_1)}\,ds + \int_{T_a}^\infty (1-\tau)\alpha I(X_k(s+T_1))e^{-r(s+T_1)}\,ds\right] = 1.$$

As the state of the chain is known at time T_1, we can change the initial condition of the Markov chain to $\bar{\mathbf{p}}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$ and remove the dependence on T_1. Note

that the time until first infection in a household is exponentially distributed with mean $1/\sigma$, and is independent of the subsequent dynamics. Because of this, the dependence on T_1 can be integrated out giving,

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s+T_1))e^{-r(s+T_1)}\,ds + \int_{T_a}^{\infty}(1-\tau)\alpha I(X_k(s+T_1))e^{-r(s+T_1)}\,ds\right] =$$
$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a}^{\infty}(1-\tau)\alpha I(X_k(s))e^{-rs}\,ds\right]\int_0^{\infty}\sigma e^{-(r+\sigma)s}\,ds = 1. \tag{2.32}$$

The final integral can be evaluated as,

$$\int_0^{\infty}\sigma e^{-(r+\sigma)s}\,ds = \frac{\sigma}{r+\sigma},$$

and moving this to the right hand side of Equation (2.32) gives,

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a}^{\infty}(1-\tau)\alpha I(X_k(s))e^{-rs}\,ds\right] = \frac{r+\sigma}{\sigma}. \tag{2.33}$$

Using the memoryless property, Equation (2.33) can be simplified to,

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\,ds\right] + \bar{p}_k(T_a)e^{-rT_a}\mathbb{E}\left[\int_0^{\infty}(1-\tau)\alpha I(X_k(s))e^{-rs}\,ds\right] = \frac{r+\sigma}{\sigma}. \tag{2.34}$$

Each term in Equation (2.34) can be evaluated using the techniques in Section 2.4, and $\bar{p}_k(T_a) = \bar{p}_k(0)e^{QT_a}$ can be evaluated numerically. We combine these equations with a numerical root finder, such as MATLAB's `fzero`.

In Chapter 3, Equations (2.22), (2.25), (2.30) and (2.34) are extended to include a finite duration of antivirals as well as modelling the alternative preallocation antiviral scheme.

2.9 Choice of Parameters

Two sets of parameters, one representing a *severe* pandemic and another representing a less severe or *mild* outbreak, are used to compare the dynamic allocation scheme to the preallocation scheme in this work. Recall from Section 2.6.1 that

the severity of an outbreak is classified based on a number of quantities, primary of which is the household reproductive ratio, R_*. We can calculate the household reproductive ratio, R_*, using the methods in Section 2.6.1 for a given set of parameters. Recall also that the Spanish Influenza pandemic was one of the worst influenza pandemics in history and is estimated to have a household reproductive ratio, R_*, of 1.8 [57], while the 2009 Swine 'Flu pandemic was far less severe and had an estimated household reproductive ratio, R_*, of 1.3 [29]. The parameters on which the comparisons in this work are carried out are chosen to give pandemics which have both of these reproductive ratios.

While not necessary, we fix $\gamma = 1$, which scales time. As a result of this time scaling, one *time unit* resembles one infectious period. The purpose of this time scaling is to reduce the number of free parameters in the model, and to avoid issues with varying estimates between strains of influenza [29, 19]. For additional realism, we set the distribution of household sizes, $\mathbf{h}$, to resemble that of Australia [4]. That is,

$$\mathbf{h} = [0.2434, 0.3397, 0.1598, 0.1569, 0.0675, 0.0231, 0.0058, 0.0039],$$

where the mean household size here is,

$$\bar{h} = 2.577.$$

This gives the *size-biased distribution*,

$$\boldsymbol{\pi} = [0.0944, 0.2636, 0.1860, 0.2435, 0.1310, 0.0537, 0.0158, 0.0120].$$

Here, there is an assumption that there are no households larger than 8 individuals. In reality, this is obviously not true. However, the number of households in Australia which have more than 8 members is not recorded by the Census, and so the number of households with more than 8 members is assumed to be negligible. These two distributions are shown in Figure 2.1. Investigation of household size distributions of Indonesia and Sudan are later investigated in order to assess some sensitivity to the distribution of household sizes, $\mathbf{h}$. The population size is fixed at 10^5 individuals,

meaning that,

$$N = \left[\frac{10^5}{\bar{h}}\right] = 38803,$$

where $[x]$ is the nearest integer to x.

Define the *mild* parameter set to have parameters $\beta = 0.9669$, $\sigma = 1$, and $\alpha = 0.8$ and the *severe* parameter set to have parameters $\beta = 1.1259$, $\sigma = 1$ and $\alpha = 1$. The mild parameter set gives a household reproductive ratio, R_*, of 1.3, approximately matching that of the 2009 Swine 'Flu pandemic [29], while the severe parameter set gives a household reproductive ratio, R_*, of 1.8, approximately matching that of the Spanish Influenza pandemic [57]. Note that there exist other combinations of parameters which give the same household reproductive ratio, R_*, as above, however, we do not consider any other combination in this work.

The average delay until antivirals arrive into a household, ζ, has not been explored in detail previously, however, an estimate of the average delay in the United Kingdom in the 2009 Swine 'Flu pandemic was approximately 1 infectious period implying $\zeta = 1$ [29]. This is clearly very high—on average an individual would have recovered before the individuals in the household receive the antivirals. We take a small delay relative to the United Kingdom estimate, in order to assess the effectiveness of dynamic allocation if significant effort was given to ensuring rapid distribution of antivirals, and set $\zeta = 0.5$, which for influenza is approximately 1.5 days. The effectiveness of antivirals is varied, with some estimates claiming a 60% reduction in susceptibility [27, 53], however, these figures have been questioned previously [43]. We take a more conservative estimate of antiviral effectiveness, and set the reduction in infectivity, τ, and susceptibility ρ, to 0.3, in line with experimental estimates [72, 36]. The average antiviral duration, κ, which is discussed in detail in Chapter 3, is set to be 1 infectious period. All parameters for both the *mild* and *severe* parameter set are listed for convenience in Table 2.3. Throughout this thesis, the units on parameter values are suppressed for clarity.

During the 2009 Swine 'Flu pandemic, Australia stockpiled enough antivirals to allocate to 41% of the population [30] but the stockpiles of other countries varied

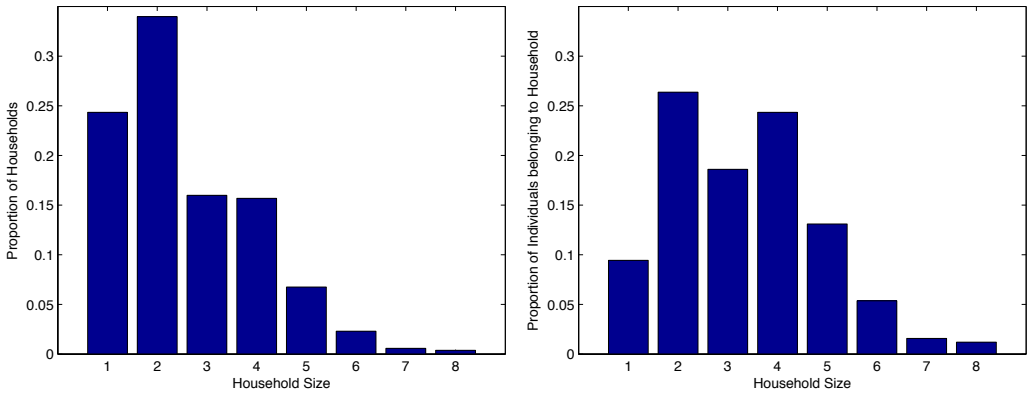
	Mild	Severe
β	0.9669 (γ^{-1})	1.1259 (γ^{-1})
α	0.8 (γ^{-1})	1 (γ^{-1})
	Both	
σ	1 (γ^{-1})	
ρ	0.3	
τ	0.3	
ζ	0.5 (γ)	
κ	1 (γ)	
γ	1	
$\mathbf{h}$	Australian Census, 2011	

Table 2.3: Definition of the *mild* and *severe* parameter sets.

by a large amount. Information about the size of the current antiviral stockpile available for use during an influenza pandemic in Australia, however is not available, and so the maximum number of available antivirals, M, cannot be set. As both approximations considered in this work are fast to compute, a wide range of available antivirals can be tested to alleviate this issue.

2.10 Summary

In this chapter, the background information about the Markov chain, the stochastic households model and the *branching process* approximation has been introduced. We demonstrated that the stochastic households model is impractical to use analytically, and can only be explored using simulation. In the branching process approximation, two key assumptions have been detailed. The first assumption is that infection happens into a *naive*, or fully susceptible population. This assumption is a property of a branching process, and cannot be removed. The second assumption



(a) Household Size Distribution, $\mathbf{h}$. (b) Size Biased Distribution, $\boldsymbol{\pi}$.

Figure 2.1: Household size distribution and the size biased distribution of Australian households.

is that antivirals have an essentially infinite duration. In Chapter 3, the branching process approximation is extended in order to remove this assumption, as well as to consider a number of other extensions. Finally, the *mild* and *severe* parameter sets have been defined. In Chapter 3, the branching process is extended to include the preallocation antiviral scheme, and some preliminary comparisons between the two antiviral allocation schemes are made using both the *mild* parameter set and the *severe* parameter set.

Chapter 3

Extensions to the Branching Process Approximation

In this chapter, the branching process approximation which was introduced in Chapter 2 is extended to allow for a finite antiviral duration, as well as to model the preallocation scheme. We also consider modelling the delay until antivirals arrive into a household and the effective duration of antivirals being of constant duration, as opposed to being exponentially distributed. Finally, preliminary results from the branching process approximation are presented.

3.1 Finite Antiviral Duration

The branching process approximation presented in Section 2.8 assumed that antivirals had an *effective duration* which was essentially infinite. In this section, the branching process approximation is extended to incorporate the finite duration of antiviral effectiveness. To facilitate this, the state space again requires extension. As before, let state $(s, e, i, 0)$ represent a household which contains individuals who are not yet taking antivirals, and state $(s, e, i, 1)$ represent a household which contains individuals who are currently taking antivirals. The antiviral status of a household, a, can now take on a new value, $a = 2$, which represents a household which

Event	Transition	Rate
Infection	$(s,e,i,a) \to (s-1,e+1,i,a)$	$\beta_k si$ for $a \neq 1$,
		$(1-\tau)(1-\rho)\beta_k si$ for $a=1$,
Progression	$(s,e,i,a) \to (s,e-1,i+1,a)$	σe
Recovery	$(s,e,i,a) \to (s,e,i-1,a)$	γi
Antivirals begin	$(s,e,i,0) \to (s,e,i,1)$	$\frac{1}{\zeta}$ for $s+e \neq k$
Antivirals end	$(s,e,i,1) \to (s,e,i,2)$	$\frac{1}{\kappa}$

Table 3.1: Possible events inside a household during a pandemic with antivirals under a dynamic allocation scheme.

contains individuals who have completed their course of antivirals. To begin with, assume that this *effective duration* of antivirals is exponentially distributed with mean effective duration κ. The Markov chain representing the status of an individual household, $X_k(t)$, has transitions detailed in Table 3.1. The possible states of $X_k(t)$ are,

$$S = \{(s,e,i,a) | s+e+i \leq k,\ s,e,i \geq 0,\ a = 0,1,2\}.$$

Note that once the course of antivirals is complete inside a household ($a = 2$), it is assumed that the individuals inside the household are never allocated a second course of antivirals. To determine the household reproductive ratio, R_*, and the Malthusian parameter, r, we can again use Equations (2.27) and (2.28), taking the reward function to be

$$f(X_k(t)) = \begin{cases} \alpha I(X_k(t)), & \text{when } a = 0 \text{ or } 2, \\ (1-\tau)\alpha I(X_K(t)) & \text{when } a = 1. \end{cases}$$

As before, the $(1-\tau)$ term accounts for a reduction in infectivity while taking antivirals. There is no benefit once the course of antivirals is complete and so when $a = 2$ the rate of infection is not reduced.

Similarly to what has been shown previously when considering the delay until antivirals arrive into a household, ζ, Figure 3.1 demonstrates that the Malthusian

parameter, r, is sensitive to the average effective duration, κ. As expected, when the effective duration of antivirals is small, the antivirals have little effect on the Malthusian parameter, r. Similarly, when the delay is large, the antivirals have little impact as all transmission of infection occurs before the antivirals arrive into the household. It can also be seen that the antivirals have the strongest effect when the delay is small and the effective duration is large. Note that as $\kappa \to \infty$, the Malthusian parameter, r, is similar to that obtained by Black *et al.* [15]. The difference between the result obtained in Figure 3.1 and the results obtained by Black *et al.* is because an SEIR model has been utilised in this work, whilst Black *et al.* utilised an SEEIIR model. Importantly, when the effective duration of antivirals, $\kappa = 5$, which is approximately 15 days for influenza, the Malthusian parameter, r, is reduced by approximately 30%. Comparatively, if the effective duration of antivirals, $\kappa = 1$, which is approximately 3 days for influenza, then the Malthusian parameter, r, is reduced by only 17%. This demonstrates that consideration of the effective antiviral duration is important when determining the effectiveness of antivirals during a pandemic.

3.2 Preallocation Scheme

Thus far, only the dynamic allocation scheme has been considered. It is also possible to model pandemic dynamics under a preallocation scheme. Under a preallocation scheme, there is no period of time where individuals in a household are waiting for antivirals to be delivered after the first infection. There is, however, a period of time where the individuals in a household have antivirals available but are not yet taking them as there has not yet been an infection inside the household. This phase where individuals inside a household have antivirals available but are not yet taking them is represented by the state $(s, e, i, 3)$. The transitions for the Markov chain, $X_k(t)$, which models the preallocation scheme, are detailed in Table 3.2. Note that the transition representing the effective duration of antivirals is the same as in Section

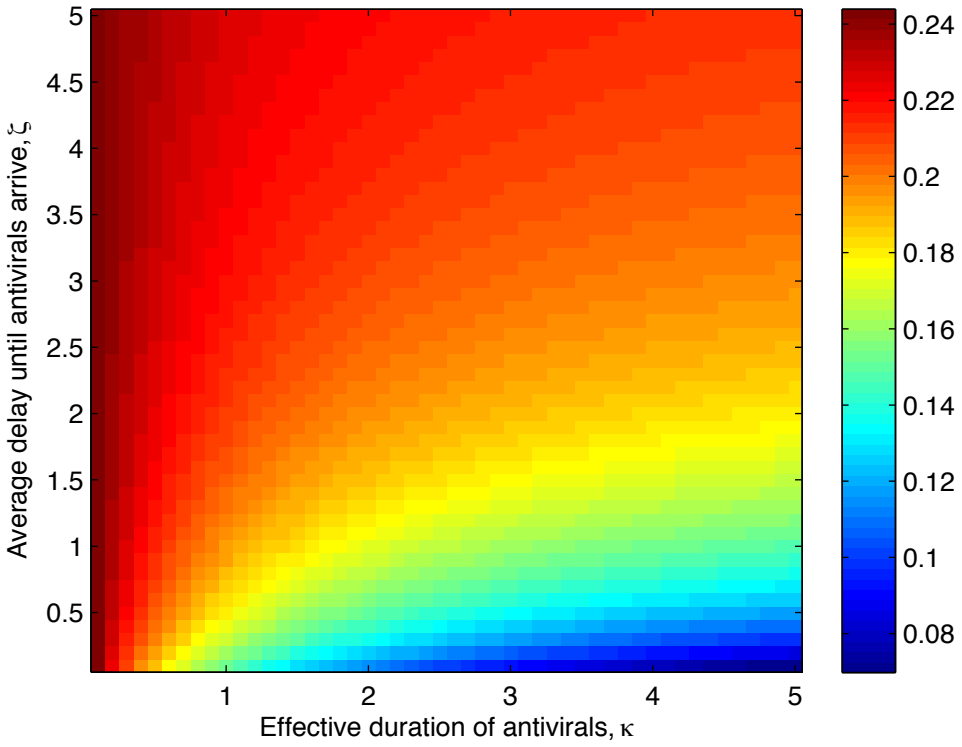


Figure 3.1: The Malthusian parameter, r, for values of mean delay until antivirals arrive in a household, ζ, and mean effective antiviral duration, κ, using the *severe* parameter set.

Event	Transition	Rate
Infection	$(s, e, i, a) \to (s-1, e+1, i, a)$	$\beta_k si$ for $a \neq 1$
		$(1-\tau)(1-\rho)\beta_k si$ for $a = 1$
Progression	$(s, e, 0, 3) \to (s, e-1, 1, 1)$	σe
	$(s, e, i, a) \to (s, e-1, i+1, a)$	
Recovery	$(s, e, i, a) \to (s, e, i-1, a)$	γi
Antivirals end	$(s, e, i, 1) \to (s, e, i, 2)$	$\frac{1}{\kappa}$

Table 3.2: Possible events inside a household during a pandemic with antivirals under a preallocation allocation scheme.

3.1, as the preallocation scheme does not affect the effective duration of antivirals. The inclusion of states of the form $(s, e, i, 3)$ again extends the state space, meaning that,

$$|S| = \frac{1}{2}(k+1)(k+2)(k+3).$$

Note that there are some states in S which are impossible to reach, such as $(s, e, i, 0)$. These states will be utilised later when considering a 'mixture' of antiviral schemes. The Markov chain, $X_k(t)$, shares many of the same transitions as the Markov chain which models the dynamic allocation scheme, but the transition representing the introduction of antivirals into a household after the first infection is no longer present. Instead, the Markov chain, $X_k(t)$, which models the preallocation scheme has a transition from the state representing individuals in a household having antivirals available but not taking them, $(s, e, 0, 3)$, to the state in which the individuals in a household are actively taking antivirals, $(s, e-1, 1, 1)$. Note that if the average duration until antivirals arrive in a household is zero, then provided the entire population has access to antivirals, the dynamic allocation scheme will produce the same reduction in infectiousness as the preallocation scheme. Should there be insufficient antivirals for the entire population, however, then even if the mean delay, ζ, is zero, the two antiviral allocation schemes will not produce the same results, as will be

discussed in Section 3.3.1.

The household reproductive ratio, R_*, and the Malthusian parameter, r, can still be calculated using the branching process approximation which includes the preallocation antiviral scheme. To do this, Equations (2.27) and (2.28) are used with the matrix, Q, which is the generator matrix for $X_k(t)$. The reward function is taken to be,

$$f(X_k(t)) = \begin{cases} \alpha I(X_k(t)) & \text{when } a = 0, 2, 3 \\ (1-\tau)\alpha I(X_k(t)) & \text{when } a = 1. \end{cases}$$

A comparison of the Malthusian parameter, r, between the dynamic allocation scheme and the preallocation scheme for a wide range of delays can be seen in Figure 3.2. When the delay is 0, the Malthusian parameter, r, is the same for the pandemic under both the dynamic allocation scheme and the preallocation antiviral scheme. Note that this comparison assumes sufficient antivirals for an entire population, so it is expected that the preallocation antiviral scheme will lead to a lower Malthusian parameter, r, than the dynamic allocation scheme, as the preallocation antiviral scheme acts just like the dynamic allocation scheme, but with zero delay. The assumption of sufficient antivirals for the entire population is likely to be unreasonable. In the next section, this assumption is removed.

3.3 Insufficient Antivirals for an Entire Population

Previously there has been an assumption that there is sufficient antivirals for every member of the population, at least in the early stages of the pandemic. When the number of antivirals available is extremely small, obviously, this assumption is not valid. However, our interest is in the range of antivirals when this assumption is likely to be reasonable, at least for a dynamic allocation scheme.

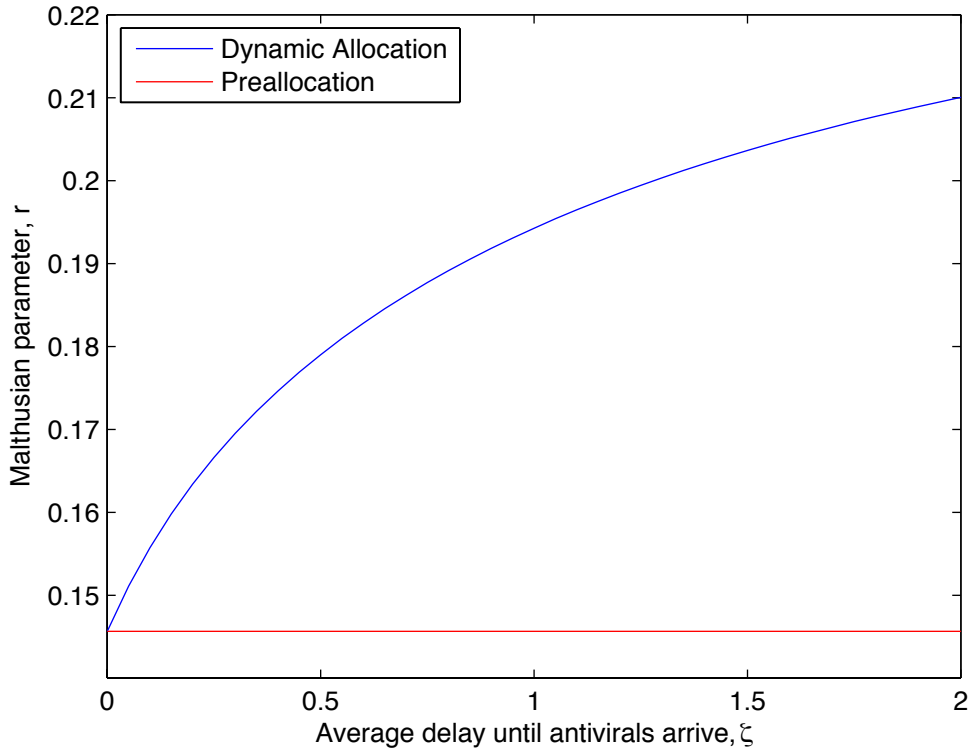


Figure 3.2: Effect of the average delay until antivirals arrive, ζ, on the Malthusian parameter, r.

Under a preallocation scheme, not all households which become infectious early in the pandemic will necessarily have access to antivirals, unlike what is likely to occur under a dynamic allocation scheme. Because of this, Equations (2.27) and (2.28) must be modified accordingly. Let ϕ_k represent the proportion of households of size k which are preallocated a supply of antivirals. Clearly, $\phi_k \in [0,1]\ \forall\, k$. If $\phi_k = 0\ \forall k$ then no households in the population are preallocated any antivirals. Similarly, if $\phi_k = 1\ \forall k$, then every household in the population is preallocated antivirals. If neither of these are true, then we have *partial preallocation*.

Assume for now that all available antivirals are preallocated. The households in the population can then be considered to have a rate of antiviral introduction, $1/\zeta$, being 0. To calculate the household reproductive ratio, R_*, the population of households is partitioned into two classes: households which have been preallocated antivirals, and households which have not. Let the branching process associated with a preallocated household of size k be $\bar{X}_k(t)$ and the branching process associated with households of size k which have not been preallocated antivirals be $X_k(t)$. These Markov chains have transitions given in Table 3.2 and Table 3.1 respectively. The equation for the household reproductive ratio is now given by the convex combination of the household reproductive ratios for $\bar{X}_k(t)$ and $X_k(t)$ respectively. That is,

$$R_* = \sum_k \pi_k p_k(0) \left(\phi_k \mathbb{E}\left[\int_0^\infty \bar{f}(\bar{X}_k(t))\, dt \right] + (1-\phi_k)\mathbb{E}\left[\int_0^\infty f(X_k(t))\, dt \right] \right) \quad (3.1)$$

with

$$\bar{f}(\bar{X}_k(t)) = \begin{cases} \alpha I(\bar{X}_k(t)) & \text{for } a = 2,3 \\ (1-\tau) I(\bar{X}_k(t)) & \text{for } a = 1 \end{cases}$$

and

$$f(X_k(t)) = \begin{cases} \alpha I(X_k(t)) & \text{for } a = 0,2 \\ (1-\tau) I(X_k(t)) & \text{for } a = 1. \end{cases}$$

Each of these path integrals can be separately evaluated using Equation (2.27). Thus, the calculation of the household reproductive ratio, R_*, is almost as straight-

forward as calculating quantities when there is sufficient antivirals for the entire population.

Similarly, the equation for the Malthusian parameter, r, is

$$\sum_k \pi_k p_k(0) \left(\phi_k \mathbb{E}\left[\int_0^\infty \bar{f}(\bar{X}_k(t)) e^{-rt} dt \right] + (1-\phi_k)\mathbb{E}\left[\int_0^\infty f(X_k(t)) e^{-rt}\, dt \right] \right) = 1 \tag{3.2}$$

with $\bar{f}(\bar{X}_k(t))$ and $f(X_k(t))$ defined as in Equation (3.1). The first term in this equation represents the households which have been preallocated antivirals, while the second term represents the households which have not been preallocated antivirals. The left hand side of Equation (3.2) can be evaluated for a given value of r, and so can be combined with a numerical root-finding algorithm to determine the Malthusian parameter, r, for which the equality is achieved.

3.3.1 Effects on a Pandemic

The effects of insufficient antivirals for an entire population can be significant. It can seen in Figure 3.3 that when there is sufficient antivirals available for the population, a preallocation scheme tends to give a lower Malthusian parameter, r, than a dynamic allocation scheme, as expected. This is because when there is a large amount of antivirals available for use, a majority of the population has antivirals available immediately under a preallocation scheme, and so the preallocation scheme leads to the same effect as the dynamic allocation scheme with zero delay. However, when there is a small number of antivirals available for use in a pandemic, the dynamic allocation scheme leads the pandemic to have a smaller Malthusian parameter, r. The dynamic allocation scheme ensures that every household which contains an infectious individual early in the pandemic receives a supply of antivirals. A preallocation scheme, however, ensures that some, roughly constant, proportion of infected individuals are taking antivirals throughout the entire pandemic at the cost of some infected individuals not receiving antivirals. The Malthusian parameter, r, is a quantity associated with the early stages of a pandemic, and so it is not unex-

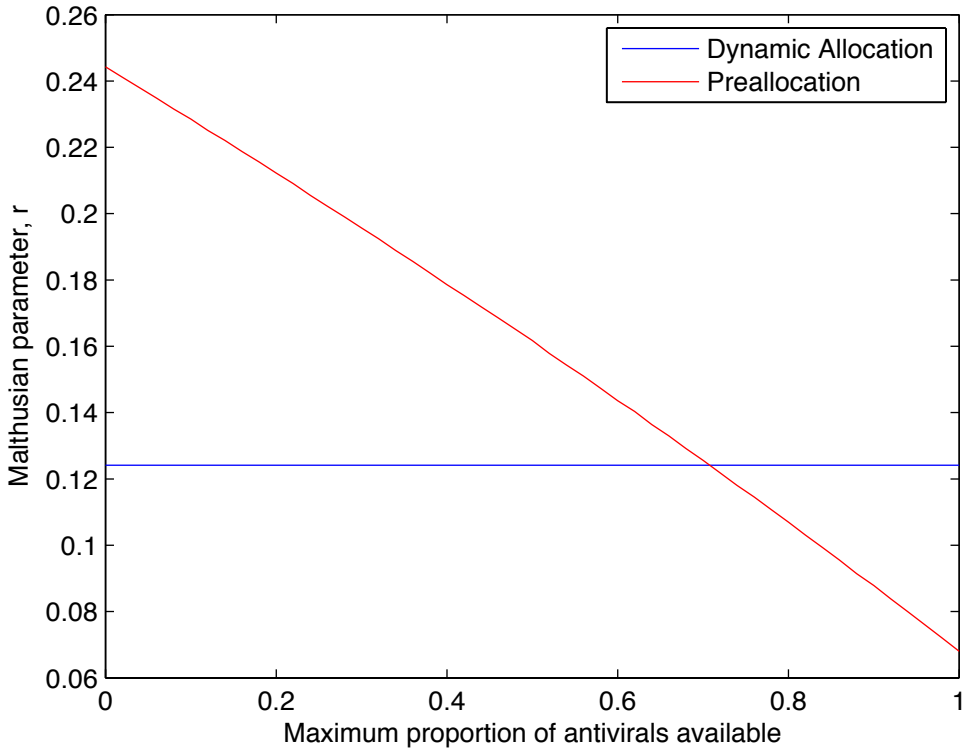


Figure 3.3: The effects of having insufficient antivirals in a population on the Malthusian parameter, r, using the *severe* parameter set.

pected that the antiviral allocation scheme which focusses on the early stages of the pandemic, that being the dynamic allocation scheme, yields a smaller Malthusian parameter, r.

3.4 Extensions to Constant-Time Events

Thus far, we have only considered event holding times which are exponentially distributed. This is because all events are modelled as transitions in a Markov chain [61]. We now consider the effective duration and the delay until antivirals arrive being of constant duration, as opposed to being exponentially distributed. The constant duration can be thought of as the opposite extreme to an exponentially distributed duration in terms of the variation of events [15]. The constant duration

means that there is no longer a Markov chain structure for each event, but quantities such as the household reproductive ratio, R_*, and the Malthusian parameter, r, can still be calculated.

3.4.1 Constant Effective Duration

For this section, we assume that the delay until antivirals arrive is exponentially distributed with mean ζ, but the effective duration is a constant time, κ. To facilitate the calculation of the household reproductive ratio, R_*, and the Malthusian parameter, r, the time of a within-household pandemic is divided into four intervals:

1. The time from first infection until the first person becomes infectious, $[0, T_1)$, where $T_1 \sim exp(1/\sigma)$,

2. The time interval until antivirals arrive to a household, $[T_1, T_1 + T_a)$, where $T_a \sim exp(1/\zeta)$,

3. The time interval that antivirals are active within a household, $[T_1 + T_a, T_1 + T_a + \kappa)$, and

4. The remainder of the pandemic, $[T_1 + T_a + \kappa, \infty)$.

Let $p_k(t)$ represent the distribution of the Markov chain, $X_k(t)$, at time t. Recall from Section 2.2 that $p_k(t)$ can be calculated as $p_k(t) = p_k(0)e^{Qt}$ where e^{Qt} is the exponential of the matrix Qt, and Q is the generator of the Markov chain, $X_k(t)$. Consider first the household reproductive ratio, R_*. Recall Equation (2.27). We calculate a single path integral from this expression,

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right],$$

with

$$f(X_k(t)) = \begin{cases} \alpha I(X_k(t)) & \text{when } a = 0, 2, \\ (1-\tau)\alpha I(X_k(t)) & \text{when } a = 1 \end{cases}.$$

The single path integral can be divided into four separate terms corresponding to the four intervals noted above as follows:

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right] = p_k(0)\mathbb{E}\left[\int_0^{T_1} \alpha I(X_k(t))\,dt + \int_{T_1}^{T_1+T_a} \alpha I(X_k(t))\,dt \right. \quad (3.3)$$
$$\left. + \int_{T_1+T_a}^{T_1+T_a+\kappa} (1-\tau)\alpha I(X_k(t))\,dt + \int_{T_1+T_a+\kappa}^{\infty} \alpha I(X_k(t))\,dt\right].$$

Note that from 0 to T_1 there is no infection in the household and so the initial state is changed from $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$ to $\bar{p}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$, where $\mathbf{u}_j$ is a vector with value 1 in the jth element, and 0 elsewhere. Using the new initial condition, $\bar{p}_k(0)$, the dependence on T_1 can be removed from Equation (3.3) giving,

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right] = \bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(t))\,dt + \int_{T_a}^{T_a+\kappa} (1-\tau)\alpha I(X_k(t))\,dt \right.$$
$$\left. + \int_{T_a+\kappa}^{\infty} \alpha I(X_k(t))\,dt\right]. \quad (3.4)$$

As expectation is linear, and $X_k(t)$ is a Markov process, the expectation of each path integral in Equation (3.4) can be taken individually, giving

$$p_k(0)\mathbb{E}\left[\int_0^\infty f(X_k(t))\,dt\right] = \bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(t))\,dt\right]$$
$$+ \bar{p}_k(T_a)\mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(t))\,dt\right]$$
$$+ \bar{p}_k(T_a+\kappa)\mathbb{E}\left[\int_0^{\infty} \alpha I(X_k(t))\,dt\right]. \quad (3.5)$$

As $T_a \sim exp(1/\zeta)$, the state of the process $X_k(t)$ at a random time T_a is required. To calculate this, a modified process, denoted $X_k^{(T_a)}(t)$ is required. The process, $X_k^{(T_a)}(t)$, is the same as $X_k(t)$ except the process is absorbed at time T_a, the time at which antivirals arrive into a household after the first infection event. Then,

$$\bar{p}_k(T_a) = \lim_{t\to\infty} \bar{p}_k(0)e^{Q^{(T_a)}t},$$

restricted to the set of transient states. Here, Q is the generator matrix of $X_k(t)$. The distribution at time T_a can be numerically estimated by evaluating $\bar{p}_k(0)e^{Q^{(T_a)}t}$ for large t. The other terms of Equation (3.5) are calculated as before.

Next, consider the Malthusian parameter, r. The situation is slightly more complex than that of the household reproductive ratio, R_*, because there is a dependence on absolute time. Recall Equation (2.28) for the Malthusian parameter. Again we expand a single term of this path integral. As before, this path integral can be divided into four distinct time periods, yielding

$$p_k(0)\mathbb{E}\left[\int_0^{T_1} \alpha I(X_k(t))e^{-rt}\,dt + \int_{T_1}^{T_1+T_a} \alpha I(X_k(t))e^{-rt}\,dt + \int_{T_1+T_a}^{T_1+T_a+\kappa} (1-\tau)\alpha I(X_k(t))e^{-rt}\,dt + \int_{T_1+T_a+\kappa}^{\infty} \alpha I(X_k(t))e^{-rt}\,dt\right] = 1, \tag{3.6}$$

with $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$. During the time period from $t = 0$ to T_1, no new infections occur in the population, and thus the first term above is also zero. Let $s = t - T_1$; then, $ds = dt$. Substituting into Equation (3.6) gives,

$$p_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-r(s+T_1)}ds + \int_{T_a}^{T_a+\kappa} (1-\tau)\alpha I(X_k(s))e^{-r(s+T_1)}ds + \int_{T_a+\kappa}^{\infty} \alpha I(X_k(s))e^{-r(s+T_1)}ds\right] = 1. \tag{3.7}$$

Since $T_1 \sim exp(\sigma)$, and the state of the Markov chain $X_k(t)$ is known at time T_1, the initial state of the Markov chain can be changed from $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$ to $\bar{p}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$. Note that this technique is valid only because the state of the process is known (deterministically) at time T_1. Integrating out the dependence on T_1 from Equation (3.7) gives,

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a}^{T_a+\kappa} (1-\tau)\alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a+\kappa}^{\infty} \alpha I(X_k(s))e^{-rs}\,ds\right] \cdot \int_0^{\infty} \sigma e^{-(r+\sigma)s}\,ds = 1.$$

The final term of this equation can be evaluated analytically and has solution $\sigma/(r+\sigma)$. Moving this to the right hand side yields

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a}^{T_a+\kappa} (1-\tau)\alpha I(X_k(s))e^{-rs}\,ds + \int_{T_a+\kappa}^{\infty} \alpha I(X_k(s))e^{-rs}\,ds\right] = \frac{r+\sigma}{\sigma}. \tag{3.8}$$

To calculate the first term of Equation (3.8), take a Markov process $\bar{X}_k(t)$ which is identical to $X_k(t)$ but is absorbed as soon as antivirals are taken. Using this process, the first term of Equation (3.8) can be rewritten as,

$$\mathbb{E}\left[\int_0^{T_a} \alpha I(X_k(s))e^{-rs}\, ds\right] = \mathbb{E}\left[\int_0^{\infty} \alpha I(\bar{X}_k(s))e^{-rs}\, ds\right] \tag{3.9}$$

which can be solved using Equation (2.17). Focussing now on the second integral of Equation (3.8), set $u = s - T_a$. Then, using the time shift technique as used when deriving Equation (3.7), the second term in Equation (3.8) can be rewritten as,

$$\mathbb{E}\left[\int_{T_a}^{T_a+\kappa} (1-\tau)\alpha I(X_k(s))e^{-rs}\, ds\right] = \mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(u+T_a))e^{-r(u+T_a)} du\right]. \tag{3.10}$$

Recall that $T_a \sim exp(\zeta)$. To remove the dependence on T_a from the expectation in Equation (3.10), we utilise the probability density function of T_a, that being $f_{T_a}(x) = \zeta e^{-\zeta x}$. Then integrating over all possible values of $T_a \in [0, \infty)$ gives,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(u+T_a))e^{-r(u+T_a)} du\right] \\
&\quad = \int_0^{\infty} \bar{p}_k(x)\zeta e^{-\zeta x} \times \mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(u))e^{-r(u+x)} du\right] dx,
\end{aligned} \tag{3.11}$$

where $\bar{p}(x)$ denotes the distribution of the Markov chain at time x. Removing terms independent of u from the expectation in Equation (3.11) gives,

$$\begin{aligned}
&\int_0^{\infty} \bar{p}_k(x)\zeta e^{-\zeta x} \times \mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(u))e^{-r(u+x)} du\right] dx \\
&\quad = \int_0^{\infty} \bar{p}_k(x)\zeta e^{-x(\zeta+r)} \times \mathbb{E}\left[\int_0^{\kappa} (1-\tau)\alpha I(X_k(u))e^{-ru} du\right] dx.
\end{aligned} \tag{3.12}$$

Note that $\bar{p}_k(x) = \bar{p}_k(0)e^{Qx}$, restricted to the set of transient states, where Q is the generator matrix of $X_k(t)$, and so the first term of Equation (3.12) is,

$$\int_0^{\infty} \bar{p}_k(x)\zeta e^{-x(\zeta+r)}\, dx = \int_0^{\infty} \bar{p}_k(0)e^{Qx}\zeta e^{-x(\zeta+r)}\, dx.$$

Moving terms independent of x outside this integral gives,

$$\int_0^{\infty} \bar{p}_k(0)e^{Qx}\zeta e^{-x(\zeta+r)}\, dx = \bar{p}_k(0)\zeta \int_0^{\infty} e^{-x((\zeta+r)I-Q)}\, dx. \tag{3.13}$$

Using the result from Section 2.3, Equation (3.13) is,

$$\int_0^\infty e^{-x((\zeta+r)I-Q)} = ((\zeta+r)I-Q)^{-1},$$

which, substituting into Equation (3.12) gives,

$$\begin{aligned}&\bar{p}_k(0)\zeta \int_0^\infty e^{-x((\zeta+r)I-Q)}dx \times \mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha I(X(k(u))e^{-ru}du\right] \\ &= \bar{p}_k(0)\zeta\left[(\zeta+r)I-Q\right]^{-1} \times \mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha I(X(k(u))e^{-ru}du\right].\end{aligned} \tag{3.14}$$

The path integral in Equation (3.14) can be evaluated using Equation (2.28).

Finally, focussing on the third integral of Equation (3.8), we have

$$\mathbb{E}\left[\int_{T_a+\kappa}^\infty \alpha I(X_k(s))e^{-rs}ds\right]. \tag{3.15}$$

Denote the generator matrix for a household not yet taking antivirals by Q_1, and the generator matrix for a household under the effect of antivirals by Q_2. Let $u = t - T_a$. Then, Equation (3.15) can be expanded as,

$$\begin{aligned}&\mathbb{E}\left[\int_{T_a+\kappa}^\infty \alpha I(X_k(s))e^{-rs}ds\right] = \\ &\int_0^\infty \bar{p}_k(0)e^{Q_1 x}\zeta e^{-\zeta x}e^{-rx}\,\mathbb{E}\left[\int_{T_c}^\infty \alpha I(X_k(u))e^{-ru}\,du\right]\,dx,\end{aligned} \tag{3.16}$$

using the same technique as in Equation (3.12). Let $v = u - \kappa$. As κ is constant, Equation (3.16) can be expanded further as,

$$\begin{aligned}&\int_0^\infty \bar{p}_k(0)e^{Q_1 x}\zeta e^{-\zeta x}e^{-rx}\,\mathbb{E}\left[\int_\kappa^\infty \alpha I(X_k(u))e^{-ru}\,du\right]\,dx \\ &= \int_0^\infty \bar{p}_k(0)e^{Q_1 x}\zeta e^{-\zeta x}e^{-rx}e^{Q_2\kappa}e^{-r\kappa}\,\mathbb{E}\left[\int_0^\infty \alpha I(X_k(v))e^{-rv}\,dv\right]\,dx,\end{aligned}$$

which, taking terms independent of x outside the first integral above gives,

$$\begin{aligned}&\int_0^\infty \bar{p}_k(0)e^{Q_1 x}\zeta e^{-\zeta x}e^{-rx}e^{Q_2\kappa}e^{-r\kappa}\,\mathbb{E}\left[\int_0^\infty \alpha I(X_k(v))e^{-rv}\,dv\right]\,dx \\ &= \zeta\bar{p}_k(0)e^{\kappa(Q_2-rI)}\int_0^\infty e^{-x((\zeta+r)I-Q_1)}\,\mathbb{E}\left[\int_0^\infty \alpha I(X_k(v))e^{-rv}\,dv\right]\,dx.\end{aligned} \tag{3.17}$$

Using the results in Section 2.3, the first integral of Equation (3.17) is,

$$\int_0^\infty e^{-x((\zeta+r)I-Q_1)}\,dx = ((\zeta+r)I-Q_1)^{-1},$$

which, substituted into Equation (3.17) gives,

$$\begin{aligned}\zeta\bar{p}_k(0)e^{-\kappa(rI-Q_2)}\int_0^\infty e^{-x((\zeta+r)-Q_1)}\,dx\,\mathbb{E}\left[\int_0^\infty \alpha I(X_k(s))e^{-rs}\,ds\right] = \\ = \zeta\bar{p}_k(0)e^{-\kappa(rI-Q_2)}\left[(\zeta+r)I-Q_1\right]^{-1}\mathbb{E}\left[\int_0^\infty \alpha I(X_k(s))e^{-rs}ds\right]\end{aligned} \tag{3.18}$$

which can again be evaluated in the same way as Equations (3.9) and (3.14). Combining Equations (3.9), (3.14) and (3.18) gives the expression for a single path integral from Equation (3.6) as,

$$\begin{aligned}&\bar{p}_k(0)\mathbb{E}\left[\int_0^{T_a}\alpha I(X_k(s))e^{-rs}\,ds\right] \\ &\quad+\bar{p}_k(0)\zeta[(\zeta+r)I-Q_1]^{-1}\mathbb{E}\left[\int_0^{\kappa}(1-\tau)\alpha I(X_k(s))e^{-rs}\,ds\right] \\ &\quad+\zeta\bar{p}_k(0)e^{-\kappa(rI-Q_2)}[(\zeta+r)I-Q_1]^{-1}\mathbb{E}\left[\int_0^\infty \alpha I(X_k(s))e^{-rs}\,ds\right] \;=\frac{r+\sigma}{\sigma}.\end{aligned} \tag{3.19}$$

The left hand side of Equation (3.19) can be evaluated for a given value of r with Equations (2.17) and (2.19) and so equality can be determined using a numerical root-finding algorithm.

To incorporate the preallocation scheme, we can break up the pandemic into the following intervals:

1. $[0, T_1)$, where T_1 is the time to the first infection. This is also the time when antivirals arrive to the household,
2. $[T_1, T_1+\kappa)$, where κ is the time when antivirals are no longer effective and,
3. $[T_1+\kappa, \infty)$, the remainder of the pandemic.

Similarly to Equation (3.4), the equation for the household reproductive ratio, R_*, can be expressed as,

$$R_* = p_k(0)\mathbb{E}\left[\int_0^{T_1}\alpha I(X_k(t))\,dt + \int_{T_1}^{T_1+\kappa}(1-\tau)\alpha I(X_k(t))\,dt + \int_{T_1+\kappa}^{\infty}\alpha I(X_k(t))\,dt\right].$$

Again, we can change the initial condition from state $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$ to $\bar{p}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$ and remove the dependence on T_1, giving,

$$\begin{aligned} R_* &= p_k(0)\mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha I(X_k(t))\,dt + \int_\kappa^\infty \alpha I(X_k(t))\,dt\right] \\ &= \bar{p}_k(0)\mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha I(X_k(t))\,dt\right] + \bar{p}(0)e^{Q\kappa}\,\mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\,dt\right], \end{aligned} \tag{3.20}$$

using the Markov property, and that $p_k(\kappa) = p_k(0)e^{Q\kappa}$ restricted to the set of transient states. The first expectation of Equation (3.20) above is over a finite time, and the second over an infinite time. Both of these expectations of path integrals can be evaluated using Equation (2.19) and (2.17) respectively.

For the Malthusian parameter, r, recall Equation (2.25). Again this can be divided into 3 intervals, giving

$$\begin{aligned} p_k(0)\mathbb{E}\left[\int_0^{T_1} \alpha e^{-rt} I(X_k(t))\,dt + \int_{T_1}^{T_1+\kappa} (1-\tau)\alpha e^{-rt} I(X_k(t))\,dt \right. \\ \left. + \int_{T_1+\kappa}^\infty \alpha e^{-rt} I(X_k(t))\,dt\right] = 1. \end{aligned} \tag{3.21}$$

As was the case when deriving the Malthusian parameter, r, previously, the first integral of Equation (3.21) zero as there are no infectious events. Moving the initial state from $p_k(0) = \mathbf{u}_{(k-1,1,0,0)}$ to $\bar{p}_k(0) = \mathbf{u}_{(k-1,0,1,0)}$ and integrating out the dependence on T_1 from Equation (3.21), gives

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha e^{-rt} I(X_k(t))\,dt + \int_\kappa^\infty \alpha e^{-rt} I(X_k(t))\,dt\right] = \frac{r+\sigma}{\sigma},$$

and using the linearity of expectations and the Markov property yields,

$$\bar{p}_k(0)\mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha e^{-rt} I(X_k(t))\,dt\right] + \bar{p}_k(\kappa)\mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\,dt\right] = \frac{r+\sigma}{\sigma}. \tag{3.22}$$

The first term of Equation (3.22) is an expectation of a path integral over a finite time, and the second term is the expectation of a path integral over infinite time, and so the left hand side can be evaluated for a given value of r using Equations (2.19) and (2.17). Recall that $\bar{p}_k(\kappa) = \bar{p}_k(0)e^{Q\kappa}$, restricted to the set of transient states, can be evaluated using EXPOKIT. This means that the left hand side of

Equation (3.22) can be calculated for a given value of r, and so the Malthusian parameter, r, for which equality holds, can be found using a numerical root-finding algorithm.

3.4.2 Constant Delay and Effective Duration

If both the delay until antivirals arrive and their active duration are constant, then the situation is much simpler as there is no need to account for the randomness of the time between the first infection event in a household and antivirals arriving or the effective duration of antivirals. Consider first a pandemic under a dynamic allocation scheme. The pandemic can be divided into 3 time intervals,

1. The time until antivirals arrive into a household, $[0, \zeta)$,
2. The time while antivirals are active inside a household, $[\zeta, \zeta + \kappa)$, and
3. The remainder of the pandemic, $[\zeta + \kappa, \infty)$.

The equation for R_* is as follows:

$$\begin{aligned} R_* &= p_k(0)\mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\,dt\right] \\ &= p_k(0)\mathbb{E}\left[\int_0^\zeta \alpha I(X_k(t))\,dt\right] + p_k(\zeta)\mathbb{E}\left[\int_0^\kappa (1-\tau)\alpha I(X_k(t))\,dt\right] \\ &\quad + p_k(\zeta+\kappa)\mathbb{E}\left[\int_0^\infty \alpha I(X_k(t))\,dt\right]. \end{aligned} \tag{3.23}$$

The first two terms of Equation (3.23) can be evaluated using Equation (2.19), while the third term can be evaluated using Equation (2.17).

Similarly, with both events being constant, the Malthusian parameter has equation,

$$\begin{aligned} &p_k(0)\mathbb{E}\left[\int_0^\zeta e^{-rt}\alpha I(X_k(t))\,dt\right] + p_k(\zeta)\mathbb{E}\left[\int_0^\kappa e^{-rt}(1-\tau) I(X_k(t))\,dt\right] \\ &\quad + p_k(\zeta+\kappa)\mathbb{E}\left[\int_0^\infty e^{-rt}\alpha I(X_k(t))\,dt\right] = 1. \end{aligned} \tag{3.24}$$

Considering a pandemic under a preallocation scheme, there is no delay until antivirals arrive. As the effective duration of antivirals is constant, the household reproductive ratio, R_*, can be calculated using Equation (3.20) and the Malthusian parameter, r, can be calculated by using Equation (3.22) with a numerical root finding algorithm.

The difference in the Malthusian parameter, r, when the effective duration of antivirals, κ, is exponentially distributed and of constant duration, is shown in Figure 3.4. Note that the exponentially distributed case yields a consistently higher Malthusian parameter, r, than the constant duration case. If the antiviral duration is exponentially distributed then there are likely to be households in the population which contain individuals who receive the benefit of antivirals for a small amount of time relative to the mean. These households which have a short effective antiviral duration do not receive much benefit from antivirals, and so the Malthusian parameter, r, increases. Note, however, that the choice of distribution of effective duration of antivirals, κ, appears to only qualitatively shift the Malthusian parameter, r. The general behaviour, that being that the dynamic allocation scheme yields a higher Malthusian parameter, r, than the preallocation scheme when the entire population has access to antivirals, still holds, regardless of whether an exponentially distributed or constant duration effective duration of antivirals, κ, is used.

3.5 Summary

In this chapter we have used a branching process approximation to calculate the household reproductive ratio, R_*, and the Malthusian parameter, r, using path integral methods. We also considered the delay until antivirals arrive into a household, ζ, and the effective duration of antivirals, κ, being of constant duration, as opposed to exponentially distributed. We showed that a constant effective duration lowered the Malthusian parameter, r, for a pandemic. This implies that assuming that the effective duration of antivirals is exponentially distributed should under-estimate

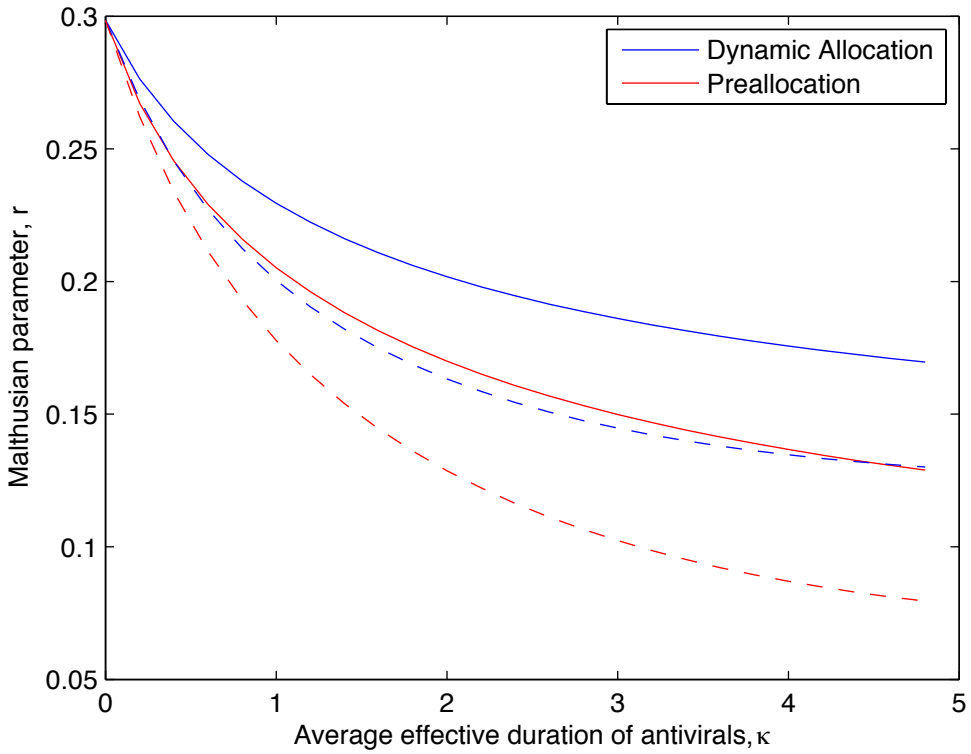


Figure 3.4: The difference in Malthusian parameter when the duration of antivirals is exponentially distributed or constant, using the *severe* parameter set from Table 2.3. Solid lines are when the effective duration of antivirals, κ, is exponentially distributed. Dashed lines are when the effective duration of antivirals, κ, is of constant duration.

the impact of antivirals. The same effect was observed for the delay until antivirals arrive into a household by Black *et al.* [15]. Importantly, the assumption of exponential duration did not largely alter the difference between the Malthusian parameter, r, under a dynamic allocation scheme and a preallocation scheme and so the assumption of this time being exponentially distributed should not affect the comparison of antiviral allocation schemes.

While the branching process is fast to compute, quantities such as the *final epidemic size*, the *peak size* and the *peak time* of the pandemic cannot be obtained. While the Malthusian parameter, r, and the household reproductive ratio, R_*, provide some indicators as to the severity of a pandemic and can give some indication as to the potential benefit of preallocation, such as a lower Malthusian parameter, r, it is still important to consider these other quantities. Intuitively, preallocation keeps an 'average level' of antivirals being taken throughout a pandemic, while dynamic allocation ensures that the level of antivirals being taken early in a pandemic is very high. This may lead to the preallocation scheme outperforming the dynamic allocation scheme over the entire pandemic, while the dynamic allocation scheme may outperform the preallocation scheme early in the pandemic. This cannot be tested using the results obtained using the branching process approximation.

The expected final epidemic size, expected peak size and expected peak time, which are associated with the long term behaviour of a pandemic, can be calculated using simulation as described in Chapter 2, however, a single realisation takes 30 minutes on an Intel Xeon 2.6GHz processor for a population of 10^5 individuals. In the next chapter, a deterministic approximation to the stochastic household model is derived. The deterministic approximation is fast to compute, but also contains all the information required to calculate the expected final epidemic size, expected peak size and expected peak time as well as the expected Malthusian parameter, r.

Chapter 4

Deterministic Approximation

In Chapters 2 and 3, we approximated the stochastic households model with a *branching process*. The key assumption in the derivation of this approximation is that the number of households in the population is infinite and so external infections happen into otherwise fully susceptible households. This means that the branching process cannot be used to accurately predict quantities such as the final epidemic size, as the number of individuals infected will grow infinitely. Recall that it is possible to estimate the expected final epidemic size, expected peak time and expected peak size by simulating the stochastic households model, however this process is time consuming. In this chapter we derive a deterministic approximation for the stochastic households model using a functional law of large numbers [49]. The deterministic nature of the approximation means that it is fast to compute, and will only need to be calculated once for each pandemic scenario. The deterministic approximation by its very nature approximates the expected value of all key quantities associated with a pandemic. However, the deterministic approximation can not be used reliably for small population sizes, or to estimate quantities very early in the pandemic. As we are looking to calculate the expected final epidemic size, expected peak size, expected peak time and expected Malthusian parameter in a large population size, these limitations of the deterministic approximation are not an issue. In this chapter we also consider a number of extensions to the stochastic households model and

the deterministic approximation, such as: having insufficient antivirals for an entire population; a *hybrid* antiviral allocation scheme; individuals in a household using antivirals incorrectly; and also the production of antivirals during a pandemic.

4.1 Development by Others

The deterministic approximation we derive is analogous to the model derived by Ball [8] for the SIS model. A deterministic approximation has been shown to be valid using the results of Kurtz [49] as the number of households $N \to \infty$ [8]. This result has also been attained using techniques from statistical physics [31]. There has also been work which outlines a procedure that constructs the system of differential equations which govern a deterministic approximation for the spread of disease through a population where the disease is modelled using the SIR model [16]. We follow a similar procedure to develop a deterministic approximation for the spread of the disease throughout a population, where the disease is modelled using the SEIR household model.

House and Keeling [39] utilise a similar deterministic approximation for an SIR model in order to investigate the effects of antivirals on a pandemic, however the model did not include the reduction to susceptibility, ρ. It was determined that as the size of households, k, increased, the amount of households required to receive antivirals to control the pandemic decreased. This is because a small number of large households receiving antivirals is equivalent to a large number of small households receiving antivirals in terms of the number of people who are taking antivirals. In this chapter, we focus on the difference between antiviral allocation schemes, however, the results of House and Keeling suggest that investigation into household size distribution may also be necessary.

4.2 Derivation

When constructing the stochastic households model, we assumed that each household was uniquely identifiable, while the individuals inside the household were not. We now seek an alternative representation of the stochastic households model in which each household is no longer uniquely identifiable [16]. Let $H_{(s,e,i)}$ represent the number of households with s susceptible, e exposed, and i infectious individuals. We initially assume that all households have a fixed size, k, an assumption which we relax in Section 4.4. We also assume that there is a fixed number of households, N, in the population, and so the total population size is Nk. Clearly, $s+e+i \leq k$, and the number of recovered individuals is $r = k-(s+e+i)$, as in Chapter 2. The set of possible household states, or configurations, is,

$$C = \{(s,e,i) | s+e+i \leq k,\ s,e,i \geq 0\}.$$

The state space for the process, $\mathbf{H}(t)$, representing the disease dynamics in a population is,

$$S = \left\{ \mathbf{H} = [H_{(s,e,i)}] \,|(s,e,i) \in C,\ H_{(s,e,i)} \geq 0 \sum_{(s,e,i)\in C} H_{(s,e,i)} = N \right\}.$$

The vector $\mathbf{H}(t)$ has a component for each configuration $(s,e,i) \in C$. As such the length of the vector, $\mathbf{H}(t)$, is,

$$|C| = \frac{1}{6}(k+1)(k+2)(k+3),$$

which is independent of N, unlike the original formulation in Section 2.5. Let,

$$\hat{H}(t) = \sum_{(s,e,i)\in C} iH_{(s,e,i)},$$

be the total number of infectious individuals in the population at time t. The process, $\mathbf{H}(t)$, has the following transitions,

$$\begin{aligned}
(H_{(s,e,i)}, H_{(s-1,e+1,i)}) &\to (H_{(s,e,i)} - 1, H_{(s-1,e+1,i)} + 1) \text{ at rate } \left(\beta_k si + \frac{\alpha}{Nk}\hat{H}(t)s\right) H_{(s,e,i)} \\
(H_{(s,e,i)}, H_{(s,e-1,i+1)}) &\to (H_{(s,e,i)} - 1, H_{(s,e-1,i+1)} + 1) \text{ at rate } \sigma e H_{(s,e,i)} \\
(H_{(s,e,i)}, H_{(s,e,i-1)}) &\to (H_{(s,e,i)} - 1, H_{(s,e,i-1)} + 1) \text{ at rate } \gamma i H_{(s,e,i)},
\end{aligned} \tag{4.1}$$

which represent infection, progression, and recovery events respectively. Here, we are effectively saying that a transition destroys a household of one type (corresponding to a -1 value), and creates a household of the new type (corresponding to a +1 value). These transitions can be represented in terms of *stoichiometric* matrices. For the three transitions in Equation (4.1), we have three stoichiometric matrices, L_1, L_2 and L_3, with (m,n)th element,

$$\begin{aligned}
L_1^{(m,n)} &= \delta_{i,i^*}(-\delta_{s,s^*}\delta_{e,e^*} + \delta_{s,s^*-1}\delta_{e,e^*+1}),\\
L_2^{(m,n)} &= \delta_{s,s^*}(-\delta_{e,e^*}\delta_{i,i^*} + \delta_{e,e^*-1}\delta_{i,i^*+1}),\\
L_3^{(m,n)} &= \delta_{s,s^*}\delta_{e,e^*}(-\delta_{i,i^*} + \delta_{i,i^*-1}),
\end{aligned}$$

for each $m = (s,e,i), n = (s^*,e^*,i^*) \in C$, which represent the infection, progression, and recovery transitions respectively; here, $\delta_{j,k}$ is the standard Kronecker delta function, that is,

$$\delta_{j,k} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k \end{cases}.$$

The index of the stoichiometric matrix determines the type of event, infection, progression, or recovery. The (m,n)th entry corresponds to the transition of a household of type m to n. The rates at which each event occurs can be encapsulated in three time dependent vectors, $\mathbf{y}_1(t), \mathbf{y}_2(t), \mathbf{y}_3(t)$, with elements

$$\begin{aligned}
\mathbf{y}_1^{(n)}(t) &= \left(\beta_k si + \frac{\alpha}{Nk}\hat{H}(t)s\right) H_n(t)\\
\mathbf{y}_2^{(n)}(t) &= \sigma e H_n(t),\\
\mathbf{y}_3^{(n)}(t) &= \gamma i H_n(t),
\end{aligned}$$

for each $n = (s,e,i) \in C$.

The state of the process, $\mathbf{H}(t)$, can be evaluated by using the event times and the sequence of events that have occurred up to time t. Let v_t represent the event type that occurred at time t. Here, $v_t = 1$ corresponds to an infection event, $v_t = 2$ corresponds to a progression event, and $v_t = 3$ corresponds to a recovery event. Let,

$$u_t = \{v_k | k < t\},$$

be the set of all events that have occurred up to time t. Also, let c_j be the household configuration which is affected by event j. Then, the state of the process, $\mathbf{H}(t)$, is,

$$\mathbf{H}(t) = \mathbf{H}(0) + \sum_{j \in u_t} L_j^{(c_j,:)},$$

where $L_j^{(c_j,:)}$ represents the c_jth row of the matrix L_j. The process $\mathbf{H}(t)$ is a Markov chain, and so can be simulated using the Gillespie Algorithm in Algorithm 1. The state space for $\mathbf{H}(t)$ is still dependent on N, meaning that for a large number of households, this state space will become large. We now seek a deterministic approximation which holds asymptotically as $N \to \infty$. Let,

$$p_{(s,e,i)} = N^{-1} H_{(s,e,i)}.$$

The transition rates in Equation (4.1) can be re-written as

$$\left(\beta_k si + \frac{\alpha}{Nk}\hat{H}(t)s\right) H_{(s,e,i)} = N\left(\beta_k si + \alpha\hat{I}(t)s\right) p_{(s,e,i)}, \tag{4.2a}$$

$$\sigma e H_{(s,e,i)} = N\sigma e p_{(s,e,i)} \tag{4.2b}$$

$$\gamma i H_{(s,e,i)} = N\gamma i p_{(s,e,i)}, \tag{4.2c}$$

where

$$\hat{I}(t) = \frac{1}{Nk}\hat{H}(t) = \frac{1}{k}\sum_{(s,e,i)\in C} i p_{(s,e,i)},$$

is the overall proportion of individuals who are infectious at time t. The transition rates in Equations (4.2a), (4.2b) and (4.2c) are in the form of

$$N f_j\left(\frac{1}{N}H_{(s,e,i)}\right),\ j = 1, 2, 3.$$

In particular, the three rates in Equations (4.2a), (4.2b) and (4.2c) have,

$$f_1 = (\beta_k si + \alpha\hat{I}(t)s)\frac{H_{(s,e,i)}}{N},$$
$$f_2 = \sigma e\frac{H_{(s,e,i)}}{N},$$
$$f_3 = \gamma i\frac{H_{(s,e,i)}}{N}.$$

This means that the transition rates in Equations (4.2a), (4.2b) and (4.2c) are *density dependent* in the sense of Kurtz [49], which implies that as $N \to \infty$, the stochastic household model, scaled by a factor of $1/N$, converges uniformly in probability over finite time intervals to the deterministic approximation which we will shortly define. This intuitively says that the deterministic approximation is accurate for a sufficiently large population size.

Let

$$\mathbf{p}(t) = N^{-1}\mathbf{H}(t).$$

The system $\mathbf{p}(t)$ represents the proportion of households in each state $(s, e, i) \in C$. Due to the scaling by a factor of $1/N$, each component of $\mathbf{p}(t)$ is between 0 and 1, unlike each component of $\mathbf{H}(t)$, and so the deterministic approximation can be used to calculate expected quantities from pandemics with arbitrarily large N. The deterministic approximation can be expressed as a sequence of matrix-vector products using the stoichiometric matrices, L_i. Let,

$$\begin{aligned}
\mathbf{w}_1^{(n)}(t) &= \left(\beta_k si + \alpha \hat{I}(t)s\right) p_n(t) \\
\mathbf{w}_2^{(n)}(t) &= \sigma e p_n(t), \\
\mathbf{w}_3^{(n)}(t) &= \gamma i p_n(t),
\end{aligned}$$

for each $n = (s, e, i) \in C$. The deterministic approximation is,

$$\frac{d\mathbf{p}(t)}{dt} = \sum_{j=1}^{3} L_j \mathbf{w}_j(t). \tag{4.3}$$

It is worth noting that for the case where all households are of size 2, House and Keeling [39] have derived an analytic solution to the set of differential equations for the SIR model. Also, for the SIS model, Ball [9] derived an analytic solution for the case where all households are of size 2. There do not appear to be analytic solutions for the SEIR model with any household size distribution, however.

To see the structure of this approximation, consider an example in a population

of households all of size $k = 2$. The set of possible household states, C, has elements,

$$\begin{aligned} c_0 &= (0,0,0) \quad & c_5 &= (0,2,0) \\ c_1 &= (0,0,1) \quad & c_6 &= (1,0,0) \\ c_2 &= (0,0,2) \quad & c_7 &= (1,0,1) \\ c_3 &= (0,1,0) \quad & c_8 &= (1,1,0) \\ c_4 &= (0,1,1) \quad & c_9 &= (2,0,0), \end{aligned}$$

and the system of differential equations is,

$$\begin{aligned} \frac{dp_0(t)}{dt} &= \gamma p_1(t), \\ \frac{dp_1(t)}{dt} &= 2\gamma p_2(t) + \sigma p_3(t) - \gamma p_1(t), \\ \frac{dp_2(t)}{dt} &= \sigma p_4(t) - 2\gamma p_2(t), \\ \frac{dp_3(t)}{dt} &= \gamma p_4(t) + \alpha \hat{I}(t) p_6(t) - \sigma p_3(t), \\ \frac{dp_4(t)}{dt} &= 2\sigma p_5(t) + \alpha \hat{I}(t) p_7(t) - (\sigma + \gamma) p_4(t), \\ \frac{dp_5(t)}{dt} &= \alpha \hat{I}(t) p_8(t) - 2\sigma p_5(t), \\ \frac{dp_6(t)}{dt} &= \gamma p_7(t) - \alpha \hat{I}(t) p_6(t), \\ \frac{dp_7(t)}{dt} &= \sigma p_8(t) - (\alpha \hat{I}(t) + \gamma) p_7(t), \\ \frac{dp_8(t)}{dt} &= 2\alpha \hat{I}(t) p_9(t) - (\alpha \hat{I}(t) + \sigma) p_8(t), \\ \frac{dp_9(t)}{dt} &= -2\alpha \hat{I}(t) p_9(t), \end{aligned}$$

with

$$\hat{I}(t) = \frac{p_1(t) + 2p_2(t) + p_4(t) + p_7(t)}{2},$$

where $p_i(t)$ is the proportion of households in state c_i. This system of differential equations can be solved numerically using Runge-Kutta methods such as those implemented in MATLAB's `ode45` function, paired with a suitable initial condition $\mathbf{p}(0)$, which we discuss next.

4.3 Initial Condition

In order to solve Equation (4.3), a suitable initial condition, $\mathbf{p}(0)$, is required. The initial condition must be such that the proportion of the population in each state is sufficiently large. Further, we want to eliminate the transient behaviour of the system and so the system essentially starts in the early growth phase of the pandemic. This allows for a fairer comparison of the general behaviour of pandemics under different antiviral allocation schemes.

To determine an initial condition which eliminates transient behaviour, first consider the steady states of the deterministic approximation. Consider the following two states of the system—the population state where every household is in state $(k, 0, 0)$, denoted $\mathbf{p_s}$, and the population state where the entire population is in state $(0, 0, 0)$, denoted $\mathbf{p}_r$. Clearly, if the system is in either state $\mathbf{p_s}$ or state $\mathbf{p_r}$, then it will be at equilibrium as the population is either fully susceptible and so there are no infectious individuals or the population is fully recovered and so there are no individuals to infect. Further, any convex combination of these two states is also an equilibrium point. To see that $\mathbf{p_s}$ is an unstable equilibrium, consider a small perturbation from $\mathbf{p_s}$. This perturbation will cause the state of the system to never be $\mathbf{p_s}$ again as individuals who are infected do not become susceptible again. As such, there exists an eigenvalue, λ, at the point $\mathbf{p_s}$, which is positive. To see the stability of the system in state $\mathbf{p_r}$, consider a small perturbation from $\mathbf{p_r}$. This perturbation will not stop the state of the system being $\mathbf{p_r}$ again, as any infectious individuals must eventually recover. This means that all eigenvalues of the system are non-positive at the point $\mathbf{p_r}$. Clearly, we want to start 'near' the state where all individuals are susceptible, $\mathbf{p_s}$, in order to capture the early growth period of the pandemic, and so we make a small perturbation from the equilibrium point, $\mathbf{p_s}$. As such, we seek an initial condition of the form,

$$\mathbf{p}(0) = \mathbf{p_s} + \omega.$$

To determine ω, a linear stability analysis is applied to the deterministic approx-

imation [44]. Let $F(\mathbf{p}(t)) = \sum_j L_j \mathbf{w}_j(t)$. The system of differential equations which forms the deterministic approximation can be expressed as,

$$\frac{d\mathbf{p}(t)}{dt} = F(\mathbf{p}(t)).$$

Let the Jacobian of this system be J, that is,

$$J_{ij} = \frac{\partial F_i}{\partial p_j},$$

and $J|_{\mathbf{x}}$ be the Jacobian evaluated at the point $\mathbf{x}$. Linearising this system about the equilibrium point $\mathbf{p_s}$, yields

$$\begin{aligned}\frac{d\mathbf{p}(t)}{dt} &= F(\mathbf{p_s}) + J|_{\mathbf{p_s}}(\mathbf{p}(t) - \mathbf{p_s}) \\ &= J|_{\mathbf{p_s}}(\mathbf{p}(t) - \mathbf{p_s}),\end{aligned}$$

as $F(\mathbf{p_s}) = 0$. Let $\mathbf{p}(t) - \mathbf{p_s} = \delta\mathbf{p}(t)$. Then, $\frac{d\delta\mathbf{p}(t)}{dt} = \frac{d\mathbf{p}(t)}{dt}$ and,

$$\frac{d\delta\mathbf{p}(t)}{dt} = J_{\mathbf{p_s}}\delta\mathbf{p}(t). \tag{4.4}$$

As Equation (4.4) is a system of constant coefficient linear differential equations, the system can be decomposed in terms of its eigenvalues and eigenvectors. Let $J|_{p_s}$ have eigenvalues $\lambda_1, \ldots, \lambda_n$, where $Re(\lambda_1) \geq Re(\lambda_2) \geq \cdots \geq Re(\lambda_n)$, with corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. It follows that

$$\delta\mathbf{p}(t) = \sum_{j=1}^{n} \epsilon_j e^{t\lambda_j}\mathbf{v}_j$$

where ϵ_j are coefficients that are yet to be determined, which will depend on the perturbation. Hence,

$$\mathbf{p}(t) = \mathbf{p_s} + \sum_{j=1}^{n} \epsilon_j e^{t\lambda_j}\mathbf{v}_j.$$

In the limit as $t \to \infty$, only the dominant eigenvalue, λ_1, will remain in the expansion, with corresponding eigenvector $\mathbf{v}_1$. We then have,

$$\mathbf{p}(t) = \mathbf{p_s} + \epsilon_1 e^{t\lambda_1}\mathbf{v}_1. \tag{4.5}$$

In order to determine ϵ_1, consider the system at time $t = 0$. From Equation (4.5), it can be seen that,

$$\mathbf{p}(0) = \mathbf{p_s} + \epsilon_1 \mathbf{v}_1. \tag{4.6}$$

We fix the initial proportion of the population infected, denoted $i_0 \in (0, 1)$. Then,

$$i_0 = \sum_{(s,e,i)\in C} p_{(s,e,i)}(0)i = \mathbf{p}(0) \cdot \mathbf{i}, \tag{4.7}$$

where $\mathbf{i}$ is the vector of the number of infectious individuals in each state $(s, e, i) \in C$. From Equation (4.6), we have

$$\mathbf{p}(0) \cdot \mathbf{i} = \mathbf{p_s} \cdot \mathbf{i} + \epsilon_1(\mathbf{i} \cdot \mathbf{v}_1),$$

but, $\mathbf{p_s} \cdot \mathbf{i} = 0$ as there are no infectious individuals when the population is in state $\mathbf{p_s}$, and so,

$$\mathbf{p}(0) \cdot \mathbf{i} = i_0 = \epsilon_1(\mathbf{i} \cdot \mathbf{v}_1).$$

Rearranging for ϵ_1, gives,

$$\epsilon_1 = \frac{i_0}{\mathbf{i} \cdot \mathbf{v}_1}. \tag{4.8}$$

Substituting Equation (4.8) into Equation (4.6) gives the formula for determining the initial condition,

$$\mathbf{p}(0) = \mathbf{p_s} + \frac{i_0}{\mathbf{i} \cdot \mathbf{v}_1}\mathbf{v}_1. \tag{4.9}$$

This procedure constructs an initial condition which satisfies the requirements of Kurtz [49], and also eliminates transient activity in the system and ensures that the system $\mathbf{p}(t)$ starts in the exponential growth phase, which allows for a fairer comparison between pandemics.

Note that in the process of determining the initial condition, we also calculate the Malthusian parameter, r; Equation (4.5) represents the system, $\mathbf{p}(t)$, growing exponentially at rate λ_1. This means that λ_1 is precisely the Malthusian parameter, r.

In order to determine the initial condition, the dominant eigenvalue and eigenvector of the Jacobian of $\mathbf{p}(t)$ evaluated at the fixed point, $\mathbf{p_s}$, is required. Recall

the system of differential equations from Equation (4.3). This system can be differentiated term by term in order to determine the Jacobian. Each matrix L_j is independent of $\mathbf{p}(t)$ and t, so we only need to differentiate the $\mathbf{w}_j(t)$ terms. That is,

$$J(t) = \sum_{j=1}^{3} L_j \frac{\partial \mathbf{w}_j(t)}{\partial \mathbf{p}}. \tag{4.10}$$

The vectors $\mathbf{w}_2(t)$ and $\mathbf{w}_3(t)$ are linear in $\mathbf{p}(t)$ and so are straightforward to differentiate, with the nth component being,

$$\frac{\partial \mathbf{w}_2^{(n)}(t)}{\partial p_m} = \begin{cases} \sigma e_n & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases}$$

and

$$\frac{\partial \mathbf{w}_3^{(n)}(t)}{\partial p_m} = \begin{cases} \gamma i_n & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases}$$

for each $n = (s_n, e_n, i_n),\ m = (s_m, e_m, i_n) \in C$. The vector $\mathbf{w}_1(t)$ is non-linear in $\mathbf{p}(t)$, as $\hat{I}(t)$ is also a function of $\mathbf{p}(t)$. To differentiate this vector, consider the nth component,

$$\mathbf{w}_1^{(n)}(t) = \left(\beta_k s_n i_n + \alpha \hat{I}(t) s_n\right) p_n(t),$$

with

$$\hat{I}(t) = \frac{1}{k} \sum_{j \in C} i_j p_j(t) = \frac{1}{k} \left(i_n p_n(t) + \sum_{\substack{j \neq n \\ j \in C}} i_j p_j(t) \right).$$

Then, differentiating term by term gives

$$\frac{\partial \mathbf{w}_1^{(n)}(t)}{\partial p_m} = \begin{cases} \beta_k s_n i_n + \dfrac{1}{k} s_n \left(2\alpha i_n p_n(t) + \alpha \displaystyle\sum_{\substack{j \neq n \\ j \in C}} i_j p_j(t) \right), & \text{if } n = m, \\ \frac{1}{k} \alpha s_n i_m p_n(t), & \text{if } n \neq m. \end{cases}$$

The fixed point $\mathbf{p_s}$ can be substituted into the Jacobian and the resulting eigenvalues can be determined numerically. The resulting dominant eigenvalue, λ_1 (which

is also the Malthusian parameter, r) and the corresponding eigenvector $\mathbf{v}_1$, can then be used in Equation (4.9) to determine the initial condition which ensures that the system $\mathbf{p}(t)$ starts in the exponential growth phase.

4.4 Heterogeneous Household Sizes

As in the branching process approximation, we want to incorporate households of different sizes. The size of a household, k, is incorporated into the set of household states in order to ensure that a transition maintains the size of a household. That is,

$$C = \{(s, e, i, k) | s + e + i \leq k,\ s, e, i \geq 0,\ k = 1, \ldots, k_{\max}\}.$$

Note that accounting for household size, k, in the set of household states, C, is equivalent to incorporating the number of recovered individuals in a household, r, however, the household size is what is often required in calculations.

The transitions of a household in the population remain the same as those in Equation (4.1), however, the matrices which represent the states in which events occur require some extension. In particular, the matrices L_1, L_2 and L_3 are now such that

$$\begin{aligned}
L_1^{(m,n)} &= \delta_{k,k^*}\delta_{i,i^*}(-\delta_{s,s^*}\delta_{e,e^*} + \delta_{s,s^*-1}\delta_{y,y^*+1}),\\
L_2^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}(-\delta_{e,e^*}\delta_{i,i^*} + \delta_{e,e^*-1}\delta_{i,i^*+1}),\\
L_3^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}\delta_{e,e^*}(-\delta_{i,i^*} + \delta_{i,i^*-1}),
\end{aligned} \tag{4.11}$$

for $m = (s, e, i, k)$ and $n = (s^*, e^*, i^*, k^*) \in C$. The three time-dependent vectors which encapsulate the rates at which these events occur are,

$$\mathbf{w}_1^{(n)}(t) = \left(\beta_k si + \alpha\hat{I}(t)s\right) p_n(t), \tag{4.12a}$$

$$\mathbf{w}_2^{(n)}(t) = \sigma e p_n(t), \tag{4.12b}$$

$$\mathbf{w}_3^{(n)}(t) = \gamma i p_n(t), \tag{4.12c}$$

for $n = (s, e, i, k) \in C$. Also, the total proportion of individuals infected in the population at time t is now,

$$\hat{I}(t) = \frac{1}{\bar{k}} \sum_{(s,e,i,k)\in C} i p_{(s,e,i,k)},$$

with the mean household size,

$$\bar{k} = \sum_{k} k h_k,$$

where h_k is the proportion of households of size k, introduced in Section 2.9. The deterministic approximation is as in Equation (4.3), with the L matrices in Equation (4.11) and $\mathbf{w}$ vectors in Equation (4.12).

4.4.1 Initial Condition

Previously, when deriving the initial condition, a perturbation was made around the point $\mathbf{p_s}$, the population state where every individual is susceptible. When there is only a single household size, the fixed point $\mathbf{p_s}$ has only a single non-zero value corresponding to state $(k, 0, 0, k)$. When there is a distribution of household sizes, the equilibrium point $\mathbf{p_s}$ has a non-zero element for each household size. For a household of size k, the value of the non-zero element of $\mathbf{p_s}$, corresponding to state $(k, 0, 0, k)$ for each k, is the proportion of households of that type in the population, h_k.

Considering the Jacobian of this system, each L_j matrix is independent of $\mathbf{p}$, and so once again we just need to differentiate each $\mathbf{w}_j$ term. As these terms have not changed, the derivatives and Jacobian of the system are as in Section 4.3. The *size-biased* distribution, $\boldsymbol{\pi}$, which was utilised in Chapters 2 and 3, appears in the initial condition. Consider the term involving α from the $\mathbf{w_1}$ vector,

$$\frac{1}{\bar{k}} \alpha s_n i_m p_n. \tag{4.13}$$

All other terms in $\frac{\partial \mathbf{w_1}}{\partial p_m}$ are zero at the point $\mathbf{p_s}$. The only non-zero p_n terms are when the household is fully susceptible, and so $s_n = k$. Also, $p_n = h_k$ from above,

so, Equation (4.13) is,

$$\begin{aligned}\frac{1}{\overline{k}}\alpha s_n i_m p_n &= \frac{kh_k}{\sum_j k_j h_j}\alpha i_m, \\ &= \pi_k \alpha i_m.\end{aligned}$$

The equation for the initial condition remains otherwise unchanged from Section 4.3.

As described in Section 4.2, the results of Kurtz [49] tell us that the deterministic approximation is valid as the number of households, $N \to \infty$, independent of the other parameters in the model. In Figure 4.1, we calculate the difference between the average of one hundred realisations from the stochastic model, using the Gillespie Algorithm in Algorithm 1, and the deterministic approximation to demonstrate the accuracy of this deterministic approximation. It can be seen that the error between the average of a number of realisations from the stochastic model and the deterministic approximation is small, even for a moderate population size of $N = 10^5$ individuals. We can see clearly that as the number of households increases, the difference between the average of realisations from the stochastic model and the deterministic approximation decreases. Interestingly, the difference between the average of the one hundred simulations and the deterministic approximation is very small at the peak of the pandemic, between $t = 28$ infectious periods and $t = 32$ infectious periods in this case. This suggests that the expected peak size and expected peak time of a pandemic are very well approximated by the deterministic approximation. Note, though, that the difference between the average of the one hundred realisations and the deterministic approximation is of order 10^{-3} across the entire pandemic, and so the deterministic approximation is a reliable method for estimating mean quantities associated with a pandemic.

The deterministic approximation solves in less than 5 seconds on an Intel Xeon 2.6GHz processor, while producing a single realisation of the stochastic households model for the *severe* parameters in Table 2.3 with a population size of 10^5 individuals

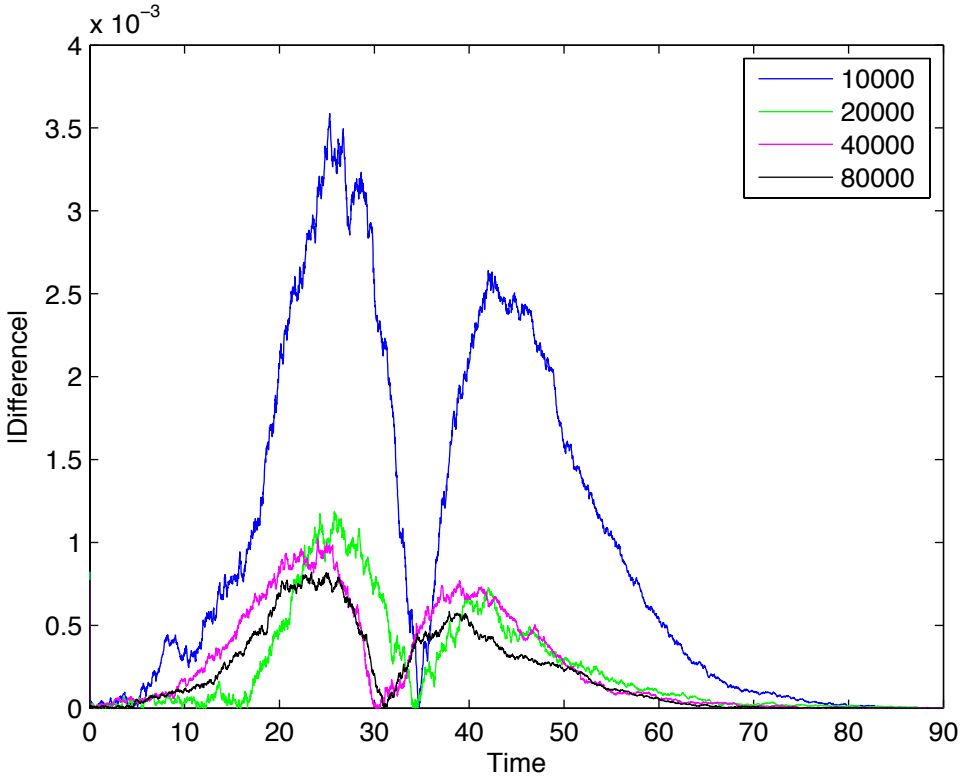


Figure 4.1: Difference between the average of 100 simulated realisations compared to the deterministic approximation using the *severe* parameter set from Table 2.3, varying population size, $N\bar{k}$.

takes approximately 30 minutes on the same infrastructure. As such, the system of differential equations which forms the deterministic approximation to the stochastic households pandemic model is a powerful tool for quickly calculating mean quantities of the pandemic.

4.5 Antiviral Allocation Schemes

To incorporate antiviral schemes, an additional class of states is required which represents the status of antivirals in a household. As in Chapter 3, we let $a = 0$ represent a household which contains individuals who have not yet begun taking antivirals, $a = 1$ represent a household which contains individuals who are currently taking antivirals, $a = 2$ represent a household which contains individuals who have completed taking antivirals and no longer receiving the effects of antivirals, and $a = 3$ represent a household which has been preallocated antivirals but has not yet begun taking them. Each household now is in a state (s, e, i, k, a) with the set of possible household configurations,

$$C = \{(s, e, i, k, a) | s + e + i \leq k,\ s, e, i \geq 0,\ k = 1, \ldots, k_{\max},\ a = 0, 1, 2, 3\}.$$

Similarly to the case without antivirals,

$$\hat{H}(t) = \sum_{(s,e,i,k,a) \in C} (1 - \delta_{a,1}\tau) i H_{(s,e,i,k,a)}(t),$$

is the total force of infection in the population at time t, but here it includes the reduction in infectivity by τ due to the effect of antivirals. The transitions for the process, $\mathbf{H}(t)$, including antiviral effects for both antiviral allocation schemes, are,

$$\begin{aligned}(H_{(s,e,i,k,a)}, H_{(s-1,e+1,i,k,a)}) \to (H_{(s,e,i,k,a)} - 1, H_{(s-1,e+1,i,k,a)} + 1) \\ \text{at rate } \left(\beta_k si + \alpha \hat{H}(t) s\right) H_{(s,e,i,k,a)} \text{ for } a \neq 1,\end{aligned}$$

which corresponds to infection into and within a household,

$$(H_{(s,e,i,k,1)}, H_{(s-1,e+1,i,k,1)}) \to (H_{(s,e,i,k,1)} - 1, H_{(s-1,e+1,i,k,1)} + 1)$$
$$\text{at rate } \left((1-\tau)(1-\rho)\beta_k si + (1-\tau)\alpha\hat{H}(t)s\right) H_{(s,e,i,k,a)},$$

corresponding to infection into and within a household which is currently taking antivirals,

$$(H_{(s,e,i,k,a)}, H_{(s,e-1,i+1,k,a)}) \to (H_{(s,e,i,k,a)} - 1, H_{(s,e-1,i+1,k,a)} + 1) \text{ at rate } \sigma e H_{(s,e,i,k,a)},$$

corresponding to infection progression,

$$(H_{(s,e,i,k,a)}, H_{(s,e,i-1,k,a)}) \to (H_{(s,e,i,k,a)} - 1, H_{(s,e,i-1,k,a)} + 1) \text{ at rate } \gamma i H_{(s,e,i,k,a)},$$

corresponding to recovery, and,

$$(H_{(s,e,i,k,1)}, H_{(s,e,i,k,2)} \to (H_{(s,e,i,k,1)} - 1, H_{(s,e,i,k,2)} + 1) \text{ at rate } \frac{1}{\kappa} H_{(s,e,i,k,1)},$$

corresponding to antivirals no longer being effective inside a household. Throughout this derivation, we assume that the effective duration of antivirals is exponentially distributed with mean κ. In the following sections, we consider the specifics of deriving the deterministic approximation for the dynamic allocation scheme and the preallocation scheme for antivirals.

4.5.1 Dynamic Allocation

Recall that in dynamic allocation there is a *delay* until antivirals arrive at a household after the first person is identified as infectious. We assume this delay is exponentially distributed with mean ζ. For a dynamic allocation scheme, we therefore have one additional transition corresponding to antiviral introduction into a household. This transition is,

$$(H_{(s,e,i,k,0)}, H_{(s,e,i,k,1)}) \to (H_{(s,e,i,k,0)} - 1, H_{(s,e,i,k,1)} + 1) \text{ at rate } \frac{1}{\zeta} H_{(s,e,i,k,0)} \text{ for } s+e \neq k.$$

Again, we can define the proportion of households in each configuration,

$$p_{(s,e,i,k,a)} = N^{-1} H_{(s,e,i,k,a)},$$

as well as the average level of infectiousness,

$$\hat{I}(t) = \frac{1}{N\bar{k}} \hat{H}(t).$$

The transition rates for the system $\mathbf{p}$ are still density dependent. The first four transition rates are identical to those discussed in Section 4.2, while the final two are of the form

$$Nx \frac{H_{(s,e,i,k,a)}}{N},$$

for some x and so the results of Kurtz [49] still apply. This means that the deterministic approximation is valid asymptotically as $N \to \infty$.

As in previous sections, the deterministic approximation can be expressed in terms of matrix-vector products. We define a matrix L_i for each transition i which encapsulates where transitions can occur within the state space, and a corresponding vector $\mathbf{w}_i$ which contains the rates at which these transitions occur. The six L matrices have elements,

$$\begin{aligned}
L_1^{(m,n)} &= \delta_{k,k^*}(1-\delta_{a,1})\delta_{i,i^*}(-\delta_{s,s^*}\delta_{e,e^*} + \delta_{s,s^*-1}\delta_{e,e^*+1}),\\
L_2^{(m,n)} &= \delta_{k,k^*}\delta_{a,1}\delta_{i,i^*}(-\delta_{s,s^*}\delta_{e,e^*} + \delta_{s,s^*-1}\delta_{e,e^*+1}),\\
L_3^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}(-\delta_{e,e^*}\delta_{i,i^*} + \delta_{e,e^*-1}\delta_{i,i^*+1}),\\
L_4^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}\delta_{e,e^*}(-\delta_{i,i^*} + \delta_{i,i^*-1}),\\
L_5^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}\delta_{e,e^*}, \delta_{i,i^*}(1-\delta_{(s+e),k})(-\delta_{a,0} + \delta_{a,1}),\\
L_6^{(m,n)} &= \delta_{k,k^*}\delta_{s,s^*}\delta_{e,e^*}, \delta_{i,i^*}(-\delta_{a,1} + \delta_{a,2}),
\end{aligned}$$

for $m = (s,e,i,k,a), n = (s^*,e^*,i^*,k^*,a^*) \in C$. The first two of these matrices correspond to infection events. The first matrix, L_1, corresponds to infection in a household without antivirals. The second matrix, L_2, corresponds to infection inside a household with antivirals. The third and fourth represent progression and recovery

respectively. The fifth matrix corresponds to the introduction of antivirals into a household that has experienced at least one infection event ($s + e \neq k$) and has not yet received antivirals, while the sixth matrix corresponds to a course of antivirals being completed in a household. The corresponding rate vectors have elements,

$$\begin{aligned}
\mathbf{w}_1^{(n)}(t) &= \left(\beta_k si + \alpha\hat{I}(t)s\right) p_n(t) \\
\mathbf{w}_2^{(n)}(t) &= \left((1-\tau)(1-\rho)\beta_k si + (1-\tau)\alpha\hat{I}(t)s\right) p_n(t) \\
\mathbf{w}_3^{(n)}(t) &= \sigma e p_n(t), \\
\mathbf{w}_4^{(n)}(t) &= \gamma i p_n(t), \\
\mathbf{w}_5^{(n)}(t) &= \frac{1}{\zeta} p_n(t), \\
\mathbf{w}_6^{(n)}(t) &= \frac{1}{\kappa} p_n(t),
\end{aligned}$$

for each $n = (s, e, i, k, a) \in C$ and,

$$\hat{I}(t) = \frac{1}{\bar{k}} \sum_{(s,e,i,k,a)\in C} i p_{(s,e,i,k,a)}(t).$$

The deterministic approximation can then be expressed as

$$\frac{d\mathbf{p}}{dt} = \sum_{j=1}^{6} L_j \mathbf{w}_j(t). \tag{4.14}$$

As before, an initial condition which starts the system in the exponential growth phase is desired in order to allow a fair comparison of the models. To do this, we consider a small perturbation from the fixed point denoted $\mathbf{p_s}$ which corresponds to the state where the entire population is susceptible. That is, $\mathbf{p_s}$ has value h_k in state $(k, 0, 0, k, 0)$ for each k with all other elements being 0. Considering the Jacobian of the system in Equation (4.14), we again have no time or state dependence in each matrix L_i. As such, we only need to consider the partial derivatives of the $\mathbf{w}_i$ vectors. These share much of the same structure as those in Section 4.3. The $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ and $\mathbf{w}_4$ terms are very similar to those in Section 4.3, with the difference being the inclusion of constants to incorporate the effects of antivirals. The system of partial derivatives is, then,

$$\frac{\partial \mathbf{w}_1^{(n)}(t)}{\partial p_m} = \begin{cases} \beta_k s_n i_n + \frac{1}{\bar{k}} s_n \left(2\alpha i_n p_n(t) + \alpha \sum_{\substack{j \neq n \\ j \in C}} i_j p_j(t) \right), & \text{if } n = m, \\ \frac{1}{\bar{k}} \alpha s_n i_n p_n(t), & \text{if } n \neq m, \end{cases}$$

$$\frac{\partial \mathbf{w}_2^{(n)}(t)}{\partial p_m} = \begin{cases} (1-\tau)(1-\rho)\beta_k s_n i_n + \frac{1}{\bar{k}}(1-\tau) s_n \left(2\alpha i_n p_n(t) + \alpha \sum_{\substack{j \neq n \\ j \in C}} i_j p_j(t) \right), & \text{if } n = m, \\ \frac{1}{\bar{k}}(1-\tau)\alpha s_n i_n p_n(t), & \text{if } n \neq m, \end{cases}$$

$$\frac{\partial \mathbf{w}_3^{(n)}(t)}{\partial p_m} = \begin{cases} \sigma e_n & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases}$$

$$\frac{\partial \mathbf{w}_4^{(n)}(t)}{\partial p_m} = \begin{cases} \gamma i_n & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases}$$

$$\frac{\partial \mathbf{w}_5^{(n)}(t)}{\partial p_m} = \begin{cases} \frac{1}{\zeta} & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases}$$

$$\frac{\partial \mathbf{w}_6^{(n)}(t)}{\partial p_m} = \begin{cases} \frac{1}{\kappa} & \text{if } n = m, \\ 0 & \text{if } n \neq m, \end{cases} \tag{4.15}$$

and the Jacobian of the system in Equation (4.14) is

$$J = \sum_{j=1}^{6} L_j \frac{\partial \mathbf{w}_j(t)}{\partial \mathbf{p}}.$$

4.5.2 Preallocation

Now, consider the preallocation scheme. A household being preallocated antivirals, that is, in state $(s, e, i, k, 3)$, will transition into state $(s, e, i, k, 1)$ upon the first infection event. After this point, the household proceeds identically to the dynamic allocation scheme. Mathematically, if $H_{(s,e,i,k,a)}$ is again the number of households

in state (s, e, i, k, a), we have the transition,

$$(H_{(s,e,0,k,3)}, H_{(s,e-1,1,k,1)}) \to (H_{(s,e,0,k,3)} - 1, H_{(s,e-1,1,k,1)} + 1) \text{ at rate } \sigma e H_{(s,e,0,k,3)}, \tag{4.16}$$

with $s + e = k$ to ensure that the household starts taking antivirals on the first infection event. The remainder of the transition rates remain the same as those in Section 4.5. While we could remove the class of states where $a = 0$, it proves useful not to do so when considering extensions to antiviral allocation schemes, discussed in Section 4.6. The rate at which the transition in Equation (4.16) occurs is clearly density dependent for the system, as the transition rate is identical to the normal progression rate considered in Section 4.2, and so the deterministic approximation will still provide valid results as the population size $N \to \infty$.

Again, the deterministic approximation is expressed in terms of matrix-vector products. Six of these matrices are identical to those in Section 4.5.1, the only addition being the matrix L_0 with (m, n)th element,

$$L_0^{(m,n)} = \delta_{k,k^*}\delta_{s,s^*}(-\delta_{a,3}\delta_{e,e^*}\delta_{i,i^*} + \delta_{a,1}\delta_{e,e^*-1}\delta_{i,i^*+1}),$$

for $m = (s, e, i, k, a), n = (s^*, e^*, i^*, k^*, a^*) \in C$ and the corresponding rate vector,

$$\mathbf{w}_0^{(n)}(t) = \sigma e p_n(t),$$

for each $n = (s, e, i, k, a) \in C$.

Looking at the initial condition for this system, a majority of the Jacobian remains the same as in the dynamic allocation case. The only additional term added to the system is the $L_0\mathbf{w}_0$ term, with L_0 being constant with respect to $\mathbf{p}$ and t, and so does not need to be considered. Also, the $\mathbf{w}_0$ term is identical to the $\mathbf{w}_3$ term and so the derivative of $\mathbf{w}_0$ is the same as that of $\mathbf{w}_3$.

4.6 Incorporating Additional Complexity

In this section, a number of extensions to the basic pandemic model are considered. We consider having,

- insufficient antivirals for the entire population,
- the potential for a hybrid antiviral allocation scheme, that is, a scheme where some fixed proportion of the available antivirals is preallocated while the remainder is kept for dynamic allocation,
- individuals in a household using their antivirals incorrectly, and
- the production of antivirals during a pandemic.

4.6.1 Insufficient Antivirals for an Entire Population

Until now, we have assumed that there is sufficient antivirals for the entire population. In reality, the amount of antivirals available in a population, M, is likely to be smaller than the population size. This can be incorporated into our model under both a dynamic and a preallocation scheme.

For a dynamic allocation scheme the total amount of antivirals used by individuals in the population at time t is,

$$A(t) = N \sum_{(s,e,i,k,a)\in C} (1-\delta_{a,0})kp_{(s,e,i,k,a)},$$

where N is the number of households in the population. It can be determined at each time point whether the supply of antivirals has been exhausted or not by calculating $A(t)$ and comparing to the amount of available antivirals, M. We use two sets of rate vectors,

$$\omega = \{\mathbf{w}_i(t)|i = 0,\ldots,6\}, \tag{4.17}$$

and,

$$\bar{\omega} = \{\bar{\mathbf{w}}_i(t)|i = 0,\ldots,6\}, \tag{4.18}$$

which contain the rates of transition when there is a supply of antivirals and when there is not a supply of antivirals respectively. Note that the only difference between ω and $\bar{\omega}$ is the vector corresponding to the introduction of antivirals, $\mathbf{w}_5$ and $\bar{\mathbf{w}}_5$

respectively. We set $\bar{\mathbf{w}}_5^{(n)} = 0 \; \forall \, n$ which ensures that no more antivirals are introduced into the population. Then, the deterministic approximation when we have insufficient antivirals available for an entire population is,

$$\frac{d\mathbf{p}(t)}{dt} = \begin{cases} \sum_{j=0}^{6} L_j \mathbf{w}_i(t) & \text{if } A(t) < M, \\ \sum_{j=0}^{6} L_j \bar{\mathbf{w}}_i(t) & \text{if } A(t) \geq M. \end{cases} \tag{4.19}$$

When considering the initial condition for this system, we make the assumption that there is sufficient antivirals for use early in the pandemic. This assumption is reasonable, because should the number of antivirals be so small that antivirals cannot be allocated early in the pandemic, the impact that the antivirals will have on the overall pandemic will be negligible. Because of this assumption, the calculation of the initial condition is unchanged from Section 4.5.1.

Next, we consider having insufficient antivirals for the entire population for a pandemic under the preallocation scheme. As in Chapter 3, denote the proportion of households of size k which are preallocated antivirals by ϕ_k. Then, the proportion of households which have not been preallocated antivirals is represented by $(1 - \phi_k)$. The equilibrium point of the system, $\mathbf{p_s}$, now has value $\phi_k h_k$ in each state $(k, 0, 0, k, 3) \in C$ and value $(1 - \phi_k)h_k$ in each state $(k, 0, 0, k, 0) \in C$. For an allocation scheme where we preallocate all available antivirals, we use the rate vectors in the set $\bar{\omega}$, which does not allow for any more antivirals to be introduced into the population. The deterministic approximation for having insufficient antivirals for the entire population under a preallocation scheme then is,

$$\frac{d\mathbf{p}}{dt} = \sum_{j=0}^{6} L_j \bar{\mathbf{w}}_j(t), \tag{4.20}$$

and the initial condition for this system can be determined using the method described in Section 4.5.2, but using the steady state $\mathbf{p_s}$ described above.

4.6.2 Hybrid Allocation Schemes

Thus far, the only allocation schemes that have been considered for antivirals are *pure* allocation schemes, those being when only a dynamic allocation scheme or pre-allocation scheme is used. Consider now a hybrid allocation scheme, where some proportion of the antiviral supply, ξ, is preallocated, and the remainder of the antiviral supply is used under a dynamic allocation scheme. To model this scenario, the ξM available antivirals to the population are preallocated, yielding the proportion of households who have been preallocated antivirals, $\boldsymbol{\phi}$, where the kth element of $\boldsymbol{\phi}$ corresponds to the proportion of households of size k which have been pre-allocated antivirals. Then, utilising the result in Section 4.6.1, the deterministic approximation is the same as in Equation (4.19).

When considering the initial condition for this system, we again assume that there is sufficient antivirals for dynamic allocation early in the pandemic. The equilibrium point of the system, $\mathbf{p_s}$, has value $\phi_k h_k$ in each $(k, 0, 0, k, 3) \in C$ and value $(1 - \phi_k)h_k$ in each $(k, 0, 0, k, 0) \in C$. The Jacobian for this system is identical to that in Section 4.6.1, using the set of rates, ω, from Equation (4.17), which allow for the introduction for antivirals into the population.

4.6.3 Incorrect Use of Antivirals

Currently, the model of the preallocation scheme assumes that each household uses their antivirals correctly. One challenge to a preallocation scheme in practice is that some households are likely to take their antivirals early, particularly due to the less precise method of identification of infection. This feature can be incorporated into our model by including a transition which allows a household to begin taking antivirals incorrectly,

$$(H_{(s,e,i,k,3)}, H_{(s,e,i,k,1)}) \to (H_{(s,e,i,k,3)} - 1, H_{(s,e,i,k,1)} + 1) \text{ at rate } \frac{1}{\psi} H_{(s,e,i,k,3)}, \quad (4.21)$$

where $a = 3$ represents a household which contains individuals who have not yet taken antivirals and $a = 1$ represents a household which contains individuals who

are currently taking antivirals. The transition rate in Equation (4.21) is clearly density dependent, being identical in form to a number of the transition rates we have considered already, and so the deterministic approximation is still a good approximation for the stochastic households model as $N \to \infty$. A seventh matrix and vector is added to the deterministic approximation to represent the states and rates at which the transition in Equation (4.21) can occur. The matrix, L_7, has (m, n)th element,

$$L_7^{(m,n)} = \delta_{k,k'}\delta_{s,s'}\delta_{e,e'}\delta_{i,i'}(-\delta_{a,3} + \delta_{a,1}),$$

with the vector $\mathbf{w}_7(t)$ having nth element,

$$\mathbf{w}_7^{(n)}(t) = \frac{1}{\psi} H_{(s,e,i,k,a)}.$$

Next, we look to calculate the steady state and initial condition. Consider the state of this system with all individuals susceptible and some households being preallocated antivirals. As individuals in a household have the ability to incorrectly use antivirals, the equilibrium state for households who have been preallocated antivirals would be $(k, 0, 0, k, 2)$. That is, each preallocated household will have used their supply of antivirals before the pandemic starts. This is clearly unreasonable, and could be avoided by rapid distribution of antivirals to households as soon as the pandemic starts. To amend this issue, we assume households do not take their antivirals incorrectly before the main growth phase of the pandemic starts, and so when considering the initial condition, we take $\mathbf{w}_7^{(n)}(t) = 0$ for each $(s, e, i, k, a) \in C$. This means that the Jacobian of the system is unchanged by the ability of a household to incorrectly take antivirals, and the equilibrium point $\mathbf{p_s}$ is unchanged also, and so the calculation of the initial condition is identical to that detailed in Section 4.6.1.

4.6.4 Production of Antivirals During a Pandemic

Currently, there is an assumption that the supply of antivirals available during a pandemic is fixed. It is possible, however, that antivirals can be produced during

a pandemic. To model this we start with some fixed number of antivirals, m, that subsequently increases by some amount s per *production period*. Denote the number of *production periods* that have passed by time t by $W(t)$. We change the maximum number of antivirals, M, to a function of t, giving

$$M(t) = m + sW(t).$$

Note that it is possible to define a 'maximum' number of antivirals that can be produced by defining $W(t)$ to have some maximum value independent of time. We preallocate out new antivirals that are produced according to the *size-biased* distribution, $\boldsymbol{\pi}$, in the same way as when initially preallocation antivirals, but now only allow households that have not previously been allocated antivirals.

The addition of the production of antivirals has added no new transitions to the system, but rather has changed which set of transition rates the system evolves by through time, according to whether antivirals can be allocated or not. The deterministic approximation is the same as that defined in Section 4.6.1, but the number of antivirals available is now a function of time, $M(t)$.

As the number of antivirals is fixed at m before the pandemic starts, calculating the Jacobian and the initial condition is identical to the case without the production of antivirals, and can be calculated by simply setting $M = m$ and using the results in Section 4.6.2.

4.7 Summary

In this chapter, a deterministic approximation to the stochastic households model has been derived. The deterministic approximation is fast to compute and gives access to quantities such as the expected final epidemic size, expected peak size, and expected peak time of a pandemic, quantities which are not calculable using the branching process approximation. Because of the fast computation speeds of the deterministic approximation, a number of extensions have been considered including

having insufficient antivirals for the population, hybrid allocation schemes, households using antivirals incorrectly and the potential production of antivirals during a pandemic. However, it is not trivial to incorporate a constant-time period in the deterministic approximation compared to the branching process approximation. In the following chapter, we apply the deterministic approximations and the extensions to the parameters described in Section 2.9. We compare the dynamic antiviral allocation scheme to the preallocation scheme and look to determine which of these antiviral allocation schemes is more effective at controlling pandemic influenza.

Chapter 5

Comparison of Antiviral Allocation Schemes

The Australian Health Management Plan for Pandemic Influenza [22] currently specifies that antivirals would be utilised according to a dynamic allocation scheme in the event of an influenza pandemic. In this chapter, we investigate the impact of antivirals on an influenza pandemic when allocated according to a dynamic allocation and a preallocation scheme. We also investigate the effect of the various extensions discussed in Chapter 4: that is, allowing individuals in a household to use antivirals incorrectly; a hybrid allocation scheme; and the production of antivirals during a pandemic. The key quantities associated with the pandemic, those being the expected Malthusian parameter, r, the expected peak size and time of a pandemic, and also the expected final epidemic size, are compared in order to help determine whether a dynamic allocation scheme or a preallocation scheme is more effective.

5.1 Dynamic Allocation

The first comparison is between a pandemic without antiviral intervention and a pandemic under a dynamic allocation scheme. The reason for this comparison is due to the discussion about the effectiveness of antivirals [43]. In particular, Jefferson *et*

al. [43] note that the effectiveness of antivirals is not fully known due to the lack of detailed clinical trial data which is publicly available. The population size is fixed, at $N\bar{k} = 10^5$, but we consider the full range of possible antiviral availability, from $M = 0$ to $M = 10^5$. Figure 5.1 shows the comparison of the final epidemic size when dynamic allocation is used, in blue, compared to having no antivirals available, in magenta and the preallocation scheme, in red. The benefits of utilising antivirals during an influenza pandemic can be clearly seen, even though we have assumed the effectiveness of the antivirals to be lower than other studies [52, 43]. There is up to a 14% difference in the final epidemic size for a *mild* outbreak. In the *severe* parameter case, the impact of antivirals is still significant, at up to 8%, however, the antivirals have less impact when compared to the *mild* parameter case. This is because the rate of infection is higher in the severe parameter case, while the delay until antivirals arrive into an infected household is the same in both cases. Hence, more infection transmission occurs before antivirals arrive in the *severe* parameter case, and so the impact of the delay under a dynamic allocation scheme is smaller. The other noticeable feature is the existence of a *saturation* point—the point at which adding more available antivirals to the pandemic has no effect. The reason for this is simple: at some point, there are enough available antivirals for every household which becomes infected. As such, having more antivirals available than this amount will not impact the pandemic. The saturation point demonstrates that there is likely to be a maximum effective stockpile necessary for use in an influenza pandemic under a dynamic antiviral allocation scheme. A key factor in this assumption is the *delay* until antivirals arrive into a household after the first infection event [15, 47]. Should the delay be large, then the antivirals will have negligible impact. Recall that an estimate of this delay from the 2009 Swine 'Flu pandemic suggested that $\zeta = 1$. In Figure 5.2, the expected final epidemic size is plotted as a function of the average delay, $\zeta \in (0, 4]$. At an average delay of 1 infectious period, the expected final epidemic size is approximately 57% of the population, while at an average delay of half an infectious period, the expected final

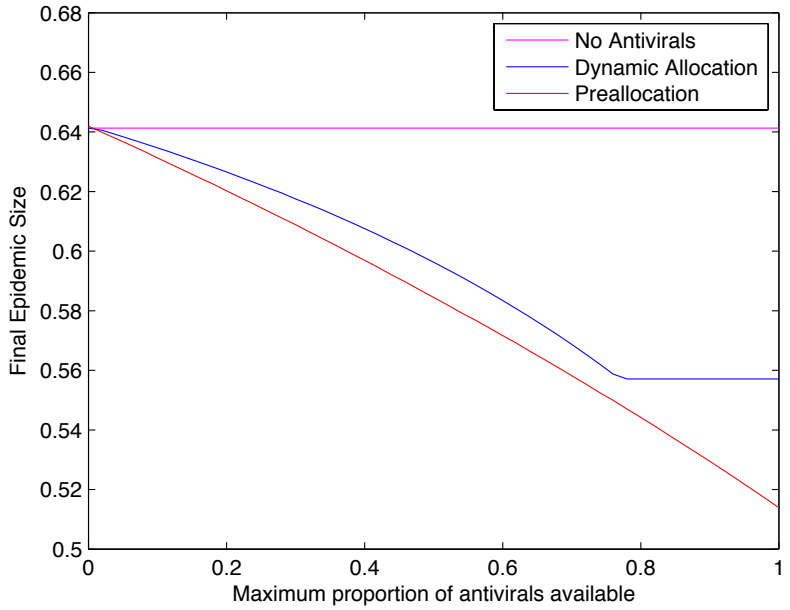
epidemic size is approximately 55% of the population. This change in expected final epidemic size is relatively small, especially compared with the change in Malthusian parameter, r, from Section 3.3.1.

5.2 Comparing Dynamic Allocation and Preallocation

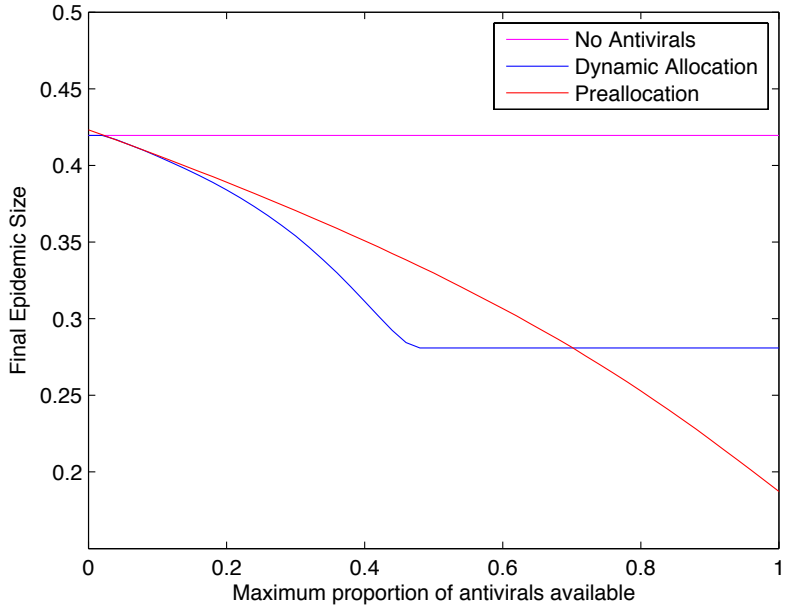
In this section, we compare pandemics under a dynamic allocation scheme and a preallocation scheme. Initially, we consider these antiviral allocation schemes without any additional modifications, and then later investigate the effects of the incorrect use of antivirals by individuals in households, hybrid allocation schemes, and the production of antivirals during a pandemic.

5.2.1 Dynamic Allocation vs Preallocation

We now investigate the difference between a pandemic under dynamic allocation and under preallocation. Figure 5.1 shows the final epidemic sizes for pandemics under both a dynamic allocation scheme and a preallocation scheme. For the case with the *severe* parameters, preallocation outperforms dynamic allocation in terms of final epidemic size regardless of the maximum amount of available antivirals. For the case with *mild* parameters, we can see that dynamic allocation outperforms preallocation until approximately 70% of the population has antivirals available. The reason that dynamic allocation outperforms preallocation during a mild outbreak is because the rate at which infection spreads is lower relative to the delay until antivirals arrive, and so more antivirals can arrive to households before other members become infectious. For large amounts of available antivirals, preallocation again gives an final epidemic size which is up to 10% smaller than the final epidemic size under a dynamic allocation scheme. This is because the benefit of having no delay until antivirals arrive outweighs the fact that there are some antivirals which are not



(a) *Severe* parameter set



(b) *Mild* parameter set

Figure 5.1: Comparisons of expected final epidemic size without antiviral intervention, and with both a dynamic allocation scheme and a preallocation scheme, for a range of maximum available antivirals. The parameter sets are taken from Table 2.3.

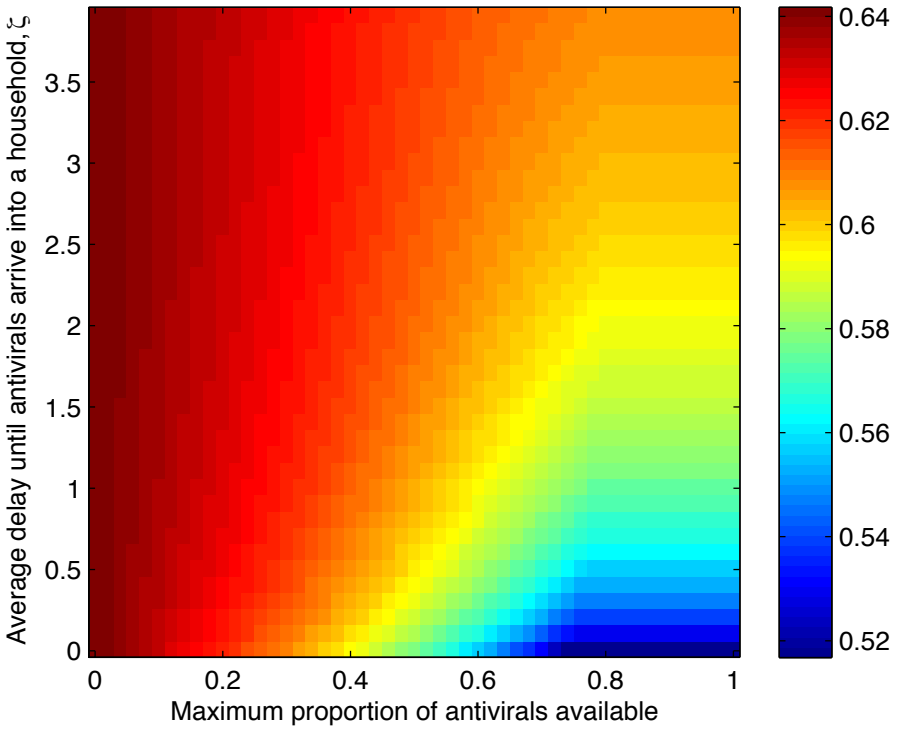


Figure 5.2: The effects of the average delay until antivirals arrive into a household, ζ, and the proportion of the population which have antivirals available, on the expected final epidemic size under a dynamic allocation scheme, using the *severe* parameter set from Table 2.3.

being used. Note, however, that when a dynamic allocation scheme is superior to a preallocation scheme, the largest difference between expected final epidemic sizes is approximately 5%, which is approximately half of the maximum possible gain of a preallocation scheme. This result indicates that selecting an antiviral allocation scheme depends on the infectiousness of the influenza strain as well as the amount of available antivirals. If the outbreak is severe, then a preallocation scheme would be better. However, for a mild outbreak, a dynamic allocation scheme may be the better choice.

Comparisons of the expected Malthusian parameter, r, expected peak size, and expected peak time for a pandemic using the *severe* parameter set are shown in Figure 5.3. Recall that a lower early growth rate and peak size is preferable, while a higher peak time is desired in order to give a longer time to control the pandemic. For the Malthusian parameter, r, we can see that the dynamic allocation scheme is outperforming the preallocation scheme until approximately 70% of the population has antivirals available. This is contrary to the results seen for final epidemic size where the preallocation antiviral scheme led to a lower final epidemic size than the dynamic allocation scheme across the entire range of available antivirals. Similarly, the expected peak time and expected peak size of a pandemic are also superior under a dynamic allocation scheme until approximately 70% of the population has antivirals available. This result is relatively intuitive; with any feasible number of antivirals, dynamic allocation guarantees that infected households will receive antivirals early in the pandemic and so it is no surprise that under a dynamic allocation scheme, early time quantities such as the Malthusian parameter, r, and peak size and peak time are all improved compared to preallocation. The preallocation scheme, however, ensures that some, roughly constant, proportion of infectious individuals are taking antivirals throughout the pandemic and this tends to lead to a smaller expected final epidemic size for a *severe* pandemic outbreak.

Figure 5.2 shows the effect of the average delay until antivirals arrive into a household, ζ, on the expected final epidemic size, for a fixed mean effective duration

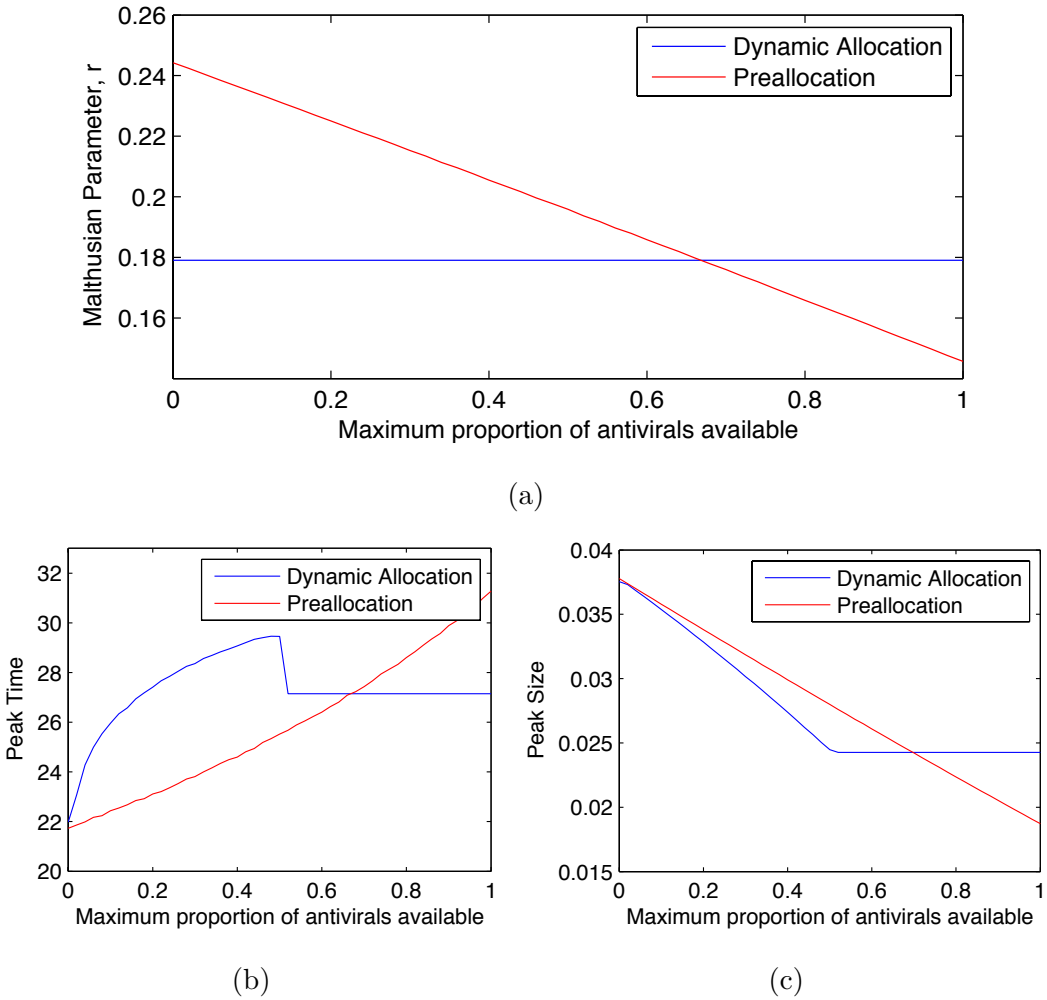


Figure 5.3: Comparisons of the expected Malthusian parameter, r, expected peak time, and expected peak size for an epidemic under a dynamic allocation scheme and a preallocation scheme using the *severe* parameter set from Table 2.3.

of antivirals, $\kappa = 1$. Only the impact on a dynamic allocation scheme needs to be considered here, as there is no delay until antivirals arrive into a household under a preallocation scheme. The figure shows that it is somewhat possible to compensate for a larger delay until antivirals arrive into a household, ζ, by having more antivirals available for the population. Similarly to when considering the final epidemic size for a fixed delay, there exists a *saturation point* of available antivirals, which is when approximately 75% of the population has antivirals available. Having more antivirals available than this amount does not affect the final epidemic size. This is relatively consistent across the range of tested values of ζ. Importantly, the figure shows that the effect of the average delay until antivirals arrive into a household, ζ, on expected final epidemic size, is dependent on the maximum proportion of antivirals available. When approximately 40% of the population has antivirals available, then the difference in expected final epidemic size when $\zeta = 0$ compared to when $\zeta = 4$ is approximately 3%. Comparatively, if 100% of the population has antivirals available, then the difference in expected final epidemic size is approximately 16%.

In Figure 5.4, the effect of the finite effective duration of antivirals, κ, on the expected final epidemic size for a fixed mean delay until antivirals arrive into a household, $\zeta = 0.5$ is shown. The colour in this figure is the difference between the expected final epidemic size under a dynamic allocation scheme and a preallocation scheme. A negative value shows that the dynamic allocation scheme leads to a smaller expected final epidemic size, while a positive value shows that the preallocation scheme leads to a smaller expected final epidemic size. It can be seen that, for a majority of the range of average effective antiviral durations, the preallocation scheme yields a smaller expected final epidemic size than the dynamic allocation scheme. The exception to this is when approximately 65% to 75% of the population has antivirals available, and the mean effective duration of these antivirals, κ, is between 2.5 and 4 infectious periods, which is approximately 7.5 to 12 days. In this case, the dynamic allocation scheme gives an expected final epidemic size which is at most 1.6% lower than the preallocation scheme. This indicates that there may

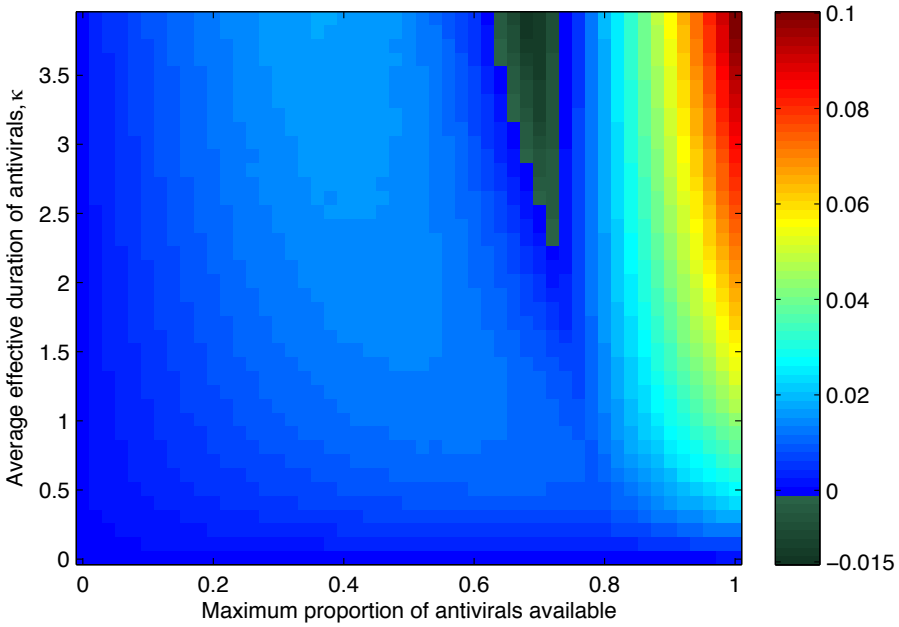


Figure 5.4: The difference between the expected final epidemic size under a dynamic allocation scheme and a preallocation scheme as a proportion of the population, depending on the finite effective duration of antivirals, κ, and the proportion of the population which has antivirals available, using the *severe* parameter set from Table 2.3.

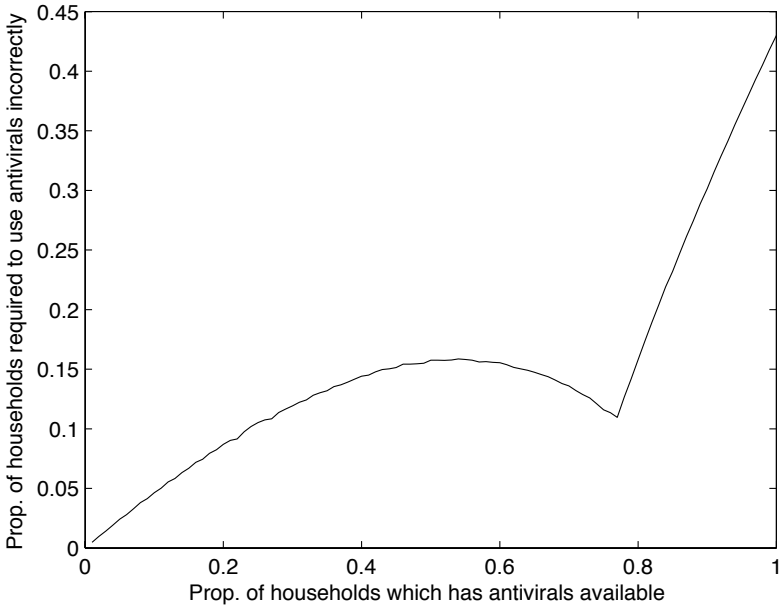
be merit to allocating long courses of antivirals over the course of a pandemic under a dynamic allocation scheme. Note, however, that the length of time for which antivirals are taken under the Australian Health Management Plan for Pandemic Influenza [22] is 4 to 6 days. In this region, the preallocation scheme yields a consistently smaller expected final epidemic size than the dynamic allocation scheme, regardless of the mean effective duration of antivirals, κ.

It is important to remember that at this point no additional features have been included in the model. In particular, we have assumed that the individuals in every household that has been preallocated antivirals will use their antivirals correctly. This is clearly a strong assumption, however, if every household was to use their antivirals correctly the preallocation scheme would be the better scheme in terms of

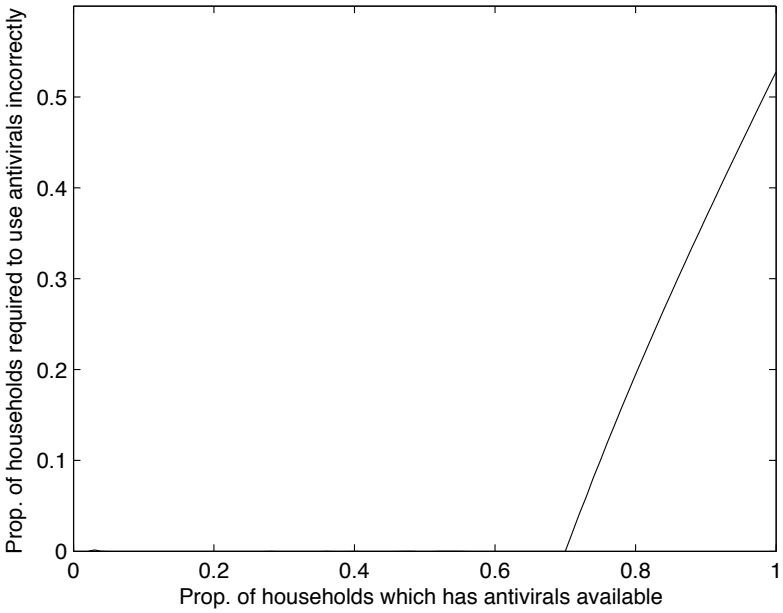
expected final epidemic size for a *severe* pandemic outbreak.

5.2.2 Incorrect Use of Antivirals

The main assumption in the previous section for the preallocation scheme is that all antivirals are used 100% effectively. Recall from Section 4.6.3 the extension to the base model where individuals in a household begin taking antivirals early at some *rate of false taking*, ψ. It is unknown what this rate would be in reality, but the number of households which contain individuals who would take antivirals incorrectly for a given ψ can be calculated by considering the proportion of households in state $(s, e, 0, k, 1)$. This is possible because a household can only be in state $(s, e, 0, k, 1)$ if the individuals inside the household take their antivirals incorrectly. The proportion of households required to incorrectly use their antivirals in order for the dynamic allocation scheme to be better than the preallocation scheme in terms of expected final epidemic size can be seen in Figure 5.5. Note that in Figure 5.5 (b), there is a large section where zero households must contain individuals who use antivirals incorrectly. This represents the fact that a dynamic allocation scheme is outperforming a preallocation scheme when all antivirals are used 100% effectively. Importantly, if between 30% and 70% of households have antivirals available and the outbreak is severe, then somewhere in the region of 10% to 15% of households would need to use antivirals incorrectly for a dynamic allocation scheme to lead to the same expected final epidemic size as the preallocation scheme. For the less severe pandemic, there is more evidence that a dynamic allocation scheme would be preferable. In particular, note that until approximately 80% of the population has antivirals available, no more than 10% of households in the population must contain individuals who use antivirals incorrectly. When a large proportion of the population has antivirals available, however, the required proportion of households who must use antivirals incorrectly increases rapidly.



(a) *Severe* parameters



(b) *Mild* parameters

Figure 5.5: Required proportion of households who use antivirals incorrectly for a dynamic scheme to be preferable to a preallocation scheme in terms of final epidemic size, for a (a) *severe* outbreak and a (b) *mild* outbreak. All parameters are taken from Table 2.3.

5.2.3 Hybrid Schemes

Another potential antiviral allocation scheme is a *hybrid* allocation scheme, that is, a scheme in which some proportion of antivirals are preallocated, while the remaining antivirals are reserved for dynamic allocation. Figure 5.6 shows the difference between a hybrid scheme and the best pure scheme, in terms of expected final epidemic size. We see that the difference between the hybrid allocation scheme and the best pure scheme is small until a relatively high proportion of the population has antivirals available. When a large proportion of the population has antivirals available, however, a hybrid scheme does not perform as well as the preallocation scheme. This is because when a small amount of antivirals is preallocated, and the remainder reserved for dynamic allocation, a hybrid scheme uses antivirals very similarly to a dynamic allocation scheme. Recall from Figure 5.1 that, for a severe outbreak, a preallocation scheme yielded a smaller expected final epidemic size than a dynamic allocation scheme. Thus, it is expected that a hybrid scheme which reserves a large amount of antivirals for dynamic allocation does not perform as well as the preallocation scheme. Note importantly that a hybrid allocation scheme does not yield a smaller expected final epidemic size than the superior pure scheme. It is possible that a hybrid allocation scheme can lead to a slightly smaller expected final epidemic size for extreme sets of parameters. However, a pandemic that follows these types of parameters is far from realistic.

5.2.4 Production of Antivirals During a Pandemic

Recall from Section 4.6.4 that it is possible to model the fact that antivirals could be produced during a pandemic, by effectively setting the amount of available antivirals, M, to be a function of time, $M(t)$. Figure 5.7 shows the effects that the production of antivirals during a pandemic could have. In Figure 5.7, the initially available amount of antivirals is assumed to be one *production* amount. We can see that there is a production amount in which dynamic allocation leads to a smaller

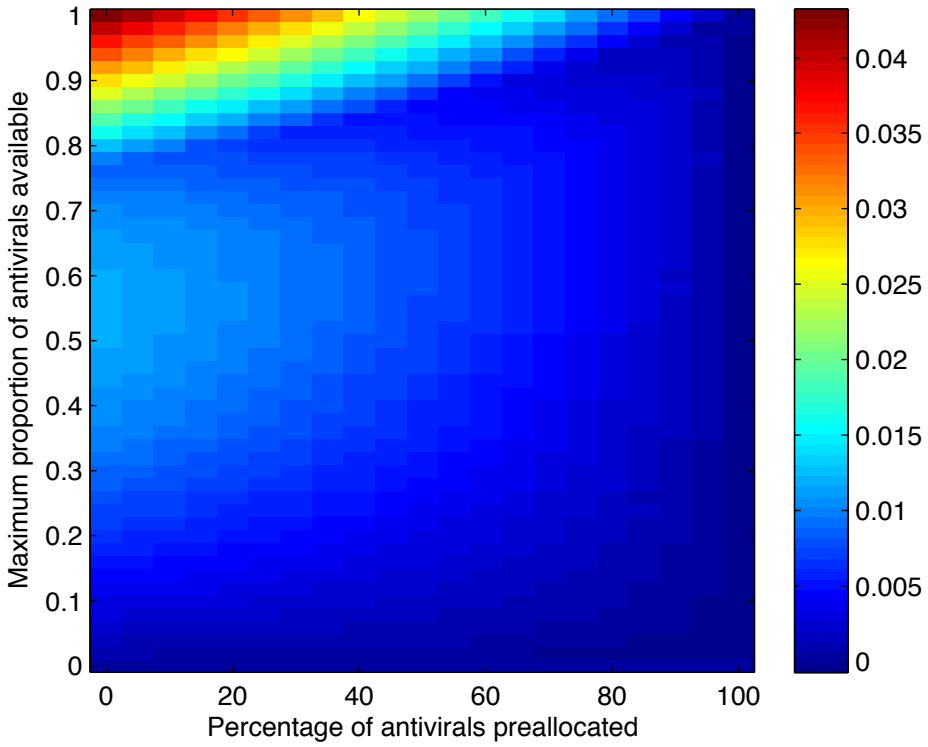


Figure 5.6: The difference in expected final epidemic size as a proportion of the total population, between a hybrid scheme and the best pure scheme using the *severe* parameter set from Table 2.3 with $M = 10^5$ available antivirals.

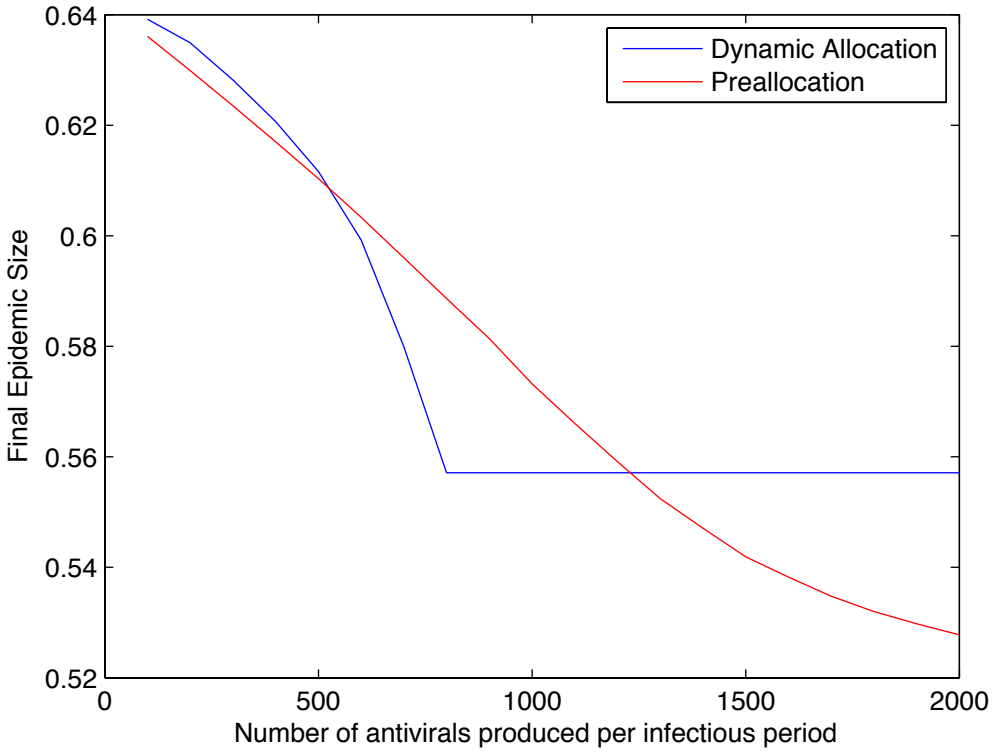


Figure 5.7: The effect of the production of antivirals throughout a pandemic on the expected final epidemic size, using *severe* parameters from Table 2.3.

expected final epidemic size than preallocation. After approximately $1,200$ antivirals are produced per time period, the preallocation scheme again leads to a smaller expected final epidemic size. This indicates that if the number of available antivirals is small, then a dynamic allocation scheme may be preferable to a preallocation scheme. If a large proportion of the population has antivirals available, though, then a preallocation scheme still leads to a smaller expected final epidemic size than the dynamic allocation scheme for a severe outbreak.

5.2.5 Sensitivity to Household Size Distribution

All the results presented in this chapter have the household size distribution, h_k, representing that of Australian households. We also investigate two other household size distributions which represent Indonesian [5] and Sudanese households [15]. Both

of these countries tend to have larger household sizes than Australia, with Sudan being the highest. The distribution of household sizes for Sudan is,

$$\mathbf{h^S} = [0.0467, 0.0705, 0.0958, 0.1080, 0.1187, 0.1153, 0.1114, 0.0948, \\ 0.0681, 0.0559, 0.0389, 0.0277, 0.0170, 0.0102, 0.0209],$$

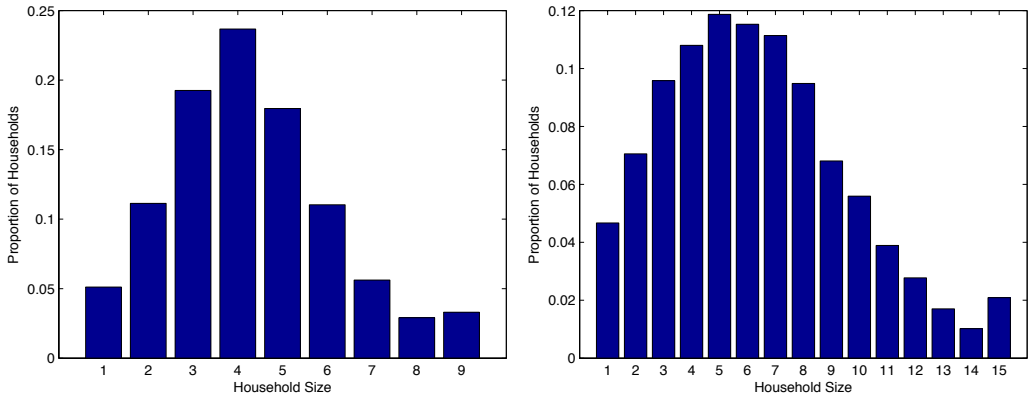
and for Indonesia is,

$$\mathbf{h^I} = [0.0512, 0.1113, 0.1926, 0.2367, 0.1795, 0.1103, 0.0562, 0.0291, 0.0331].$$

Both of these household size distributions are shown in Figure 5.8. Figure 5.9 shows the expected final epidemic size for a range of available antivirals, across the household size distributions of Australia, Indonesia and Sudan for a severe pandemic outbreak. It can be seen that while the household size distribution shifts the graphs, it does not generally alter the fact that a preallocation scheme consistently yields a smaller expected final epidemic size than the dynamic allocation scheme. The effects of each antiviral scheme are qualitatively similar, and so we conclude that these results seem to hold across a range of household size distributions. Note that even though the population size in each of these pandemics is fixed at $N = 10^5$, the expected final epidemic size tended to increase with the mean household size of the population. This result has been seen previously [39]. Having more members inside a household means that a single infectious individual is likely to, on average, create more secondary infectious cases, when compared to an individual in a smaller household, due to the higher mixing rate inside a household compared to in the general population. This larger amount of secondary infectious cases leads to the higher expected final epidemic size that is observed.

5.3 Summary

In this chapter, comparisons have been made between the dynamic allocation scheme and the preallocation scheme. First, it was seen that even under a dynamic allocation scheme, antivirals can provide a substantial reduction in the expected final



(a) Indonesia (b) Sudan

Figure 5.8: Household size distribution of (a) Indonesia and (b) Sudan

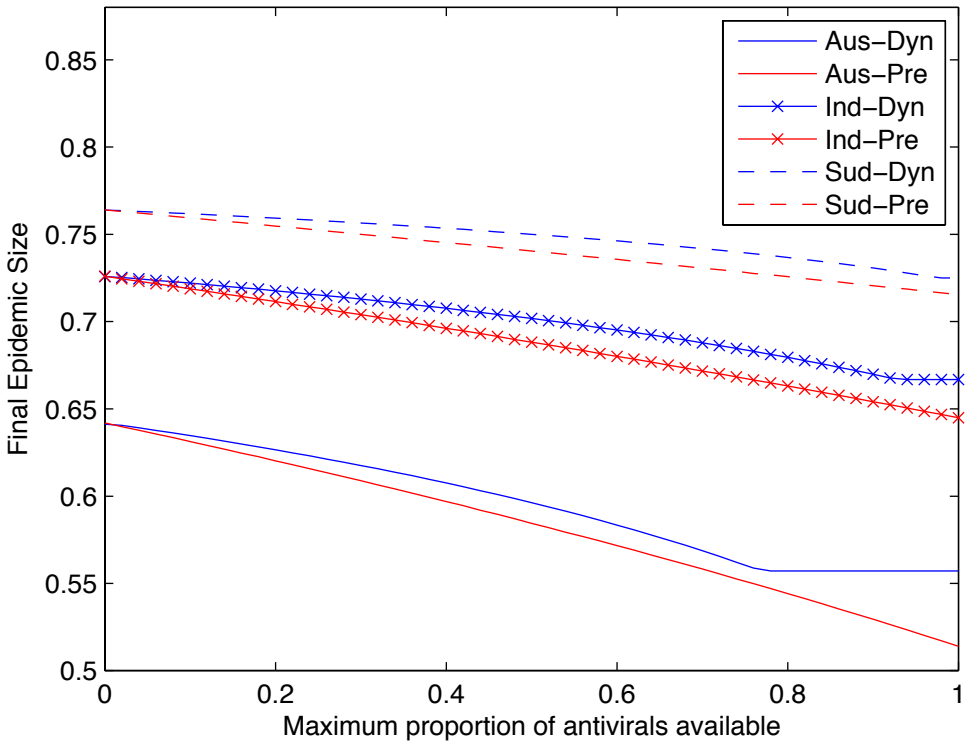


Figure 5.9: The effect of household size distribution on the expected final epidemic size for a range of antivirals available for the population. All parameters, except for the household size distribution, are taken from the *severe* parameter set from Table 2.3.

epidemic size. Also demonstrated was that, for a severe pandemic, a preallocation scheme leads to a lower expected final epidemic size than a dynamic allocation scheme. Contrary to the expected final epidemic size, however, the expected Malthusian parameter, r, expected peak size, and expected peak time indicated that until a large proportion of the population has antivirals available, a dynamic allocation scheme would be better. This suggests that the preferable allocation scheme depends on which key quantity is the focus of pandemic control. It is worth noting that extremely high stockpiles of antivirals are somewhat unrealistic and expensive [18, 30]. Also considered were extensions to the stochastic households model. We demonstrated that, for a severe pandemic outbreak, somewhere between 10% and 15% of households must contain individuals who incorrectly use antivirals for the dynamic allocation to give a smaller expected final epidemic size than the preallocation scheme. We also showed that a hybrid allocation scheme does not produce a smaller expected final epidemic size than a pure scheme. Finally, we also showed that changes in the household size distribution do not appear to qualitatively alter the expected final epidemic size. These results indicate that a preallocation scheme could potentially lead to a smaller expected final epidemic size, particularly for a severe outbreak.

Chapter 6

Conclusion

In this thesis, two potential antiviral allocation schemes, dynamic allocation and preallocation, have been compared in order to determine which allocation scheme is better in terms of a number of performance measures. To facilitate this, two approximations for the *stochastic household model* have been considered and extended.

6.1 Approximations

In Chapters 2 and 3, the *branching process* approximation was introduced and extended. While branching processes have been used to assess the severity of an influenza outbreak previously [71], and also used to determine the effectiveness of antivirals [15], the finite duration of antivirals has not been previously explored in this context. Many simulation studies incorporate a finite duration of antivirals [53, 28, 14], however, this finite duration is assumed fixed, and the effect has not been assessed. We have demonstrated, using a branching process, that the finite duration of antivirals can have significant impact on the Malthusian parameter, r. For example, if the average delay until antivirals arrive in the household is exponentially distributed with a mean of 1 infectious period, then having a the effective duration of antivirals be exponentially distributed with a mean of 5 infectious periods gives a Malthusian parameter which is approximately 25% lower than the situation where

the mean effective duration is 1 infectious period. This implies that selecting an antiviral allocation scheme depends not only on the delay until antivirals arrive into a household after the first infection event, as has been determined previously [15, 27], but also on the effective duration of these antivirals. Note that as the mean effective duration, $\kappa \to \infty$, there is an approximately 36% reduction in the Malthusian parameter, compared to the case with no antiviral intervention. This estimate is similar to the result obtained by Black *et al.* [15]. In the case when $\kappa = 1$, the Malthusian parameter is reduced by approximately 17%, in line with estimates from Ghani *et al.* [29]. This reaffirms that the finite duration of antivirals is an important consideration, even for a quantity associated with the early stages of the pandemic such as the Malthusian parameter.

We also considered having the delay until antivirals arrive into the household after the first infection event, and the effective duration of antivirals, being of constant length as opposed to being exponentially distributed. The results in Section 3.4 demonstrate that assuming an exponentially distributed period gives a generally higher Malthusian parameter as opposed to a constant duration, in line with previous results [71, 15]. Note importantly that the predominant impact of the effective duration of antivirals is to lower the Malthusian parameter, independent of the distribution of this effective duration.

The main issue with using a branching process approximation is that it cannot be easily used to calculate quantities such as the final epidemic size, or peak time and peak size. It is the required assumption for a branching process that the population size is infinite and so infection happens into a naive, or fully susceptible household, that leads to the fact that the branching process cannot easily approximate the long-term behaviour of a pandemic. As discussed throughout this work, the dynamic allocation scheme ensures that every household which experiences an infection event early in the pandemic will receive antivirals, while the preallocation scheme ensures that some, roughly constant, proportion of infectious individuals is taking antivirals throughout the pandemic. The benefits of a preallocation scheme then may not

be seen in quantities associated with the early stages of a pandemic, such as the Malthusian parameter.

In Chapter 4, a *deterministic* approximation is derived to calculate the expected final epidemic size, expected peak size, and expected peak time of a pandemic. One limiting factor to the deterministic approximation is that we assume that all event times are exponentially distributed. While other distributions can be incorporated through systems of delay differential equations [1], they have not been considered in this work. The deterministic approximation utilises an alternative representation of the stochastic households model. This representation means that households are no longer uniquely identifiable, unlike in the representation described in Chapter 2. This alternative representation condenses the state space of the process, making the model more tractable [16]. When formulated in this way, we have shown that the stochastic households model is a *density-dependent* process in the sense of Kurtz [49], and so the deterministic approximation is accurate in the limit as the number of households $N \to \infty$ [8]. This type of deterministic approximation has been utilised previously to assess the impact of antivirals, however, these models have either not assumed household structure [34] or have not focussed on the method of antiviral distribution [16, 39], both of which are contained within this work. As the state space for the deterministic approximation is relatively small, a number of extensions to the stochastic households model were also incorporated. These extensions are: having insufficient antivirals for an entire population; members of households using their supply of antivirals incorrectly; and the production of antivirals during an epidemic.

6.2 Comparison of Antiviral Allocation Schemes

In Chapter 5, the two antiviral allocation schemes, dynamic allocation and preallocation, were compared. Firstly, it was noted that a dynamic allocation scheme, which is the antiviral allocation scheme that would currently be utilised in the event of an influenza epidemic in Australia [22], has a significant impact on final epidemic

size for both a mild and severe epidemic outbreak. A dynamic allocation scheme reduced the final epidemic size by up to 14% in the mild case and 8% in the severe case. This reduction is small in comparison to reductions in the Malthusian parameter considered both here and by others [15, 29]. One factor which is often not considered is the impact that the delay until antivirals arrive into a household and the effective duration of antivirals has on the expected final epidemic size. We showed that it is possible to compensate somewhat for the delay until antivirals arrive into a household by having more antivirals available for the population, although this compensation is clearly limited. We also showed that the effect of a finite duration on antivirals can be profound, even for quantities associated with the early stages of the pandemic, such as the Malthusian parameter. When assuming a long mean effective duration of antivirals, we showed in Chapter 3 that antivirals reduced the Malthusian parameter by approximately 30% for a mean delay of 1 infectious period, while assuming a small mean effective duration of antivirals only reduced the Malthusian parameter, by approximately 17%. Both of these percentages have been determined previously [15, 29], and this result demonstrates the importance of the finite effective duration of antivirals.

One important consideration for any antiviral allocation scheme is the proportion of the population which has antivirals available. Many studies in the past assume that the entire population has access to antivirals [14, 15, 53], however, should this not be the case, then the effectiveness of antivirals can be significantly impacted. We demonstrated that for a mild outbreak, the allocation scheme which produced the smaller expected final epidemic size was dependent on how many antivirals were available. Should more than 70% of the population have antivirals available, then a preallocation scheme gives a smaller expected final epidemic size, while if less than 70% of the population has antivirals available, then a dynamic allocation scheme gives a smaller expected final epidemic size. For a severe outbreak, however, a preallocation scheme always gives a lower expected final epidemic size than a dynamic allocation scheme, regardless of how many antivirals are available. Interestingly,

when considering just the Malthusian parameter, this is not the case. Even for a severe outbreak, the dynamic allocation scheme gives a smaller Malthusian parameter if less than 70% of the population has antivirals available. This occurrence justifies the consideration of the expected final epidemic size, as opposed to just the Malthusian parameter when comparing the antiviral allocation schemes. The dynamic allocation scheme also gives a lower expected peak size, and higher expected peak time, when less than 70% of the population has antivirals available, similarly to the Malthusian parameter. Again, this is because the dynamic allocation scheme ensures that all individuals inside households which experience an infection event early in the pandemic have access to antivirals, and so the quantities associated with the early stages of the pandemic are reduced compared to preallocation. A preallocation scheme, however, ensures that some, roughly constant, proportion of infectious individuals is taking antivirals throughout the pandemic, and in a severe outbreak, this leads to a smaller expected final epidemic size than the dynamic allocation scheme. This result highlights the importance of considering a number of key quantities associated with an epidemic when choosing which antiviral allocation scheme is best. Unfortunately, with nine parameters in the model, seven of which must be estimated based on pandemic data, obtaining good predictions of quantities associated with a pandemic would prove challenging in practice. It is worth noting that the cost of maintaining large stockpiles is estimated to be high [18], and so in reality, a country is unlikely to have stockpiles as high as 70% of the population.

The main assumption in the preallocation scheme is that individuals inside a household always use their antivirals correctly. In reality, this is unlikely to be true [28]. We showed that, for a severe outbreak, if the proportion of the population which has antivirals available is greater than 20%, then more than 10% of the population would have to use antivirals incorrectly for a dynamic allocation scheme to give a lower expected final epidemic size than a preallocation scheme. For a mild outbreak, though, a dynamic allocation scheme already gives a lower expected final epidemic size than a preallocation scheme until approximately 70% of the population has

antivirals available. After this point, however, the proportion of households required to use antivirals incorrectly increases steeply. These results again suggest that should the amount of antivirals available for the population be high, then a preallocation scheme would lead to a lower expected final epidemic size than a dynamic allocation scheme, even when individuals in a household are able to use antivirals incorrectly.

Another potential allocation scheme is a hybrid allocation scheme. A hybrid allocation scheme preallocates some proportion of the antiviral stockpile, and reserves the rest of the antivirals for a dynamic allocation scheme. We demonstrated that, for feasible parameters, a hybrid allocation scheme does not lower the expected final epidemic size compared to a pure allocation scheme for both a mild and severe pandemic.

In other studies, the supply of antivirals is assumed to be infinite [15, 27] or fixed [24, 28, 29, 55]. In this work, we have considered the effects of a production of antivirals during a pandemic. We showed that the production of antivirals can have a noticeable impact on expected final epidemic size for a severe pandemic outbreak. The expected final epidemic size under preallocation scheme in particular was increased noticeably, even when the amount of antivirals produced per infectious period was relatively small. While no data is available with respect to the capacity of antiviral production, we expect that this number will be large enough to ensure that there is sufficient antivirals for the dynamic allocation scheme. With our parameters, sufficient antivirals for a dynamic allocation scheme is 0.08% of the population per day. We demonstrated that producing more antivirals beyond this amount will not change the expected final epidemic size under a dynamic allocation scheme, but will only improve the expected final epidemic size under a preallocation scheme.

This work has shown overall that for a severe pandemic outbreak, a preallocation scheme could yield a smaller expected final epidemic size than a dynamic allocation scheme. However, quantities such as the expected Malthusian parameter, expected peak size, and expected peak time may indicate that a dynamic allocation scheme would be preferable. For a mild pandemic outbreak, this work has demonstrated

that a preallocation scheme would yield a lower expected final epidemic size than preallocation only if the amount of antivirals available for the population is high, otherwise a dynamic allocation scheme would yield the lower expected final epidemic size.

6.3 Limitations and Potential Extensions

While the results throughout this thesis provide a broad comparison of antiviral allocation schemes for an influenza pandemic, there are some limitations of the model which has been utilised. The model that has been utilised throughout this work is the SEIR model, as opposed to the SIR model which is commonly used for influenza [16, 39, 66]. However, the model could be further refined for a better approximation. This would entail adding additional phases to the model, which will increase the size of the state space. The deterministic approximation is fast to compute because of a relatively small state space, but some extension should not affect this computation time too drastically. For example, adding an additional exposed and infectious phase would mean that the distribution for the exposed and infectious periods would follow an Erlang-2 distribution [2, 66], which is more in line with observations from the 2009 Swine 'Flu pandemic [29]. This extension, however, lifts the number of possible household configurations for a household of size 3 from 20 to 56.

One key limitation of our model is the assumption that there are negligible births and deaths throughout the duration of the pandemic. While death is somewhat similar to recovery in terms of a modelling assumption, in that the individual has no further impact on the pandemic, the minimisation of death through a pandemic has been a focus of other studies [33, 34]. We expect that minimisation of final epidemic size would also contribute to a lower number of deaths, however, this concept is not verified here. The inclusion of births has been shown to play an important role in long-term models [33], however, this work is focussed on the impact of a single

pandemic, the likely span of which is a few months, and so the amount of births is unlikely to have a substantial impact.

Another limitation of the model utilised in this work is that there is only a single strain of influenza. This is closely linked to the time-scale for which these comparisons are relevant, however, consideration of multi-strain dynamics may prove important for a large-scale pandemic [58]. This extension may also require having the transition rates being time-dependent rather than constant throughout the pandemic. The analytical form of the Jacobian may no longer exist, however, the differential equations should still be solvable numerically.

In this work, we have assumed *homogeneous mixing* at a population level. It would be possible to overlay a network structure into the model, similarly to other studies [69]. This would allow for more targeted antiviral distribution to people who have recently come in contact with an infectious individual, in a similar vein to *contact tracing* during a pandemic [40, 67, 77].

The model could be extended to further accurately represent a population by including a third phase of mixing which could represent a workplace or school environment. This third level would add significant complexities to the stochastic households model, but would allow exploration into the effects of closing schools and workplaces, which is a control measure discussed in the Australian Health Management Plan for Pandemic Influenza [22]. The closing of workplaces and schools has previously been shown to reduce the severity of a pandemic by 30% to 70% [38]. Incorporating this extension would also allow testing of the two antiviral allocation schemes—dynamic allocation and preallocation—to determine the effects of the delay until antivirals arrive with a more detailed transmission model. Preliminary work based on the concept of large *patches* has been undertaken [16], but as of yet, there have been no approximations for models consisting of more than two levels of mixing.

Another potential extension to this work is to investigate different methods of preallocating the antivirals, in a similar vein to the optimal vaccination policy ques-

tion [12, 54, 62]. In this thesis, the antivirals are preallocated to entire households in the population according to the distribution of household sizes, **h**. A potential investigation would be to change this method, say by giving antivirals to all the large households first, or according to the size-biased distribution, and testing the effects.

Also of interest is to calculate the *diffusion approximation* for the mean pandemic. This would allow for an approximation of the variance of the process which could help for planning against worst-case and best-case pandemic scenarios. The diffusion approximation for an SIR model has been completed [16], however, this is yet to be extended to the SEIR model that was utilised throughout this work. A method which could potentially be applied to the SEIR model has been established [64], however, expressions for the diffusion approximation for the SEIR stochastic households model have not yet been studied.

New strains of influenza have caused pandemics approximately every 30 years. Events of the past would indicate that control of future pandemics is of utmost importance. Further research is needed to fully understand the best use of antivirals. In this thesis, we have shown that a potential new allocation scheme, known as preallocation, could have potential benefits for pandemic control. The preallocation scheme yields a consistently smaller final epidemic size than the dynamic allocation scheme for a *severe* pandemic outbreak. However, a dynamic allocation scheme is more robust in terms of the number of antivirals available for the population and the severity of the pandemic outbreak. The extensions and ideas presented throughout this work should lead to a more efficient use of antivirals, leading to a smaller impact of future pandemics.

Bibliography

[1] M. E. Alexander, S. M. Moghadas, G. Röst, and J. Wu. A delay differential model for pandemic influenza with antiviral treatment. *Bulletin of Mathematical Biology*, 70(2):382–397, 2008.

[2] D. Anderson and R. Watson. On the spread of a disease with gamma distributed latent and infectious periods. *Biometrika*, 67(1):191–198, 1980.

[3] R. M. Anderson and R. M. May. *Infectious diseases of humans: dynamics and control*. Oxford and New York: Oxford University Press, 1991.

[4] Australian Bureau of Statistics. NPRD Number of Persons Usually Resident in Dwelling. *Census 2011*, 2011.

[5] Badan Pusat Statistik-Statistics Indonesia (BPS) and ORC Macro. Indonesia Demographic and Health Survey 2002-2003. *Calverton, Maryland, USA: BPS and ORC Macro*, 2003.

[6] F. 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] F. Ball. Threshold behaviour in stochastic epidemics among households. In C. Heyde, Y. Prohorov, R. Pyke, and S. Rachev, editors, *Athens Conference on Applied Probability and Time Series Analysis*, volume 114 of *Lecture Notes in Statistics*, pages 253–266. Springer New York, 1996.

[8] F. Ball. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. *Mathematical Biosciences*, 156(1):41–67, 1999.

[9] F. Ball and P. Donnelly. Branching process approximation of epidemic models. *Theory of Probability & Its Applications*, 37(1):119–121, 1993.

[10] F. Ball and P. Donnelly. Strong approximations for epidemic models. *Stochastic Processes and their Applications*, 55(1):1–21, 1995.

[11] F. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with Two Levels of Mixing. *The Annals of Applied Probability*, 7(1):46–89, 1997.

[12] F. G. Ball and O. D. Lyne. Optimal vaccination policies for stochastic epidemics among a population of households. *Mathematical Biosciences*, 177:333–354, 2002.

[13] A. D. Barbour and G. Reinert. Approximating the epidemic curve. *Electronic Journal of Probability*, 18(54):1–30, 2013.

[14] N. G. Becker and D. Wang. Can antiviral drugs contain pandemic influenza transmission? *PloS ONE*, 6(3):e17764, 2011.

[15] A. J. Black, T. House, M. J. Keeling, and J. V. Ross. Epidemiological consequences of household-based antiviral prophylaxis for pandemic influenza. *Journal of The Royal Society Interface*, 10(81), 2013.

[16] A. J. Black, T. House, M. J. Keeling, and J. V. Ross. The effect of clumped population structure on the variability of spreading dynamics. *Journal of Theoretical Biology*, 359(0):45–53, 2014.

[17] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 1. Springer, 2001.

[18] L. R. Carrasco, V. J. Lee, M. I. Chen, D. B. Matchar, J. P. Thompson, and A. R. Cook. Strategies for antiviral stockpiling for future influenza pandemics: a global epidemic-economic perspective. *Journal of The Royal Society Interface*, 8(62):1307–1313, 2011.

[19] F. Carrat, C. Sahler, S. Rogez, M. Leruez-Ville, F. Freymuth, C. Le Gales, M. Bungener, B. Housset, M. Nicolas, and C. Rouzioux. Influenza burden of illness: Estimates from a national prospective survey of household contacts in France. *Archives of Internal Medicine*, 162(16):1842–1848, 2002.

[20] V. Colizza, A. Barrat, and M. Barthelemy. Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Medicine*, 4(1):e13, 2007.

[21] Commonwealth of Australia. Interim National Pandemic Influenza Clinical Guidelines, 2006. `http://www.health.gov.au/internet/panflu/publishing.nsf/Content/interim-pandemic-clinical-guidelines-1`.

[22] Commonwealth of Australia. Australian Health Management Plan for Pandemic Influenza, 2009. `http://www.health.gov.au/internet/panflu/publishing.nsf/Content/ahmppi-2009`.

[23] Commonwealth of Australia. Review of Australia's Health Sector Response to Pandemic (H1N1) 2009: Lessons Identified, 2011. `http://www.flupandemic.gov.au/internet/panflu/publishing.nsf/Content/review-2011-l`.

[24] B. S. Cooper, R. J. Pitman, W. J. Edmunds, and N. J. Gay. Delaying the international spread of pandemic influenza. *PLoS Medicine*, 3(6):e212, 2006.

[25] P. C. Cross, P. L. Johnson, J. O. Lloyd-Smith, and W. M. Getz. Utility of R0 as a predictor of disease invasion in structured populations. *Journal of the Royal Society Interface*, 4(13):315–324, 2007.

[26] J. M. Epstein, D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev. Controlling pandemic flu: the value of international air travel restrictions. *PloS ONE*, 2(5):e401, 2007.

[27] N. M. Ferguson, D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke. Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452, 2006.

[28] T. C. Germann, K. Kadau, I. M. Longini, and C. A. Macken. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences*, 103(15):5935–5940, 2006.

[29] A. Ghani, M. Baguelin, J. Griffin, S. Flasche, A. J. van Hoek, S. Cauchemez, C. Donnelly, C. Robertson, M. White, J. Truscott, et al. The early transmission dynamics of H1N1pdm influenza in the United Kingdom. *PLoS Currents*, 1:RRN1130, 2009.

[30] S. Ghosh and J. Heffernan. Influenza pandemic waves under various mitigation strategies with 2009 h1n1 as a case study. *PLoS ONE*, 5(12):e14307, 12 2010.

[31] G. Ghoshal, L. Sander, and I. Sokolov. SIS epidemics with household structure: The self-consistent field method. *Mathematical Biosciences*, 190(1):71–85, 2004.

[32] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[33] K. Glass, J. McCaw, and J. McVernon. Incorporating population dynamics into household models of infectious disease transmission. *Epidemics*, 3(3):152–158, 2011.

[34] E. Goldstein, J. C. Miller, J. O'Hagan, and M. Lipsitch. Predispensing of Antivirals to High-Risk Individuals in an Influenza Pandemic. *Influenza and Other Respiratory Viruses*, 4(2):101–112, 2010.

[35] E. Goldstein, K. Paur, C. Fraser, E. Kenah, J. Wallinga, and M. Lipsitch. Reproductive numbers, epidemic spread and control in a community of households. *Mathematical Biosciences*, 221(1):11 – 25, 2009.

[36] F. G. Hayden, R. Belshe, C. Villanueva, R. Lanno, C. Hughes, I. Small, R. Dutkowski, P. Ward, and J. Carr. Management of Influenza in Households: A Prospective, Randomized Comparison of Oseltamivir Treatment With or Without Postexposure Prophylaxis. *Journal of Infectious Diseases*, 189(3):440–449, 2004.

[37] F. Hoppensteadt and P. Waltman. A problem in the theory of epidemics. *Mathematical Biosciences*, 9:71–91, 1970.

[38] T. House, M. Baguelin, A. J. Van Hoek, P. J. White, Z. Sadique, K. Eames, J. M. Read, N. Hens, A. Melegaro, W. J. Edmunds, and M. J. Keeling. Modelling the impact of local reactive school closures on critical care provision during an influenza pandemic. *Proceedings of the Royal Society of London B*, 2011.

[39] T. House and M. J. Keeling. Deterministic epidemic models with explicit household structure. *Mathematical Biosciences*, 213(1):29–39, 2008.

[40] T. House and M. J. Keeling. The impact of contact tracing in clustered populations. *PLoS Computational Biology*, 6(3):e1000721, 2010.

[41] T. House, J. V. Ross, and D. Sirl. How big is an outbreak likely to be? Methods for epidemic final-size calculation. *Proceedings of the Royal Society A*, 469(2150), 2013.

[42] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15124–15129, 2004.

[43] T. Jefferson, M. A. Jones, P. Doshi, C. B. Del Mar, C. J. Heneghan, R. Hama, and M. J. Thompson. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev*, 1(1), 2012.

[44] D. W. Jordan and P. Smith. *Nonlinear ordinary differential equations*. Clarendon Press Oxford, 1987.

[45] M. J. Keeling and P. Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.

[46] M. J. Keeling and A. Shattock. Optimal but unequitable prophylactic distribution of vaccine. *Epidemics*, 4(2):78 – 85, 2012.

[47] J. K. Kelso, N. Halder, and G. J. Milne. The impact of case diagnosis coverage and diagnosis delays on the effectiveness of antiviral strategies in mitigating pandemic influenza A/H1N1 2009. *PLoS ONE*, 5(11):e13797, 2010.

[48] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A*, 115(772):700–721, 1927.

[49] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability*, 7(1):49–58, 1970.

[50] K. O. Kwok, G. M. Leung, P. Mak, and S. Riley. Antiviral stockpiles for influenza pandemics from the household perspective: Treatment alone versus treatment with prophylaxis. *Epidemics*, 5(2):92–97, 2013.

[51] J. Lessler, N. G. Reich, and D. A. Cummings. Outbreak of 2009 pandemic influenza A (H1N1) at a New York City school. *New England Journal of Medicine*, 361(27):2628–2636, 2009.

[52] I. M. Longini, M. E. Halloran, A. Nizam, and Y. Yang. Containing pandemic influenza with antiviral agents. *American Journal of Epidemiology*, 159(7):623–633, 2004.

[53] I. M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. Cummings, and M. E. Halloran. Containing pandemic influenza at the source. *Science*, 309(5737):1083–1087, 2005.

[54] L. Matrajt, M. Halloran, and I. M. Longini Jr. Optimal Vaccine Allocation for the Early Mitigation of Pandemic Influenza. *PLoS Computational Biology*, 9(3):e1002964, 2013.

[55] J. M. McCaw and J. McVernon. Prophylaxis or treatment? optimal use of an antiviral stockpile during an influenza pandemic. *Mathematical Biosciences*, 209(2):336 – 360, 2007.

[56] J. A. Metz and O. Diekmann. The dynamics of physiologically structured populations. *Lecture notes in biomathematics*, 68, 1986.

[57] C. E. Mills, J. M. Robins, and M. Lipsitch. Transmissibility of 1918 pandemic influenza. *Nature*, 432(7019):904–906, 2004.

[58] P. Minayev and N. Ferguson. Improving the realism of deterministic multi-strain models: implications for modelling influenza A. *Journal of The Royal Society Interface*, 6(35):509–518, 2009.

[59] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, 45(1):3–49, 2003.

[60] P. Neal. Compound poisson limits for household epidemics. *J. Appl. Probab.*, 42(2):334–345, 06 2005.

[61] J. R. Norris. *Markov chains*. Cambridge University Press, 1998.

[62] R. Patel, I. M. Longini Jr, and M. E. Halloran. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology*, 234(2):201–212, 2005.

[63] L. Pellis, F. Ball, and P. Trapman. Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R0. *Mathematical Biosciences*, 235(1):85 – 97, 2012.

[64] P. Pollett. Diffusion approximations for ecological models. *Proceedings of the International Congress on Modelling and Simulation*, 2:843–848, 2001.

[65] P. K. Pollett and V. T. Stefanov. Path Integrals for Continuous-Time Markov Chains. *Journal of Applied Probability*, 39(4):901–904, 2002.

[66] J. Ross, T. House, and M. Keeling. Calculation of disease dynamics in a population of households. *PLoS ONE*, 5(3):e9666, 2010.

[67] J. V. Ross and A. J. Black. Contact tracing and antiviral prophylaxis in the early stages of a pandemic: the probability of a major outbreak. *Mathematical Medicine and Biology*, 2014.

[68] E. Seneta. *Non-negative matrices and Markov chains*. Springer, 2006.

[69] K. Sharkey. Deterministic epidemic models on contact networks: correlations and unbiological terms. *Theoretical Population Biology*, 79(4):115–29, June 2011.

[70] R. Sidje. Expokit. A software package for computing matrix exponentials. *ACM Transactions on Mathematical Software (TOMS)*, 24(1):130–156, 1998.

[71] S. E. F. Spencer and P. D. O'Neill. Assessing the impact of intervention delays on stochastic epidemics. *Methodology and Computing in Applied Probability*, 15(4):803–820, 2013.

[72] G. Stiver. The treatment of influenza with antiviral drugs. *Canadian Medical Association Journal*, 168(1):49–57, 2003.

[73] A. R. Tuite, A. L. Greer, M. Whelan, A.-L. Winter, B. Lee, P. Yan, J. Wu, S. Moghadas, D. Buckeridge, B. Pourbohloul, and D. N. Fisman. Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. *Canadian Medical Association Journal*, 182(2):131–136, 2010.

[74] J. Wallinga and M. Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B*, 274(1609):599–604, 2007.

[75] World Health Organisation. Global Alert and Response (GAR), Accessed June 2014. `http://www.who.int/csr/disease/swineflu/notes/h1n1_vaccine_20090806/en/`.

[76] World Health Organisation. Global Alert and Response (GAR) - 2009, Accessed June 2014. `http://www.who.int/csr/don/archive/year/2009/en/`.

[77] J. T. Wu, S. Riley, C. Fraser, and G. M. Leung. Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Medicine*, 3(9):e361, 2006.