The Effects of Clustered Unvaccinated Individuals on Epidemic Outbreak

Hamilton Hoxie Ackerman[1], Isaac J. Michaud[2], Shannon Reed[3], Andre Robinson[4]

Graduate Mentors: Naala Brewer[5], Griselle Torres-Garcia[5]

Faculty Advisor: David E. Hiebeler[2]

[1] Department of Mathematics and Statistics, Boston University, Boston, MA 02215-2496

[2] Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752

[3] Department of Mathematics, University of Missouri - Columbia, MO 65211-0001

[4] Department of Mathematics, Medgar Evers College, Brooklyn, NY 11225-2010

[5] Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287-1804

February 6, 2009

Abstract

An SIR epidemiological household model is constructed and studied to understand the effects that clusters of unvaccinated individuals have on the disease dynamics of a population. The model contains two levels of mixing where individuals make more intra-household than inter-household contacts. Stochastic simulations and numerical solutions are utilized to explore the model. A new extension of the basic reproductive number that incorporates more spatial information by predicting the average number of tertiary infections caused by a single infected individual is introduced to describe the threshold behavior and severity of an epidemic. Using these methods we show that clustering of unvaccinated individuals always leads to more severe epidemics.

1 Introduction

Modern vaccination programs have succeeded in reducing or eliminating many infectious diseases around the world and are generally regarded as one of the greatest public health achievements of the twentieth century by the US Centers for Disease Control and Prevention (CDC) [1, 2, 3]. Multiple analyses have confirmed that they are a cost-effective asset to modern public health policy, and that high vaccination rates for children should be a top priority for health officials [4, 5].

Despite the success of vaccination initiatives in reducing infectious disease levels worldwide, the topic of vaccination is and always has been controversial for a number of reasons. Some individuals question the government's prerogative in mandating vaccinations or suspect the government of financially profiting from mandatory vaccination programs [6, 7]. Many question the efficacy and safety of vaccines, often expecting vaccines to provide near perfect immunity with little or no risk [8]. According to the CDC, commonly held public misconceptions surrounding vaccines include: that diseases had already begun to disappear without the influence of vaccines, that the majority of people who catch a disease have been vaccinated, that there are "hot lots" of vaccines that have been associated with more adverse events and deaths than others, that vaccines cause many harmful side effects, illnesses, and even death, that vaccine-preventible diseases have been virtually eliminated from the US, and that administering multiple vaccinations for different diseases simultaneously can overload the immune system and increase the risk of harmful side effects [9]. As these vaccine-preventible diseases disappear from the public eye, individuals simultaneously forget about the diseases being vaccinated against and believe that the risks associated with vaccination exceed those posed by the targeted illness [7, 8, 10]. Finally, controversial vaccine-related incidents such as the implied link between an MMR (measles, mumps, rubella) vaccine and autism in children and the use of thimerosal, a preservative containing mercury, in vaccine preparations have made it difficult to convince some members of the public that vaccines are one of our best lines of defense against many serious illnesses [11, 12].

The controversy of vaccines has also been fueled in recent years by vocal "anti-vaccination" groups. While there have always been critics of the vaccination process, the recent availability of mass media and the Internet as

outlets for the dissemination of information has enabled members of these groups to spread their message easily and inexpensively. There have been a number of papers published on the mindsets, motivations, and vehicles of these groups, and many agree that their arguments, though often based more on emotional appeal and temporal correlations than on medical fact, are capable of reaching and influencing a non-negligible proportion of the population [10, 13, 14, 15, 16].

The net effect of these controversies and this anti-vaccination influence is a growing concern in the medical community over rising rates of individuals opting out of vaccination. This trend has been most apparent in the rising numbers of parents acquiring vaccination exemptions for their children based on their religious, philosophical, or personal beliefs. A number of papers have been written studying the effects of these exemptions [17, 18, 19, 20], and a correlation between ease of exemption and pertussis incidence rate by state has been demonstrated [21]. (For an overview of the vaccination laws these parents are avoiding, see [22]). It has also been noted that homeschooled children are not required to comply with state vaccination laws, providing another opportunity for individuals to avoid vaccinating their children in compliance with school entry laws [23].

If these unvaccinated individuals were present in sufficiently small numbers and spread randomly throughout the population, the effect of herd immunity could still prevent an epidemic. However, there is reason to believe that these unvaccinated individuals could be geographically and socially clustered, which could in turn increase the risk of an outbreak or epidemic. This clustering could stem from a variety of sources. Many religious groups, including the Amish, Christian Scientists, and Jehovah's Witnesses, have resisted vaccination programs for religious and philosophical reasons [17, 24]. Socioeconomic standing is also commonly associated with disparities in local vaccination rates [25, 26, 27]. The previously mentioned ease of exemption issue has led some states in the US to possess lower vaccination rates than others. Finally, ideas from economics and game theory have been applied to this situation, with authors suggesting that individuals unsure whether or not to vaccinate their children may be influenced by these growing trends in non-vaccination around them [17, 24, 25, 28]. While the clustering of unvaccinated individuals has received a fair amount of attention in the medical community, to our knowledge, there have been no models built to quantify the increased risks associated with clusters of unvaccinated individuals. It is this risk that we explore here.

While we believe our approach in analyzing clustered unvaccinated individuals is novel, we are building on many ideas previously explored in mathematical biology. Our model is an SIR (susceptible $\rightarrow$ infected $\rightarrow$ recovered/removed) model, which was proposed in 1927 [29] and has been analyzed extensively in years since. To take the effects of vaccination into account, two classes of susceptible are included, an idea that was presented at least as early as 1968 [30] and has been analyzed in both deterministic and stochastic settings since then [31, 32, 33]. In an attempt to improve upon the common assumption of homogeneous mixing between all individuals of the population, we consider a population partitioned into households, where individuals within households contact each other at a different (typically much higher) rate than they contact individuals in the general population. This approach was first discussed by Rushton and Mautner [34] and has been considered extensively in the years since [35, 36, 37, 38, 39, 40]. As will be described later in our paper, an infinite system of differential equations arises when the moments of the probability distributions describing infectious individuals are computed. To truncate this system, moment-closure approximations are used; see [41] for an excellent introduction to the topic of moment approximations and [42] for similar approximations to what are used here. Finally, suggestions for reducing the impact of clustered unvaccinated individuals on epidemic dynamics are included at the end of the paper based on our findings. These suggestions follow an extensive line of public health investigation on optimal vaccination strategies in a variety of model settings [40, 43, 44, 45, 46].

Section 2 presents the household equations, the measure of clustering of unvaccinated individuals used throughout the paper, the moment equations, the moment closure methods used to truncate the infinite system of moment equations, and the derivation of the first and second-generation R_0 equations for the model. Section 3 contains an explanation of the methods utilized in studying the behavior of our model. Section 4 contains the results, which were found via simulations, numerical solutions, and analysis. In Section 5, we discuss the meaning and implications of the results. Finally, in Section 6, we extend our theoretical results to public health policy and propose model modifications that could be made in order to analyze more specific scenarios.

Monte Carlo R_0 simulator converges nicely to the deterministic value of R_0. Thus, we can say that the numerical values derived from our analysis line up with our stochastic expectations.

It should be noted that the differences between the numerical solution and the simulation in Figure 3 of our model do not impact the value of either the first or second-generation R_0. It may seem like analysis drawn from the household equations should not align with the stochastic simulation due to the discrepancies in Figure 3, but luckily here the discrepancies do not affect our analysis. This is due to the fact that the simulation and numerical solutions deviate from each other long after the time of the initial infection. As can be seen in Figure 3, for $t < 5$ the two curves are nearly indistinguishable from one another. Since numerical solutions and stochastic simulations both exhibit the same initial behavior, which is important to R_0, we can safely use our analytically derived R_0 values when discussing the simulations.

After verifying that our deterministic expressions for the first and second-generation R_0 values align with their expected value over many stochastic realizations, it is important to verify that the original epidemiological meaning behind R_0 as a threshold parameter still holds. By computing values of the first and second-generation R_0 values which live around the threshold value of 1, we were able to run simulations and observe the resulting epidemic behavior. To produce Figures 6-8, we ran many full stochastic simulations with borderline threshold parameter values. The general trend shows that the second-generation R_0 values are better predictors of whether or not an epidemic will occur than the first-generation values.

It is important to realize that even though the second-generation R_0 is deterministically greater than or less than 1, this fact will not perfectly predict epidemic behavior. This deterministic second-generation R_0 value is an average taken over all individuals and households, whereas the outcome of a single infection will depend greatly on its starting location. Infectious individuals placed in completely unvaccinated households are more infectious than individuals placed in purely vaccinated households due to the spatial structure of the system. From the point of view of the entire population, the global R_0 is a random variable which depends on the initially infected individual.

Figures 6-8 gave reason to believe that the second-generation R_0 was a better predictor of epidemiological outbreaks when considering the effect of clustering than the traditional R_0 was. Now that our derived expression for the second-generation R_0 is an adequate measures of the expected severity of an epidemic outbreak, we can use it to finally demonstrate the general effect that clustering has on an epidemic.

Since the second-generation R_0 describes the behavior of an epidemic outbreak in our model, we can continue analyzing it to determine the impact of increased clustering on epidemics. By first taking the derivative of the second-generation R_0 with respect to the other parameters, we derived exactly what intuition would suggest. The most important final result is that the derivative of the second-generation R_0 with respect to Q is always positive (Equation (22)). This implies that increasing clustering always increases the value of the second-generation R_0, which by the rest of our analysis shows that clustering will always, on average, create larger and more probable epidemics when an outbreak occurs.

The result that we have found for households is analogous to the result that was found in lattice population models: that clustering of suitable habitat was found to beneficial to a locally dispersing population [53]. Here we see the same effect where increased clustering causes the final proportion of recovered individuals in the population to increase. In addition to this, we found the second-generation R_0 to be an adequate threshold estimator which allows us to predict not only the severity of an epidemic but also the likelihood of an outbreak turning into full blown epidemic.

6 Conclusion

Some epidemiological models are used to describe the spread of disease in the presence of vaccinated individuals. We believe that it would be advantageous to incorporate clustering into these models in order to gain a more thorough understand of the effect of the unvaccinated population on disease dynamics, especially considering the current trends of clustering mentioned previously. Incorporating clustering into a model is a fairly general procedure and could be applied to the many vaccines that are currently under some measure of public scrutiny.

Our investigation prompted many questions about current and future public health policy. How could the medical community target these unvaccinated clusters in order to reduce the chances of an outbreak? When and how should school entry vaccination law exemptions be given to avoid clustering of the unvaccinated? Perhaps a more involved and taxing exemption process could efficiently reduce a significant portion of the school based clustering in many communities. In addition, more involved health education, both at the student and parent levels, could change some of the negative attitudes surrounding vaccines. It would be beneficial to derive strategies concerning vaccination education that might be implemented to balance the onslaught of pseudo-medical information found on the Internet and in popular periodic publications. More future research into the matter will be necessary to understand just how the academic and medical communities can impact

this troubling trend towards increasing numbers of clusters of unvaccinated individuals.

There are many modifications that could be implemented to create a more realistic epidemiological model. For instance, many vaccine-preventible diseases, including pertussis, measles, mumps, tuberculosis, and smallpox, have a latent period after initial exposure. An SEIR model could be built with a exposed/latent class to include this time delay. Outbreaks of these diseases can last for months, so it could also be worthwhile to incorporate population birth and death rates. Perhaps even more worthwhile would be to consider an active vaccination program during an epidemic, with unvaccinated individuals becoming vaccinated at a rate related to the population infection rate.

Our spatial structure contained two levels of mixing, but this could be augmented to three or more. This extension would permit the modeling of situations such as the mixing of individuals living in dorms within colleges within cities or churches within suburbs within larger urban areas. Variable mixing could also be incorporated, with mixing between groups i and k occurring at rate α_{ik}. This could be used to model geographic spatial interactions, where New York and New Jersey would mix more frequently than New York and Arizona, or social boundaries where households of similar socioeconomic status would be more likely to interact than with others of significantly different socioeconomic status. These or perhaps even more complex contact structures could be included in the model [58]. Combining these ideas with a model that does not assume that all households are of the same size would allow the model to be applied to specific diseases or geographic locations. Finally, reparameterization could be done to force the moment equations to be numerically stable.

Despite the fact that this model leaves room for these extensions, many important facts concerning the behavior of infections exacerbated by clustering have been illuminated. Herd immunity is more easily diminished in communities with relatively high vaccination rates than many well-mixed models have the ability to demonstrate and than most of the general public has cause to believe. Although the effects of clustering on epidemiological systems have been hypothesized by many [21, 24, 42, 59, 60], they have not previously been quantified, nor have they been applied to current health standards in order to maximize public safety. It now has been shown that there is not only an undeniable upward trend in the overall size of an epidemic and the level of clustering, but also that clustering allows for epidemics to exist under conditions that would otherwise be unfavorable for an outbreak. (Instances of this occurring in various communities around the world have been documented in recent years [24, 61, 62, 63, 64, 65]). All of these results contribute to the mounting evidence that clustering and the spatial considerations that it entails are crucial to the study of epidemiological dynamics.

7 Acknowledgments

We would like to thank Carlos Castillo-Chavez for giving us the opportunity to participate in MTBI 2008 summer program. We thank David Hiebeler, Mat Gluck, Naala Brewer, Griselle Torres-Garcia and other MTBI faculty for their help, guidance, and suggestions in developing this report. Without their support, none of this would have been possible. This material is based upon work supported by the National Science Foundation under Grant No. DMS-0718786 to D.H. The MTBI/SUMS Summer Undergraduate Research Program is supported by The National Science Foundation (DMS-0502349), The National Security Agency (DOD-H982300710096), The Sloan Foundation, and Arizona State University.

2 Model

Our model is an SIR household model describing the spread of a theoretical epidemic. Consider a population consisting of n_2 households, each containing n_1 individuals, such that total population size is $n_1 n_2 = N$. As in most household models, individuals within the same household interact with one another at a higher rate than they do with individuals in other households. This model is an extension of the SIS model proposed by Hiebeler [42].

To model the impact of clustered unvaccinated individuals on the spread of an epidemic, we introduce two separate classes of susceptible individuals: vaccinated and unvaccinated. Unvaccinated individuals are always infected when contacted by an infected individual, and vaccinated individuals are infected at a lower rate which is proportional to the efficacy of the vaccine. To simplify the analysis of the system, a number of assumptions are made. First, it is assumed that the vaccine in question does not grant perfect immunity, so individuals who have been vaccinated can still become infected but will do so at a reduced rate. In reality, no vaccine grants perfect immunity one hundred percent of the time it is administered; reported efficacies can range anywhere from 10% to 97% effective and depend on a number of factors, including the disease being vaccinated against, the type of vaccine administered, the age of the target population, the number of doses administered, and whether or not the vaccines are administered during an outbreak [47, 48, 49, 50, 51].

It is also assumed that there is no incubation period, so infected individuals can immediately infect susceptible individuals. Vaccinated and unvaccinated individuals are equally infective once infected. Once an individual is infected, he infects people independently and at a constant rate over the course of his infection period. There are also no deaths due to infection. Finally, it is assumed that there is no vaccination during the outbreak; the initial distribution of vaccinated individuals determines the vaccination status of the system.

2.1 The Household Equations

Let $U_k(t)$, $V_k(t)$, $I_k(t)$, and $R_k(t)$ represent the proportion of individuals of household k at time t who are unvaccinated, vaccinated, infected, and recovered (immune), respectively. Then the dynamics are described by the following set of equations:

$$\frac{dU_k}{dt} = -[\phi(1-\alpha)U_k I_k + \phi\alpha E[I]U_k] \tag{1}$$

$$\frac{dV_k}{dt} = -(1-\rho)[\phi(1-\alpha)V_k I_k + \phi\alpha E[I]V_k] \tag{2}$$

$$\frac{dI_k}{dt} = \phi(1-\alpha)I_k(U_k + (1-\rho)V_k) + \phi\alpha E[I](U_k + (1-\rho)V_k) - \mu I_k \tag{3}$$

$$\frac{dR_k}{dt} = \mu I_k \tag{4}$$

where $k = 1, 2, ..., n_2$, ϕ is the per capita infection rate of the disease, μ is the per capita recovery rate, and ρ is the efficacy of the vaccine. If $\rho = 1$, the vaccine provides perfect immunity, and $\rho = 0$ implies that vaccinated and unvaccinated individuals are equally susceptible. The parameter α is the proportion of global infection events. If $\alpha = 1$, then all events are global and the household system becomes a traditional well-mixed SIR model, and when $\alpha = 0$ there is no interaction between households. Let $E[I]$ denote the average proportion of infected individuals at time t, where the average is taken over the n_2 households. Then

$$E[I] = \frac{1}{n_2}\sum_{k=1}^{n_2} I_k(t)$$

$E[U]$ and $E[V]$ are similarly defined as the average proportion of unvaccinated and vaccinated individuals in the population at time t, respectively. By assuming that the timescale of a single epidemic outbreak is relatively short, we assume that the total population N remains fixed, and natural births and deaths are neglected. Since the population is constant, then $U_k + V_k + I_k + R_k = 1$ for each household, so we can do without Equation (4) when analyzing the system. Thus, the k-household dynamics are governed by three equations, that is, we have a system of $3n_2$ equations describing the disease dynamics for the entire population. Also note that, for any time t, we have $E[U] + E[V] + E[I] + E[R] = 1$. Finally, we assume that n_1 is sufficiently large to use continuous state variables for $U_k(t)$, $V_k(t)$, $I_k(t)$, and $R_k(t)$. In the case when n_1 is small, a branching process is typically used [38, 52].

2.2 Initial Conditions

Let $U(t)$, $V(t)$, and $I(t)$ be sets of the proportions $U_k(t)$, $V_k(t)$, and $I_k(t)$, respectively, for $k = 1, ..., n_2$. The initial household distributions are thus given by $U(0)$, $V(0)$, and $I(0)$. We can average over all households to compute the mean proportions of unvaccinated, vaccinated, and infected individuals, denoted $E[U(0)]$, $E[V(0)]$, and $E[I(0)]$, respectively. These average values are important initial conditions for our household system, and we use special notation to denote them:

$$\mathcal{U}_0 = E[U(0)] = \text{proportion of the total population initially unvaccinated}$$
$$\mathcal{V}_0 = E[V(0)] = \text{proportion of the total population initially vaccinated}$$

Note that when our household system is initially constructed, there are no infected individuals present. With these initial conditions defined, we can now derive the most important initial condition of our model: Q, the measure of clustering of unvaccinated individuals.

2.3 Initial Clustering

To study the impact of clustering of unvaccinated individuals on the dynamics of an epidemic, we need to consider the initial distribution of unvaccinated individuals within the system of households. When introducing an infected individual into a susceptible population, it should be intuitive that an outbreak is more likely to occur in a household with a high initial proportion of unvaccinated individuals than in a household with a high proportion of vaccinated individuals. Because this infected individual can be introduced into any household, we'd like to be able to quantify the overall disparities in initial conditions among households i.e. how clustered unvaccinated individuals are in the population.

A recent approach by Becker and Dietz used mean and variance to quantify differences in household composition [36]. In our model, we have already defined the mean of interest as $\mathcal{U}_0$, and we can begin by using the variance of $U(0)$ to measure the clustering of unvaccinated individuals within our system of households. Consider the bounds of the variance of $U(0)$:

$$0 \leq Var(U(0)) \leq E[U(0)] - E[U(0)]^2$$

Dividing by E[U(0)]:

$$0 \leq \frac{Var(U(0))}{E[U(0)]} \leq 1 - E[U(0)]$$

Adding E[U(0)]:

$$E[U(0)] \leq \frac{Var(U(0))}{E[U(0)]} + E[U(0)] \leq 1 \tag{5}$$

Let us now define the value Q as:

$$Q = \frac{Var(U(0))}{E[U(0)]} + E[U(0)] = \frac{E[U(0)^2] - E[U(0)]^2}{E[U(0)]} + E[U(0)] = \frac{E[U(0)^2]}{E[U(0)]} \tag{6}$$

The value Q still contains the same information as the variance of $U(0)$, but it is a more intuitive expression for measuring clustering in our model. This is because Q can be interpreted in a probabilistic sense: Q is the probability that a randomly selected neighbor (individual from the same household) of a randomly selected unvaccinated individual is also unvaccinated [42]. Thus, higher values of Q signify a greater degree of clustering of unvaccinated individuals. Based on the bounds established in Equation (5), the extreme cases are $Q = E[U]$ and $Q = 1$. In the former case, every household contains exactly $\mathcal{U}_0$ proportion of unvaccinated individuals i.e. unvaccinated individuals are distributed uniformly among the households. In the latter case, $\mathcal{U}_0$ proportion of the households are completely unvaccinated. Using Q, we can now precisely describe the initial clustering present in our model.

Before we can either numerically solve or simulate our model, we must have an initial distribution of unvaccinated and vaccinated individuals among the households. A computer script was written to create this distribution, and uses a method similar to an algorithm used previously to generate heterogeneous lattice landscapes [53]. The algorithm works as follows: first, the desired proportion of unvaccinated individuals is specified ($\mathcal{U}_0$), as is the ultimate desired level of clustering (Q). Also specified are n_1 and n_2. Next, n_2 households are

created, each containing proportion $\mathcal{U}_0$ of unvaccinated individuals and proportion $(1 - \mathcal{U}_0) = \mathcal{V}_0$ of vaccinated individuals. Since there is no variance among households in terms of proportion unvaccinated, the current Q value, which will be denoted Q^*, is equal to $\mathcal{U}_0$.

In order to move this current Q^* closer to the desired Q, two households are selected at random and their proportions of unvaccinated individuals are compared. If Q^* is less than Q, we move an unvaccinated individual from the household with the lower proportion of unvaccinated individuals to the household with the higher proportion of unvaccinated individuals. To keep household sizes the same, we then move a vaccinated individual from the household with the higher proportion of unvaccinated individuals to the household with the lower proportion of unvaccinated individuals; in effect, we have swapped an unvaccinated individual and a vaccinated individual between these two households in a way that puts more unvaccinated individuals into the same household. On the other hand, if Q^* is greater than Q, we simply swap two individuals in the opposite direction - move an unvaccinated individual into the household with the lower proportion of unvaccinated individuals and replace him with a vaccinated individual. (Since the initial Q^* is equal to its lower bound, the swaps will almost always tend to increase Q^*.) After every swap, Q^* is recomputed and compared to Q, and this entire swapping process is repeated until the current Q^* is within some specified tolerance, ϵ, of Q. At this point, we have created a system of households with the desired level of clustering and $Q^* = Q$.

This algorithm is a simplified version from Hiebeler, where our $\mathcal{U}_0$ and Q are equivalent to Hiebeler's p_0 and q_{ss} in both use and interpretation [53]. However, our system is less spatially structured than a lattice landscape. Hiebeler's original algorithm uses the notion of adjacency of sites to define clustering, but here we have defined adjacency of individuals to mean that they both belong to the same household. This change of definition is used because there are no spatial relationships within a single household, so the most that can be said about the spatial correlation between individuals is whether or not they belong to the same household.

To illustrate this notion of Q, consider the examples of systems of households in Figure 1. Each system was generated with constant $\mathcal{U}_0 = 0.25$ but varying Q values of 0.255, 0.3, 0.5, and 1.0, respectively. When Q is near $\mathcal{U}_0$, the unvaccinated individuals of the population are distributed approximately homogeneously among households. As the value of Q increases, the majority of unvaccinated individuals become markedly clustered into a smaller number of households.

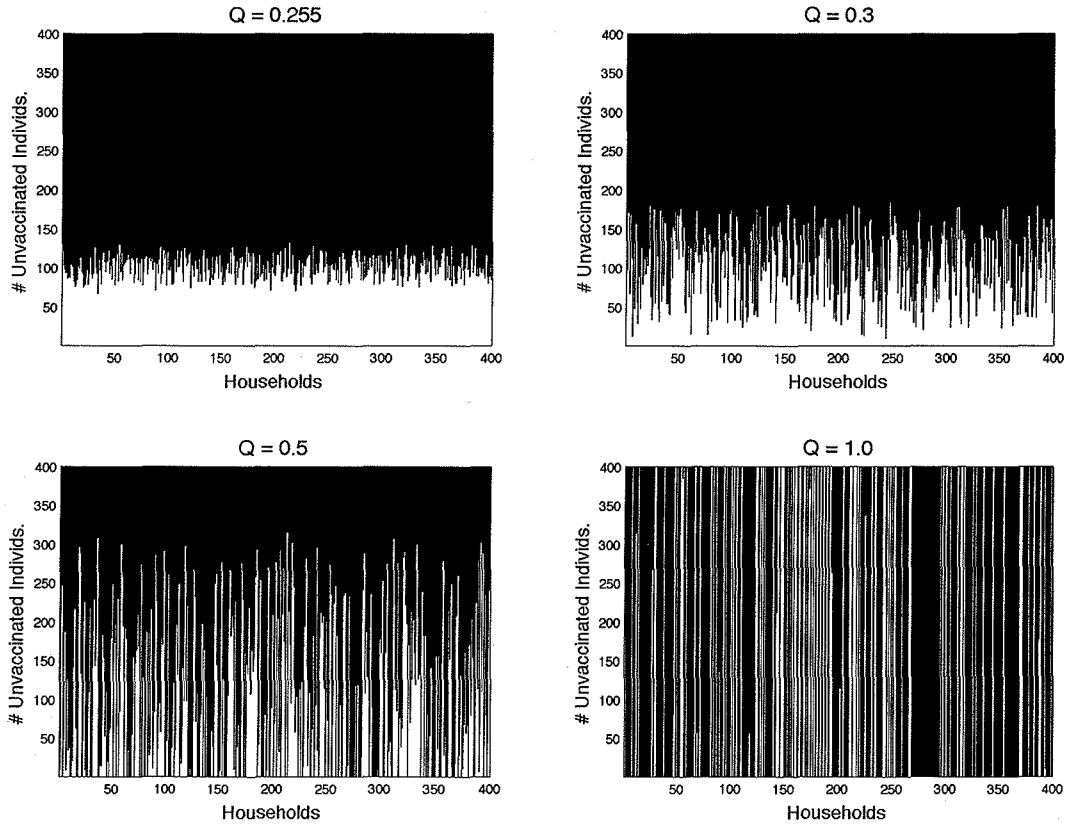


Figure 1: A graphical illustration of Q, the overall measure of clustering of unvaccinated individuals in a population. In each system of 400 households with 400 individuals each, 25% of individuals are unvaccinated ($\mathcal{U}_0 = 0.25$). However, as Q is increased, the unvaccinated individuals become increasingly clustered into the same households. For Q = 0.255, unvacinated individuals are distributed essentially homogeneously throughout the households. For Q=0.3 and 0.5, increasingly large disparities in non-vaccination levels arise between households. For Q=1, theoretically any household containing one unvaccinated individual must be full of unvaccinated individuals. Those households pictured with both vaccinated and unvaccinated individuals present in the case of Q=1 exist because the algorithm stops within ϵ of Q^*=1.

2.4 Moment Equations

When considering the entire system of household equations, one must keep track of $3n_2$ equations. An alternative approach is to consider the entire system by writing differential equations which describe the change in the expected value of U, V, and I, where the expectation is taken over all households [41, 42, 54]. The equations can be derived via the following process:

$$\begin{aligned}\frac{d}{dt}E[U] = \frac{d}{dt}(\frac{1}{n_2}\sum_{k=1}^{n_2}U_k) = (\frac{1}{n_2}\sum_{k=1}^{n_2}\frac{d}{dt}U_k) &= E[\frac{d}{dt}U_k] \\ &= E[-[\phi(1-\alpha)U_kI_k + \phi\alpha E[I]U_k]] \\ &= -(\phi(1-\alpha)E[UI] + \phi\alpha E[I]E[U])\end{aligned}$$

By the same process, equations for the derivative of the expected value of V and I can be computed. The entire system can thus be reduced to:

$$\frac{d}{dt}E[U] = -(\phi(1-\alpha)E[UI] + \phi\alpha E[I]E[U]) \tag{7}$$

$$\frac{d}{dt}E[V] = -(1-\rho)(\phi(1-\alpha)E[VI] + \phi\alpha E[I]E[V]) \tag{8}$$

$$\frac{d}{dt}E[I] = \phi(1-\alpha)(E[UI] + (1-\rho)E[VI]) + E[I](\phi\alpha(E[U] + (1-\rho)E[V]) - \mu) \tag{9}$$

These moment equations reduce the original system of household equations by a factor of n_2, a significant improvement. However, this simplification is not without a cost because these three moment equations do not form a closed system, indicated by the presence of the joint moment terms $E[UI]$ and $E[VI]$. Equations for these two joint moments are as follows (see Appendix A for derivations):

$$\begin{aligned}\frac{d}{dt}E[UI] =& -\phi(1-\alpha)E[I^2U] - \phi\alpha E[I]E[UI] + \phi(1-\alpha)(E[U^2I] + (1-\rho)E[UIV]) \\ &+ \phi\alpha E[I](E[U^2] + (1-\rho)E[UV]) - \mu E[UI]\end{aligned} \tag{10}$$

$$\begin{aligned}\frac{d}{dt}E[VI] =& -(1-\rho)\phi(1-\alpha)E[I^2V] - \phi\alpha E[I]E[VI] + \phi(1-\alpha)(E[UIV]) \\ &+ (1-\rho)E[V^2I]) + \phi\alpha E[I](E[UV] + (1-\rho)E[V^2]) - \mu E[VI]\end{aligned} \tag{11}$$

Within these equations, we see three new second-order moments: $\{E[UV], E[U^2], E[V^2]\}$. We also see a number of third-order moments arise, which are the terms that will eventually be approximated using first and second order moments. First, equations for these three new joint moments are given (see Appendix A for derivations):

$$\begin{aligned}\frac{d}{dt}E[UV] =& -(1-\rho)\phi(1-\alpha)E[UIV] - \phi\alpha E[I]E[UV] - \phi(1-\alpha)E[UIV] \\ &- \phi\alpha E[I]E[UV]\end{aligned} \tag{12}$$

$$\frac{d}{dt}E[U^2] = -\phi(1-\alpha)E[U^2I] - \phi\alpha E[I]E[U^2] \tag{13}$$

$$\frac{d}{dt}E[V^2] = -2(1-\rho)\phi\left((1-\alpha)E[V^2I] + \alpha E[I]E[V^2]\right) \tag{14}$$

In these five equations for the second-order moments, we see five third-order moments: E$[I^2U]$, E$[I^2V]$, E$[U^2I]$, E$[V^2I]$, and E$[UIV]$. Explicitly finding equations for these moments would produce fourth-order moments, which would in turn produce fifth-order moments, etc. Instead of considering an infinite system of moment equations, we use the following moment-closure approximations [42] (see Appendix B for derivations):

$$E[UIV] \approx \frac{E[UI]E[UV]}{E[U]}$$

$$E[U^2I] \approx \frac{E[U^2]E[UI]}{E[U]}$$

$$E[I^2V] \approx \frac{E[VI]^2}{E[V]}$$

$$E[V^2I] \approx \frac{E[VI]^2}{E[I]}$$

$$E[I^2U] \approx \frac{E[UI]^2}{E[U]}$$

Substituting these approximations back into Equations (10-14) and combining them with Equations (7-9), we have the final eight-dimensional system of moment equations for this model:

$$\frac{d}{dt}E[U] = -(\phi(1-\alpha)E[UI] + \phi\alpha E[I]E[U])$$
$$\frac{d}{dt}E[V] = -(1-\rho)(\phi(1-\alpha)E[VI] + \phi\alpha E[I]E[V])$$
$$\frac{d}{dt}E[I] = \phi(1-\alpha)(E[UI] + (1-\rho)E[VI]) + E[I](\phi\alpha(E[U] + (1-\rho)E[V]) - \mu)$$
$$\frac{d}{dt}E[IU] \approx -\phi(1-\alpha)\frac{E[IU]^2}{E[U]} - \phi\alpha E[I]E[IU] + \phi(1-\alpha)\left(\frac{E[U^2]}{E[U]}E[IU] + (1-\rho)\frac{E[IU]E[UV]}{E[U]}\right)$$
$$+ \phi\alpha E[I](E[U^2] + (1-\rho)E[UV]) - \mu E[IU]$$
$$\frac{d}{dt}E[IV] \approx -(1-\rho)\phi(1-\alpha)\frac{E[IV]^2}{E[V]} - \phi\alpha E[I]E[IV] + \phi(1-\alpha)\left(\frac{E[IU]E[UV]}{E[U]} + (1-\rho)\frac{E[VI]^2}{E[I]}\right)$$
$$+ \phi\alpha E[I](E[VU] + (1-\rho)E[V^2]) - \mu E[IV]$$
$$\frac{d}{dt}E[UV] \approx -(1-\rho)\phi(1-\alpha)\frac{E[IU]E[UV]}{E[U]} - \phi\alpha E[I]E[UV] - \phi(1-\alpha)\frac{E[IU]E[UV]}{E[U]} - \phi\alpha E[I]E[UV]$$
$$\frac{d}{dt}E[U^2] \approx -\phi(1-\alpha)\frac{E[U^2]E[IU]}{E[U]} - \phi\alpha E[1]E[U^2]$$
$$\frac{d}{dt}E[V^2] \approx -2(1-\rho)\phi\left((1-\alpha)\frac{E[VI]^2}{E[I]} + \alpha E[I]E[V^2]\right)$$

Thus, this system of differential equations describes the approximate behavior of the entire system of households. Both household and moment systems are solved numerically, and the results of these solutions are discussed in later sections of the paper.

2.5 First and Second-Generation R_0

The basic reproductive number R_0 is a classic component of epidemic modeling and is generally defined as the expected number of secondary infections an infected individual will create when placed into a population consisting entirely of susceptible individuals [55, 56, 57]. In our model, a single infectious individual placed into household k can initiate four types of infection events: long-distance unvaccinated, long-distance vaccinated, local unvaccinated, and local vaccinated. Based on the parameters of the model, we can compute the expected number of each type of event this individual will propagate. The sum of these four quantities is the total expected number of secondary infections caused by the initial infected individual when beginning in household k:

$$R_{0,k} = \underbrace{\frac{\phi}{\mu}\alpha E[U]}_{\text{long-dist. U's}} + \underbrace{\frac{\phi}{\mu}\alpha(1-\rho)E[V]}_{\text{long-dist. V's}} + \underbrace{\frac{\phi}{\mu}(1-\alpha)U_k}_{\text{local U's}} + \underbrace{\frac{\phi}{\mu}(1-\alpha)(1-\rho)V_k}_{\text{local V's}} \tag{15}$$

To find the equation for the traditional first-generation R_0, take the expected value over all households:

$$E[R_{0,k}] = \left(\frac{\phi}{\mu}\right)(\alpha(E[U] + (1-\rho)E[V]) + (1-\alpha)(E[U] + (1-\rho)E[V]))$$

Substituting in $E[U(0)] = \mathcal{U}_0$ and $E[V(0)] = \mathcal{V}_0$, we have:

$$= \left(\frac{\phi}{\mu}\right)(\mathcal{U}_0 + (1-\rho)\mathcal{V}_0)$$
$$= \left(\frac{\phi}{\mu}\right)(\mathcal{U}_0 + (1-\rho)(1-\mathcal{U}_0))$$
$$= \left(\frac{\phi}{\mu}\right)(1-\rho(1-\mathcal{U}_0)) = {R_0}^{(1)} \tag{16}$$

While this is certainly a useful equation to consider, it fails to incorporate the clustering of unvaccinated individuals present in the population. For clustering to play a role, we consider the fact that each secondary infected individual can, in turn, create tertiary infections of four types: long-distance unvaccinated, long-distance vaccinated, local unvaccinated, and local vaccinated. The expected number of tertiary infections created by each secondary infected individual are as follows:

$$\text{Long-dist. U's:} \quad \left(\frac{\phi}{\mu}\right)\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(Q_{U|U}+(1-\rho)Q_{V|U}\right) \tag{17}$$

$$\text{Long-dist. V's:} \quad \left(\frac{\phi}{\mu}\right)\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(Q_{U|V}+(1-\rho)Q_{V|V}\right) \tag{18}$$

$$\text{Local U's:} \quad \left(\frac{\phi}{\mu}\right)\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(U_k+(1-\rho)V_k\right) \tag{19}$$

$$\text{Local V's:} \quad \left(\frac{\phi}{\mu}\right)\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(U_k+(1-\rho)V_k\right) \tag{20}$$

In Equations (17-20), the notation $Q_{X|Y}$ is used to denote the conditional probability of choosing an individual of type X from a given household, given that an individual of type Y was previously selected from the same household. These probabilities relate directly to the clustering of unvaccinated (and implicitly, unvaccinated) individuals. Using Equation (6) as an example, we have the following formulas:

$$Q_{U|U}=\frac{E[U^2]}{E[U]}, Q_{V|V}=\frac{E[V^2]}{E[V]}, Q_{V|U}=\frac{E[UV]}{E[U]}, Q_{U|V}=\frac{E[UV]}{E[V]}$$

To compute the "second-generation R_0" for the k^{th} household, denoted $R_{0,k}^{(2)}$, compute the product of each corresponding component of Equation (15) and its corresponding equation in Equations (17-20). This gives:

$$\begin{aligned}R_{0,k}^{(2)}=\left(\frac{\phi}{\mu}\right)^2 &\left[\alpha E[U]\left\{\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(Q_{U|U}+(1-\rho)Q_{V|U}\right)\right\}\right.\\ &+\alpha(1-\rho)E[V]\left\{\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(Q_{U|V}+(1-\rho)Q_{V|V}\right)\right\}\\ &+(1-\alpha)U_k\left\{\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(U_k+(1-\rho)V_k\right)\right\}\\ &\left.+(1-\alpha)(1-\rho)V_k\left\{\alpha\left(E[U]+(1-\rho)E[V]\right)+(1-\alpha)\left(U_k+(1-\rho)V_k\right)\right\}\right]\end{aligned}$$

Note that there are four terms which have the common factor $\alpha\left(E[U]+(1-\rho)E[V]\right)$. Factoring this term out and rearranging the remaining terms, we have

$$\begin{aligned}R_{0,k}^{(2)}=&\alpha\left(\frac{\phi}{\mu}\right)R_{0,k}(E[U]+(1-\rho)E[V])\\ &+\left(\frac{\phi}{\mu}\right)^2(1-\alpha)\left\{\alpha E[U^2]+\alpha(1-\rho)E[UV]+\alpha(1-\rho)E[UV]+\alpha(1-\rho)^2E[V^2]\right.\\ &\left.+(1-\alpha)U_k^2+(1-\alpha)(1-\rho)U_kV_k+(1-\alpha)(1-\rho)U_kV_k+(1-\alpha)(1-\rho)^2V_k^2\right\}\end{aligned}$$

This expression is still in terms of k, but taking the average over all households gives

$$E[R_{0,k}^{(2)}]=\left(\frac{\phi}{\mu}\right)R_0^{(1)}\alpha(E[U]+(1-\rho)E[V])+\left(\frac{\phi}{\mu}\right)^2(1-\alpha)E[U^2]+2(1-\rho)E[UV]+(1-\rho)^2E[V^2]$$

which can be rewritten as

$$E[R_{0,k}^{(2)}]=\left(\frac{\phi}{\mu}\right)^2\left(\alpha(E[U]+(1-\rho)E[V])^2+(1-\alpha)(E[U^2]+2(1-\rho)E[UV]+(1-\rho)^2E[V^2])\right)$$

Since $E[U(0)^2]=\mathcal{U}_0Q$, $E[U(0)V(0)]=\mathcal{U}_0(1-Q)$, and $E[V(0)^2]=E[(1-U(0))^2]=1-\mathcal{U}_0(2-Q)$, this reduces to

$$=\left(\frac{\phi}{\mu}\right)^2\left((\rho\mathcal{U}_0+1-\rho)^2+(1-\alpha)\mathcal{U}_0\rho^2(Q-\mathcal{U}_0)\right)=R_0^{(2)} \tag{21}$$

where Equation (21) is our final equation for the second-generation R_0.

It is important to recognize that the second-generation R_0 given in Equation (21) is a slight overestimate. When computing the traditional first-generation R_0, the assumptions are made that the primary infected individual is counted as both an infected individual as well as a susceptible individual, and furthermore that the population is large enough to ensure that the primary infected individual does not attempt to infect the same individual more than once. We take this one step further for the second-generation R_0 by applying these same two assumptions to *every* secondary infected individual in the population. For reasonable first and second-generation

R_0 values and sufficiently large households, the true values and our computed values given by Equations (16) and (21) should be very similar.

Even in the case when true values and computed values differ by a non-negligible amount, we believe that from a public health policy perspective, it is probably better to overestimate these values than to underestimate them. While overestimating these values may occasionally result in preparing for an epidemic that is of a smaller magnitude than anticipated, this outcome is far better than under-preparing for an epidemic that is much more serious than predicted. Given reasonable parameter values, however, this should not be an issue.

3 Methods

3.1 Numerical Solutions

Numerically solving the $3n_2$ household differential equations may seem like a daunting challenge, but because of the symmetry and simplicity of our deterministic model and the speed of modern computers, this system takes only a fraction of a second to solve. Our numerical solutions were obtained using MATLAB's `ode45` ordinary differential equation solver. Solving this system gives us values for U_k, V_k, I_k, and R_k for each household over the course of the epidemic. Taking these values and computing the expected value of each class over time yields $E[U], E[V], E[I]$, and $E[R]$ as functions of time, which can be graphed and provide a perspective on the epidemic's course from the point of view of the population. We also attempted to numerically solve the system of eight moment equations.

3.2 Stochastic Initial Conditions

Even though numerically solving the system of household equations is a purely deterministic process, every run of the numerical solution is unique because the initial conditions of the differential equations are stochastic. First, the household algorithm described in Section 2.3 produces systems of households which are randomly generated with specified $\mathcal{U}_0$ and Q values. Next, two important new initial conditions are specified: the number of households which will initially contain infected individuals (δ_2) and the number of individuals within each of these households to be initially infected (δ_1). These δ_2 households are chosen randomly from the n_2 possible households, and then δ_1 out of n_1 individuals from each household are randomly infected, with no preference given to vaccinated or unvaccinated individuals. Without clustering to consider, this would not cause any difference between runs, but because clustering is present, different initial conditions will produce different epidemic outcomes. To account for stochastic disparities, we complete multiple runs of each numerical solution, average the runs together, and give an indication of the amount of variability within the repeated runs.

3.3 Stochastic Simulation

A stochastic simulation of our theoretical epidemic was programmed in MATLAB. The simulations are computationally slower than numerically solving the system of household equations, but they give a much more realistic description of the dynamics of an epidemic as it spreads through the spatial structure of the households for a number of reasons. In the stochastic simulations, household sizes are specified, whereas no such specification is made for the household equations. More importantly, however, events are discrete and random in the stochastic simulations, whereas they are continuous and completely deterministic in the household equations. For the same exact parameter values, initial conditions, and household distribution, the `ode45` solutions will be the same every time, whereas the stochastic simulations will vary based on the outcomes of random events. This latter case is more realistic when considering the spread of disease in a dynamic population, and so stochastic simulations are favored over numerical solutions for analyzing the model and understanding the phenomenon under consideration.

3.4 Monte Carlo Second-Generation R_0 Simulator

Using the traditional interpretation of first-generation R_0, it is possible to stochastically compute R_0 by introducing one infected individual into a population of susceptible individuals and counting the number of secondary infections that are propagated. Repeating this process many times and averaging the results should produce an accurate first-generation R_0. Much in the same way, it is also possible to stochastically compute second-generation R_0 by introducing an infected individual into an entirely susceptible population and counting the number of tertiary infections produced indirectly by the initially infected individual. Again, the average of many repetitions should result in an accurate second-generation R_0. Using this Monte Carlo method, we are able to verify the formulas given by Equations (16) and (21).

4 Results

4.1 Epidemic Progression Visualized

In Figure 2 we can see the progression of an epidemic with respect to the proportion of each type of individual in every household. As the epidemic proceeds it is clear that households which contain more unvaccinated people are much more vulnerable to the epidemic. From these snapshots it is obvious that clusters of unvaccinated individuals pose a problem because the epidemic can grow quickly through local infection events.

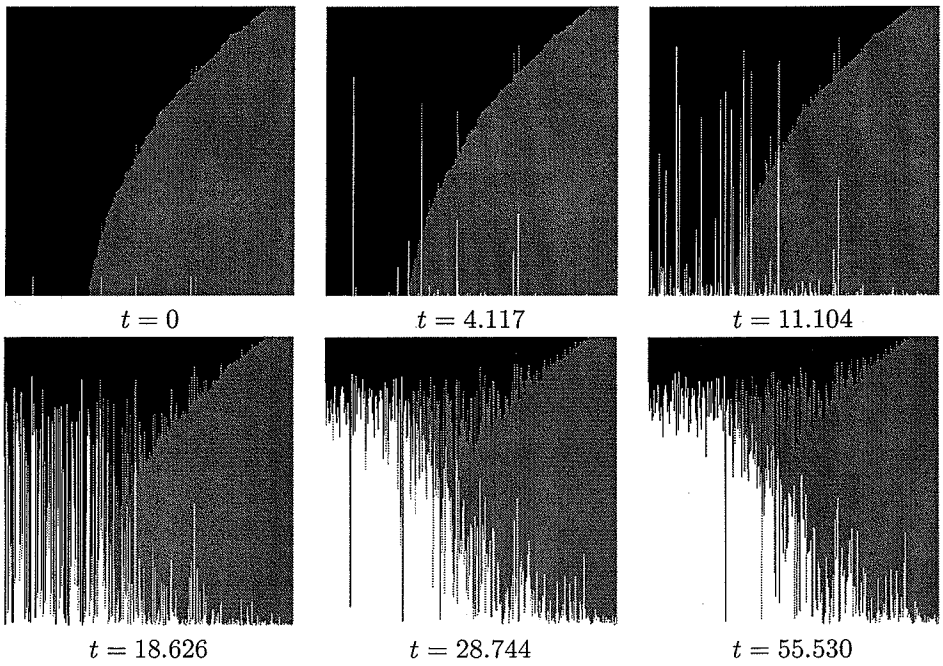


Figure 2: A sequence of images displaying the entire course of a stochastic simulation over time. The simulation consisted of 300 households, each with 300 individuals. Initial conditions were $\mathcal{U}_0 = 0.5$ and $Q = 0.8$, with 5 randomly-selected households each having 20 infectious individuals. Parameter values were $\phi = 2$, $\mu = 1$, $\rho = 0.6$, and $\alpha = 0.05$. Within each image, each column represents a household. Within a household, individuals (arranged from top to bottom) are: black = unvaccinated; dark gray = vaccinated; light gray = infectious; white = recovered. Note that the left-right order of the households has been sorted according to the number of vaccinated individuals initially present. At the end of the simulation, the number of resistant individuals who have been infected in a given household is inversely statistically related to the initial number vaccinated there.

4.2 Comparing Numerical Solutions and Stochastic Simulations

Our next result in Figure 3 demonstrates the connection between the original deterministic model given by Equations (1-4) and the stochastic simulation of the system as a continuous-time Poisson process with the time between events being exponentially distributed. The numerical solution for the specified parameter values predicts a larger and faster epidemic, whereas the simulation realizes a smaller and slower epidemic. The final proportion of individuals in each class is the same in both the numerical solutions and the stochastic simulations. Clearly there are differences in the processes, but since they predict the same initial and final values, we will consider them equivalent for our purposes.

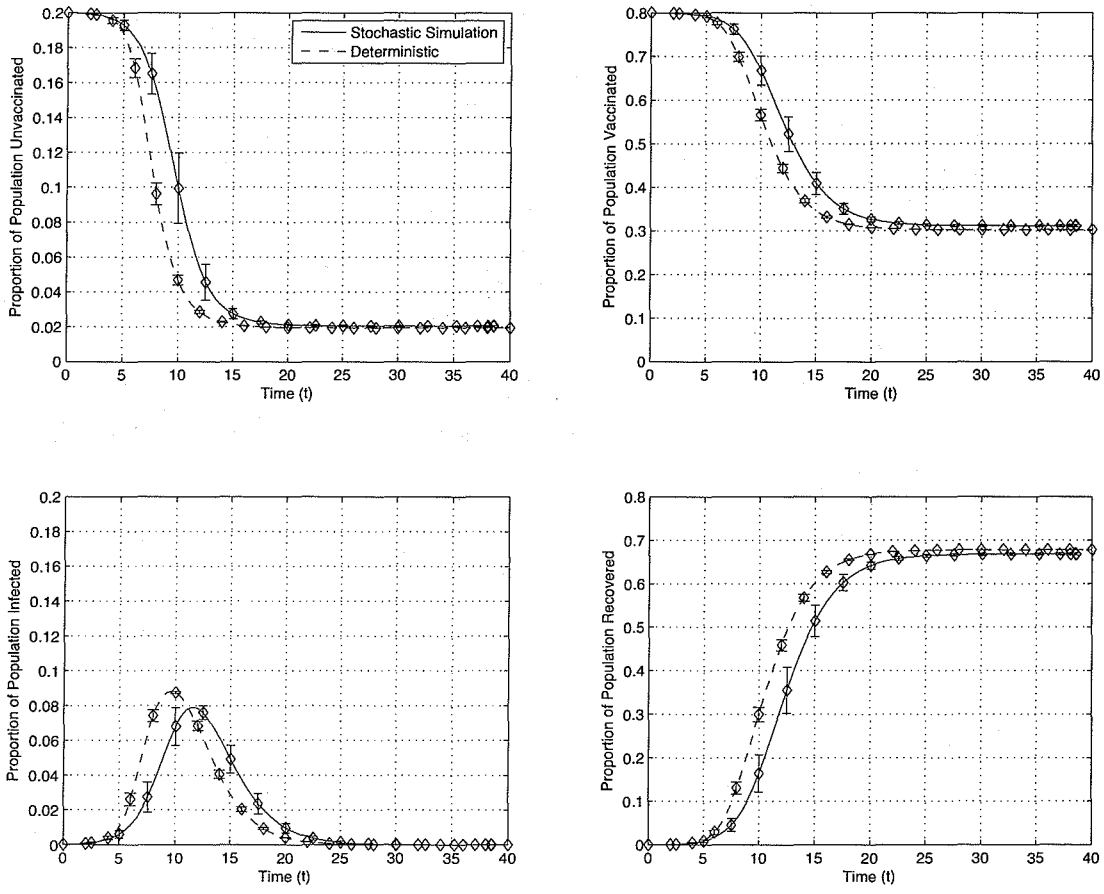


Figure 3: The above plots contain the resulting curves of both the numerical solution of the system of household equations and the corresponding stochastic simulations. The curves were obtained using 400 households, each with 400 individuals. Initial conditions were $\mathcal{U}_0 = 0.5$ and $Q = 0.8$, with five randomly-selected households each having five infected individuals. The epidemic parameter values were $\alpha = 0.1$, $\phi = 3$, $\mu = 1$, and $\rho = 0.5$. Each subplot contains the curve of one of the state variables E[U],E[V],E[I], or E[R]. Each curve is the resulting average of twenty runs of the simulation or numerical solutions exclusively. Error bars were plotted showing $\pm$ one standard deviation of the runs. The differences between the simulations and the numerical solutions is most notable in their transitory behavior during the epidemic. The simulations lag the numerical solutions, but when the epidemic does die out, both systems' state variables reach the same final values. The reason for this time difference is believed to be due to spatial correlations seen in the stochastic simulation, which are absence in the numerical solution, but this was not investigated further because the beginning and final behavior of the system is currently the only thing of interest.

4.3 The Effects of Varying Q (Stochastic Simulations)

In Figure 4, we can see the effect of changing the clustering for a single set of initial parameter values, as predicted by repeated simulation. Increasing clustering results in a non-trivial increase in the final size of the epidemic. Also, the results show that for moderate levels of clustering there is tremendous variability. This implies that the epidemic's outcome is sensitive to variations in the stochastic initial conditions within the middle range of Q values.

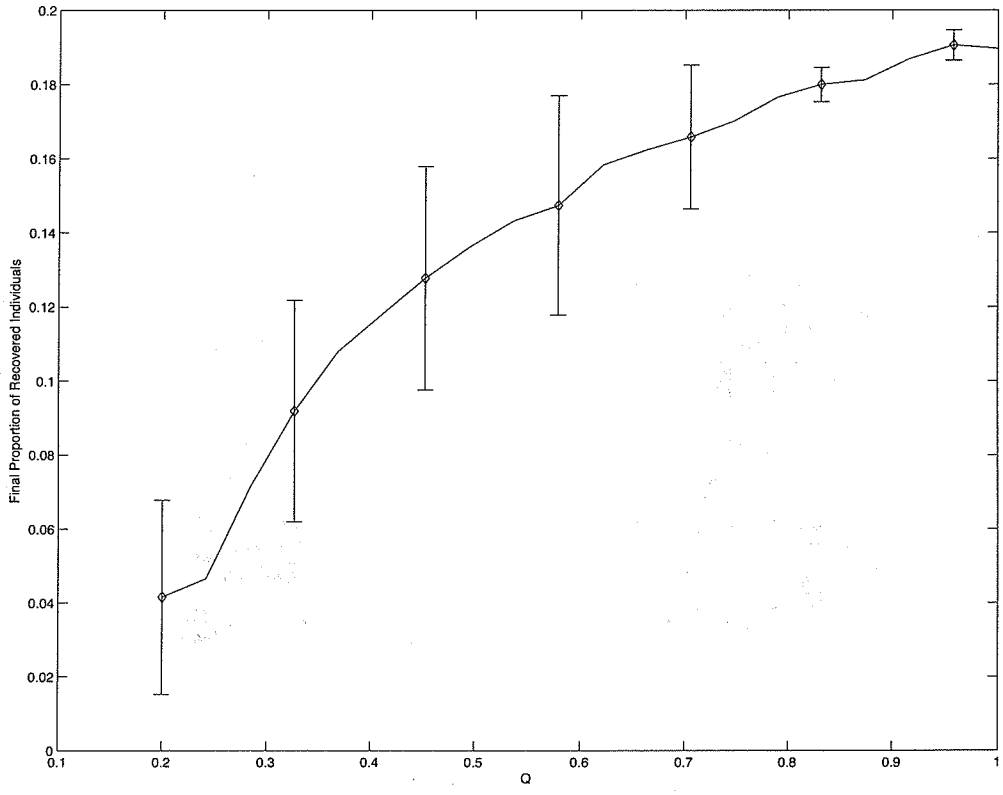


Figure 4: The final proportion of recovered individuals after the epidemic occurs as the amount of clustering is varied from its minimum value to its maximum value. The curve was generated from stochastic simulations using 400 households, each with 400 individuals. Initial conditions were $\mathcal{U}_0 = 0.2$ with five randomly-selected households containing 5 infected individuals each. Parameters were $\phi = 3$, $\mu = 1$, $\rho = 0.5$, and $\alpha = 0.1$. The amount of clustering Q varied from 0.2 to 1.0, its minimum and maximum values. Each data point is the average of 40 simulations and the errorbars show $\pm$ one standard deviation of the sampled simulations for each data point. It is clear that as clustering increases the final proportion of recovered increases, therefore clustering does cause larger epidemics. It should also be noticed that there is much more viability in the results for small values of Q, this is probably due to the increase in variability in the stochastic initial conditions for those small values of Q.

4.4 Stochastically Verifying First and Second-Generation R_0 Equations

Using a Monte Carlo simulation, we were able to repeatedly infect a susceptible population with a single infected individual and count the number of secondary and tertiary infections which that individual caused in his infectious lifetime. By sampling many populations, as displayed in Figure 5, we can see that averaging values over repeated simulations produces estimates of the first and second-generation R_0 values that almost exactly match those predicted by Equations (16) and (21), respectively.

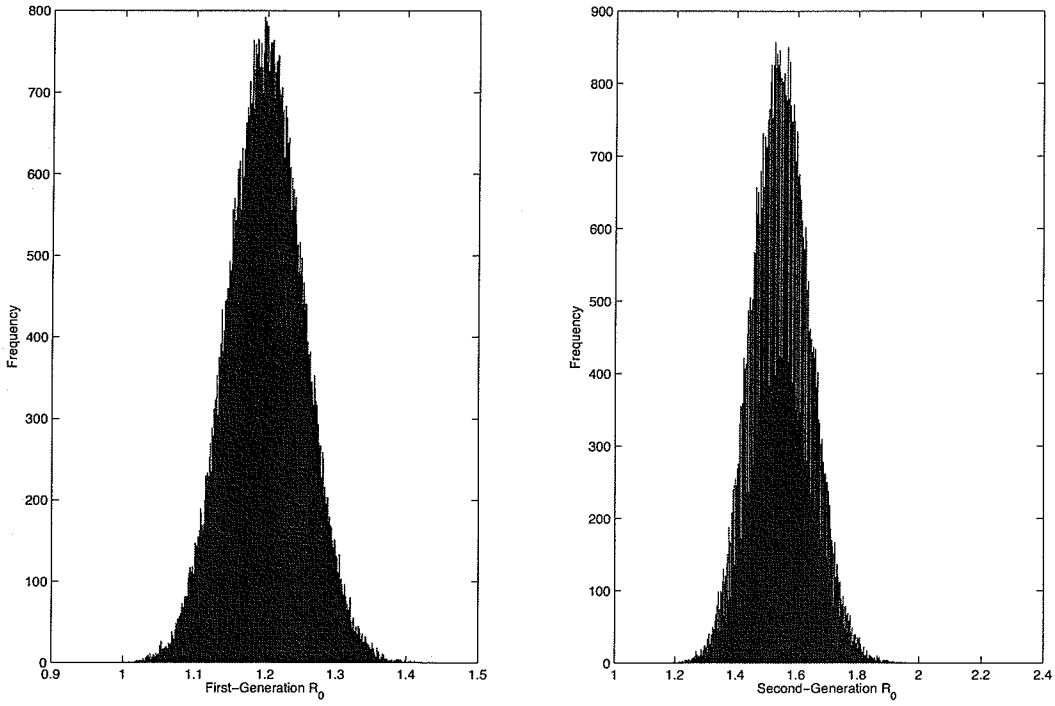


Figure 5: First and second generation R_0 values were stochastically simulated in a system of 200 households, each with 200 individuals, as follows: a system of households was generated with initial proportion of unvaccinated individuals = 0.2, Q = 0.8, and epidemic parameters α=0.1, $\phi = 2$, $\mu = 1$, and $\rho = 0.5$. One individual was randomly selected and infected, and the number of secondary and tertiary infections caused by the infected individual was counted. All individuals were reset to susceptible, and this process was repeated an additional 999 times. The average number of secondary and tertiary infections was then computed for the 1000 individuals. This process of sampling was repeated for 100,000 randomly generated systems of households, and the histograms of the data are plotted above. For the first-generation R_0, the simulated mean was 1.2001 with variance of 0.0028; Equation (16) predicts a value of 1.2. For the second-generation R_0, our simulations produced a mean of 1.5484 with a variance of 0.0093, and deterministically, Equation (21) predicts a value of 1.5484. Thus, it is quite clear that the derivation of the first and second-generation R_0 values from our system of household equations are the true expectation of an epidemic's R_0 values averaged over all stochastic realizations.

4.5 Threshold Behavior for First and Second-Generation R_0

Epidemic theory suggests that R_0 is one of the most important values in assessing epidemic dynamics of a population. "In the stochastic model, the epidemic never takes off if $R_0 \leq 1$ and it takes off with a non-zero probability if $R_0 > 1$" [38]. However, the second-generation R_0 is at least as important to consider in a household setting with clustering. To illustrate the importance of $R_0^{(2)}$, stochastic simulations were performed using predetermined parameter values such that:

i) first-generation $R_0 < 1$, second-generation $R_0 < 1$

ii) first-generation $R_0 < 1$, second-generation $R_0 > 1$

iii) first-generation $R_0 > 1$, second-generation $R_0 > 1$

Working with Equations (16) and (21), it is obvious that there are many combinations of ϕ, μ, ρ, α, $\mathcal{U}_0$, and Q that will result in first and second-generation R_0 values in the above desired ranges. For this reason, the biologically-feasible parameter values in Table 1 were chosen.

$\phi = 2$	$\mu = 1$	$\rho = 0.8$
$\alpha = 0.1$	$\mathcal{U}_0 = 0.2$	$Q = 0.8$

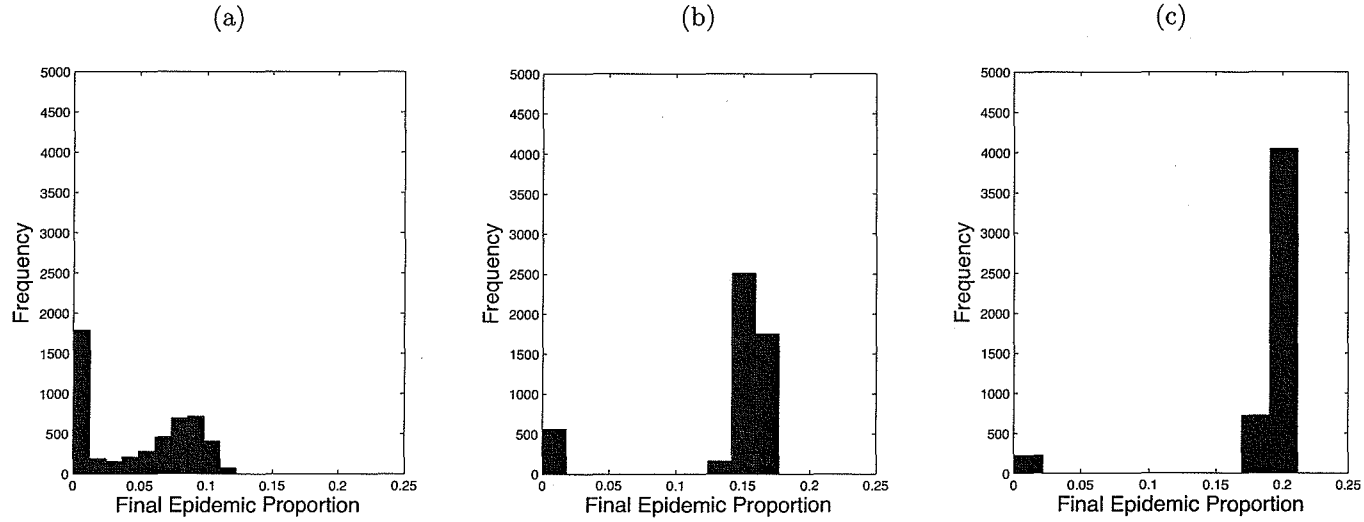
Table 1: Biologically feasible epidemic parameters used as a starting point in the investigation of threshold behavior.

One parameter was then changed while the others were held constant to yield first and second-generation R_0 in ranges i-iii above. The parameter values that were changed and the resulting first and second-generation R_0 values from Equations (16) and (21) are given in Table 2.

Parameter Changed	New Value	First-Gen R_0	Second-Gen R_0
ϕ	2.00	0.72	0.7949
ϕ	2.51	0.9036	1.2521
ϕ	3.00	1.08	1.7885
ρ	0.88	0.5920	0.6850
ρ	0.66	0.9440	1.0793
ρ	0.44	1.2960	1.7633
$\mathcal{U}_0$	0.15	0.64	0.6343
$\mathcal{U}_0$	0.30	.88	1.12
$\mathcal{U}_0$	0.45	1.12	1.6173

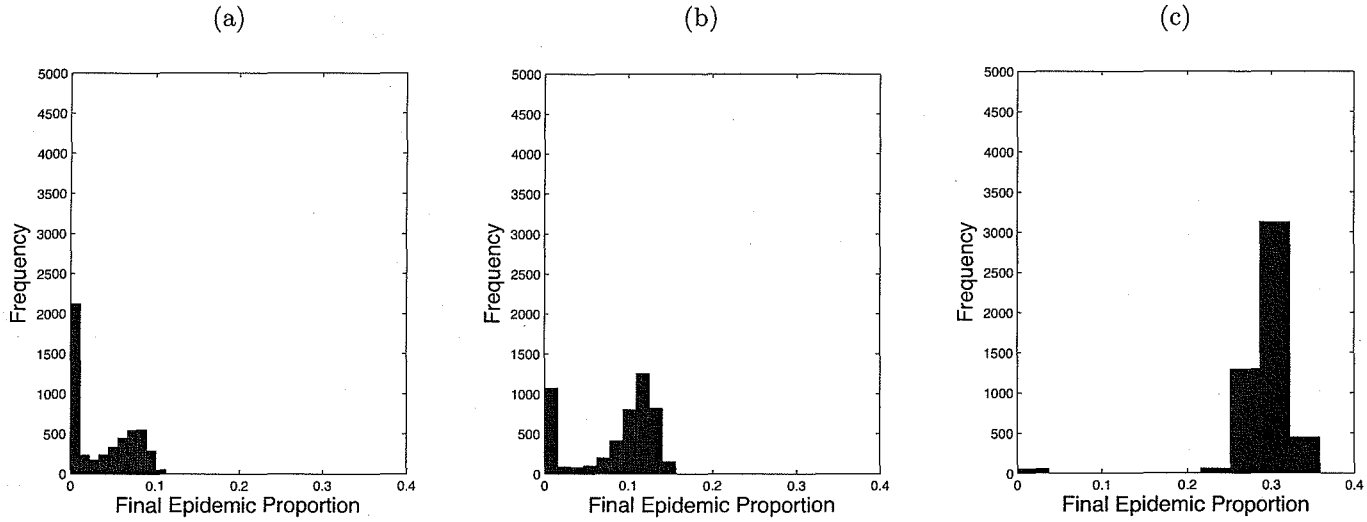
Table 2: Changes made to Table 1 parameters values and the resulting first and second-generation R_0 values used to investigate epidemic behavior near threshold parameter values. It should be clear that for each parameter, the first R_0 values satisfy condition i, the second R_0 values satisfy condition ii, and the third R_0 values satisfy condition iii.

For each of the above parameter values, 5000 full simulations were run in a system with $n_1 = n_2 = 200$ and with 5 randomly-selected households each containing 5 randomly selected individuals. For each simulation, the proportion of the total population in the recovered class (i.e. the proportion of people affected by the epidemic) was recorded. The resulting histograms of these simulations are given in Figures 6-8. For condition i, the systems behave as we would expect: most of the epidemics do not progress to a substantial size. Condition iii also behaves according to epidemic theory, with a majority of the epidemics increasing to a substantial size. The interesting behavior occurs for condition ii; even though first-generation R_0 is less than 1, a majority of epidemics reach a fairly substantial size. These simulations reveal the importance of considering the second-generation R_0 value in a system of households with clusters of unvaccinated individuals.



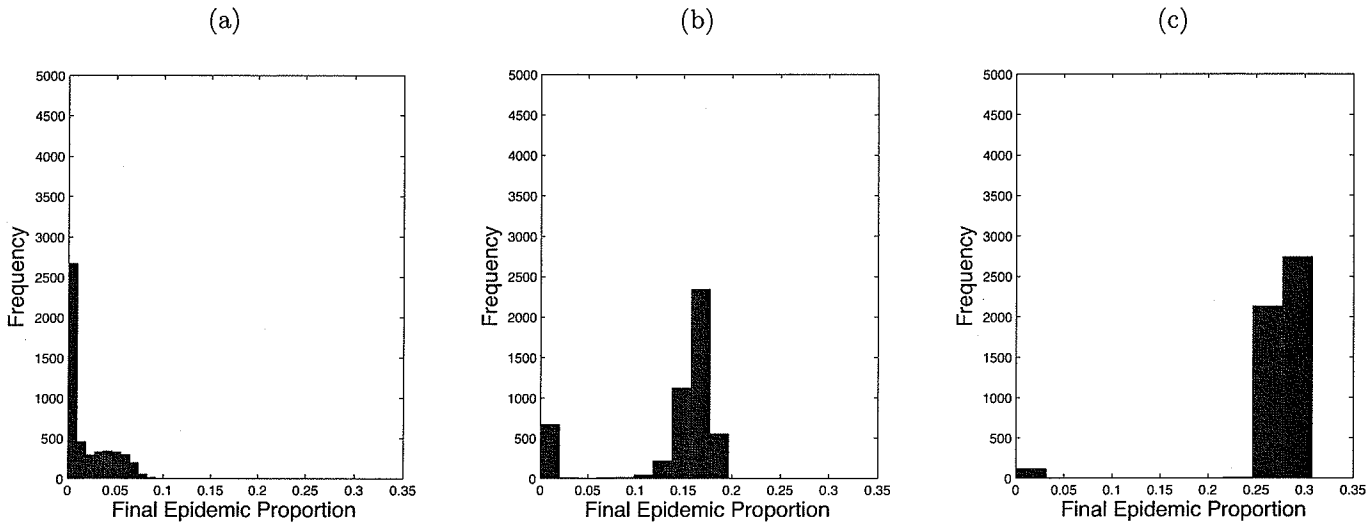
ϕ=2, $R_0^{(1)} = 0.72$, $R_0^{(2)} = 0.7949$ ϕ=2.51, $R_0^{(1)} = 0.9036$, $R_0^{(2)} = 1.2521$ ϕ=3, $R_0^{(1)} = 1.08$, $R_0^{(2)} = 1.7885$

Figure 6: The outcomes of 5000 full epidemic simulations in a system of 200 households with 200 individuals, with 5 randomly-selected households each containing 5 randomly-selected individuals. **a)** ϕ=2, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. As predicted by epidemic theory, a majority of epidemics failed to grow substantially large. **b)** ϕ=2.51, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. Even though first-generation $R_0 < 1$, most of the epidemics grow substantially large. **c)** ϕ=3, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. As predicted by epidemic theory, most epidemics grow quite large.



ρ=0.88, $R_0^{(1)} = 0.5920$, $R_0^{(2)} = 0.6850$ ρ=0.66, $R_0^{(1)} = 0.9440$, $R_0^{(2)} = 1.0793$ ρ=0.44, $R_0^{(1)} = 1.2960$, $R_0^{(2)} = 1.7633$

Figure 7: The outcomes of 5000 full epidemic simulations in a system of 200 households with 200 individuals, with 5 randomly-selected households each containing 5 randomly-selected individuals. **a)** ϕ=2, μ=1, ρ=0.88, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. As predicted by epidemic theory, a majority of epidemics failed to grow substantially large. **b)** ϕ=2, μ=1, ρ=0.66, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. Even though first-generation $R_0 < 1$, a majority of the epidemics grow substantially large. **c)** ϕ=2, μ=1, ρ=0.44, α=0.1, $\mathcal{U}_0$=0.2, Q=0.8. As predicted by epidemic theory, most epidemics grow quite large.



$\mathcal{U}_0$=.15, $R_0^{(1)} = 0.64$, $R_0^{(2)} = 0.6343$ $\mathcal{U}_0$=.30, $R_0^{(1)} = 0.8800$, $R_0^{(2)} = 1.12$ $\mathcal{U}_0$=.45, $R_0^{(1)} = 1.12$, $R_0^{(2)} = 1.6173$

Figure 8: The outcomes of 5000 full epidemic simulations in a system of 200 households with 200 individuals, with 5 randomly-selected households each containing 5 randomly-selected individuals. **a)** ϕ=2, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.15, Q=0.8. As predicted by epidemic theory, a majority of epidemics failed to grow substantially large. **b)** ϕ=2, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.30, Q=0.8. Even though first-generation $R_0 < 1$, a majority of the epidemics grow substantially large. **c)** ϕ=2, μ=1, ρ=0.8, α=0.1, $\mathcal{U}_0$=0.45, Q=0.8. As predicted by epidemic theory, most epidemics grow quite large.

4.6 Second-Generation R_0 Sensitivity Analysis

Given the expression for $R_0^{(2)}$ in Equation (21), it is instructive to verify that changing various epidemic parameters affects the second-generation R_0 in the expected way. We do so by taking the partial derivative with respect to the epidemic parameters of interest and qualitatively analyzing the resulting behavior.

$$\frac{\partial R_0^{(2)}}{\partial \phi} = 2\left(\frac{\phi}{\mu^2}\right)\left((\rho\mathcal{U}_0 + (1-\rho))^2 + (1-\alpha)\mathcal{U}_0\rho^2(Q-\mathcal{U}_0)\right) \geq 0$$

Since ϕ is present only in the first term, $R_0^{(2)}$ and ϕ are positively correlated. So as infection rates increase, we would expect the average number of tertiary infections to increase.

$$\frac{\partial R_0^{(2)}}{\partial \mu} = -2\left(\frac{\phi^2}{\mu^3}\right)\left\{(\rho\mathcal{U}_0 + (1-\rho))^2 + (1-\alpha)\mathcal{U}_0\rho^2(Q-\mathcal{U}_0)\right\}$$

With the only μ term in the denominator, it is clear that increasing recovery rate will cause $R_0^{(2)}$ to decrease and vice versa.

$$\frac{\partial R_0^{(2)}}{\partial \alpha} = \rho^2\left(\frac{\phi}{\mu}\right)^2 \mathcal{U}_0(\mathcal{U}_0 - Q) = -\left(\frac{\rho\phi}{\mu}\right)^2 \mathrm{Var}(U) \leq 0$$

For positive parameter values, this partial derivative is always negative for $Q > \mathcal{U}_0$, which implies that α and $R_0^{(2)}$ are negatively correlated. If no clustering is present in the unvaccinated population, then $Var(U(0)) = 0$ and α would not effect the value of $R_0^{(2)}$ because $\frac{\partial R_0^{(2)}}{\partial \alpha} = 0$.

$$\frac{\partial R_0^{(2)}}{\partial \rho} = \left(\frac{\phi}{\mu}\right)^2\left(-2\rho[\mathcal{U}_0{}^2 - 2\mathcal{U}_0 + 1 + \mathcal{U}_0 Q(1-\alpha) + \alpha\mathcal{U}_0{}^2] + 2\mathcal{U}_0 - 2\right)$$

The only term involving ρ is negative, so as the vaccine becomes more effective, $R_0^{(2)}$ decreases.

$$\frac{\partial R_0^{(2)}}{\partial \mathcal{U}_0} = \rho \left(\frac{\phi}{\mu}\right)^2 \left(2(1-\rho) + \rho Q(1-\alpha) + \rho\alpha\mathcal{U}_0\right)$$

This partial derivative is always non-negative, so $\mathcal{U}_0$ and $R_0^{(2)}$ are positively correlated. This makes sense, since we would expect the potential for epidemic to be greater in a population with a greater initial proportion of individuals unvaccinated.

$$\frac{\partial R_0^{(2)}}{\partial Q} = \left(\frac{\phi}{\mu}\right)^2 \left(\mathcal{U}_0\rho^2(1-\alpha)\right) \tag{22}$$

This partial derivative is always non-negative, so clustering and $R_0^{(2)}$ are directly proportional. This finding agrees with the trend seen in Figure 4 and demonstrates analytically that increased clustering leads to greater second-generation R_0 values, which in turn create larger and more probable epidemics.

4.7 Numerically Solving the Moment Equations

All attempts to numerically solve the system of moment equations failed. The exact reason why the system is so numerically unstable is not precisely known. An investigation of why the equations failed numerically was not undertaken because it is not necessary to answer our question about clustered unvaccinated individuals. We believe, however, that the first problem with the system is that when it is solved, some of the moments become negative and other moments become divergent, tending towards either positive or negative infinity. These degenerative behaviors, which are biologically unfeasible, are probably caused by the fact that at times, the higher order moments, being many orders of magnitude smaller than lower order moments, tend towards zero in such a way that our numerical method will overstep the moments' proper bounds.

In previous work [42], drastic measures were taken to ensure that the system of moment equations could be numerically solved by reparameterizing the equations so that each of the moments were kept within their correct bounds on the interval [0,1]. However, this was not seen as worthwhile in order to answer our question about clustering. It would be interesting to continue working with the moment equations because they pose an interesting view of the system. If the moment equations could be made to be well-behaved, then they could be integral to understanding the dynamic of spatial correlation as they change through the course of an epidemic.

5 Discussion

We had hypothesized that groups of unvaccinated individuals within a contained geographic or social population could serve as a likely starting point for an outbreak and significantly affect epidemiological dynamics, even while the vaccination rate of the greater population remains constant. Our system of household equations describes a simplified system of such a partitioned population, and our model represents a simple generic epidemic.

Taking our basic model of household equations, we developed a stochastic simulation for a more realistic representation of the theoretical epidemic under consideration. In Figure 3, we can see that the stochastic simulation and the numerical solution of the household equations vary in their transient behavior, but their initial and final states are equal to one another. Since our model is not fitted to specific data, we assume that the stochastic simulations are a more realistic interpretation of the epidemic because interactions between households and individuals occur as random events instead of as deterministic flows. Also, the simulations allow us to consider populations which are smaller than those that might be necessary to satisfy all of the assumptions required for a continuous differential equation model.

Using the stochastic simulations, it was straightforward to simulate systems of households with identical epidemic parameters and initial vaccination coverage levels while varying the amount of clustering. As is shown in Figure 4, clustering does increase the size of the recovered class at the end of epidemic. An exhaustive search of the parameter space could have been carried out to understand the effect of clustering in all possible cases. Instead, a more general result came from reverting back to the household equations and analytically deriving both the first and second-generation R_0 equations (Equations (16) and (21)). These two expressions are tools that allow us to measure the secondary and tertiary impact of a disease prescribed by a specific infection rate, recovery rate, vaccine efficacy, proportions of long-term and short-term interactions, and the level of clustering of unvaccinated individuals.

The first step in exploring the importance of the first and second-generation R_0 equations with respect to our system was to stochastically simulate both R_0 values using a Monte Carlo method. We found that by taking a sufficiently large number of samples, our simulation's computed R_0 values converged to the predicted R_0 values that we get from Equations (16) and (21). Figure 5 clearly shows that the average of many realizations of the

A Derivations of Moment Equations

The second-order moment equations given in Section 2.4 are derived as follows.

$$\begin{aligned}
\frac{d}{dt}E[UI] =& E\Big[\frac{d}{dt}UI\Big] \\
=& E\Big[I\frac{dU}{dt} + U\frac{dI}{dt}\Big] \\
=& -\phi(1-\alpha)E[I^2U] - \phi\alpha E[I]E[UI] + \phi(1-\alpha)(E[U^2I] + (1-\rho)E[UIV]) \\
&+ \phi\alpha E[I](E[U^2] + (1-\rho)E[UV]) - \mu E[UI]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}E[VI] =& E\Big[\frac{d}{dt}VI\Big] \\
=& E\Big[I\frac{dV}{dt} + V\frac{dI}{dt}\Big] \\
=& -(1-\rho)\phi(1-\alpha)E[I^2V] - \phi\alpha E[I]E[VI] + \phi(1-\alpha)(E[UIV] + (1-\rho)E[V^2I]) \\
&+ \phi\alpha E[I](E[UV] + (1-\rho)E[V^2]) - \mu E[VI]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}E[UV] =& E\Big[\frac{d}{dt}UV\Big] \\
=& E\Big[U\frac{dV}{dt} + V\frac{dU}{dt}\Big] \\
=& -(1-\rho)\phi(1-\alpha)E[UIV] - \phi\alpha E[I]E[UV] - \phi(1-\alpha)E[UIV] - \phi\alpha E[I]E[UV]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}E[U^2] =& 2E\Big[U\frac{dU}{dt}\Big] \\
=& 2E[-\phi(1-\alpha)IUU - \phi\alpha E[I]UU] \\
=& -\phi(1-\alpha)E[U^2I] - \phi\alpha E[I]E[U^2]
\end{aligned}$$

$$\begin{aligned}
\frac{d}{dt}E[V^2] =& 2E\Big[V\frac{dV}{dt}\Big] \\
=& -2(1-\rho)\phi\left((1-\alpha)E[V^2I] + \alpha E[I]E[V^2]\right)
\end{aligned}$$

B Derivations of Moment Closure Approximations

The moment-closure approximations given in Section 2.4 are below, and the following conventions are used: E[XYZ] is the expectation of choosing three individuals from one household, one being of type X, one of type Y, and one of type Z. In the probabilities, each $*$ is a place holder and denotes choosing either an Unvaccinated individual (U), a vaccinated individual (V), or an infected individual (I) without preference. In this way, for example, $P(**V|IU*)$ reads as "the probability you choose a V from a particular household given that you have already selected one I and one V from that same household". In each approximation we can assume that $P(X**|*YZ) \approx P(X**|*Y*)$ since there is little difference between the probability of choosing an X individual given a Y and a Z individual from the same household and the probability of choosing an X individual given a Y individual and another unspecified indivdual from the same household. To make this assumption, we must first assume two proportions to be independent for each distinct approximations. These proportional assumptions can be seen below with their corresponding approximations.

$$\begin{aligned} E[IUV] =& P(IUV) = P(IU*)P(**V|IU*) \\ \approx& P(IU*)P(**V|*U*) \\ \approx& \frac{P(IU)P(UV)}{P(U)} = \frac{E[IU]E[UV]}{E[U]} \end{aligned}$$

It is assumed that that I and V are independent.

$$\begin{aligned} E[U^2I] =& P(UUI) \approx P(UU*)P(**I|*U*) \\ \approx& \frac{P(UU)P(IU)}{P(U)} = \frac{E[UU]E[IU]}{E[U]} \end{aligned}$$

It is assumed that I and U are independent.

$$\begin{aligned} E[I^2V] =& P(IIV) \approx P(I*V)P(*I*|**V) \\ \approx& \frac{P(IV)P(IV)}{P(V)} = \frac{E[IV]^2}{E[V]} \end{aligned}$$

It is assumed that I and I are independent.

$$\begin{aligned} E[V^2I] =& P(VVI) \approx P(*VI)P(V**|**I) \\ \approx& \frac{P(VI)P(IV)}{P(I)} = \frac{P(VI)^2}{P(I)} = \frac{E[VI]^2}{E[I]} \end{aligned}$$

It is assumed that V and V are independent.

$$\begin{aligned} E[I^2U] =& P(IIU) \approx P(I*U)P(*I*|**U) \\ \approx& \frac{P(IU)P(IU)}{P(U)} = \frac{E[IU]^2}{E[U]} \end{aligned}$$

It is assumed that I and I are independent.

References

[1] Centers for Disease Control and Prevention. Achievements in Public Health, 1990-1999: Impact of Vaccines Universally Recommended for Children – United States, 1990-1998. *MMWR Morb Mortal Wkly Rep*, 48:241–248, 1999.

[2] Centers for Disease Control and Prevention. Ten Great Public Health Achievements - United States, 1900-1999. *MMWR Morb Mortal Wkly Rep*, 48:241–243, 1999.

[3] SL Plotkin and SA Plotkin. *Vaccines*, chapter 1, pages 1–12. W.B. Saunders Company, 3 edition, 1999.

[4] AB Coffield, MV Maciosek, JM McGinnis, JR Harris, MB Caldwell, SM Teutsch, D Atkins, JH Richland, and A Haddix. Priorities among recommended clinical preventive services. *American Journal of Preventive Medicine*, 21(1):1–9, 2001.

[5] WA Orenstein, RG Douglas, LE Rodewald, and AR Hinman. Immunizations in the united states: Success, structure, and stress. *Health Affairs*, 24(3):599–610, 2005.

[6] P Davies. Antivaccination web sites. *JAMA*, 288(14):1717–1718, Oct 2002.

[7] SP Calandrillo. Vanishing vaccinations: Why are so many americans opting out of vaccinating their children? *University of Michigan Journal of Law Reform*, 37:353–440, 2004.

[8] RT Chen. *Vaccines*, chapter 49, pages 1144–1163. WB Saunders Company, 3 edition, 1999.

[9] Centers for Disease Control and Prevention. Some Common Misconceptions about vaccination and how to respond to them. `http://www.cdc.gov/vaccines/vac-gen/6mishome.htm`, Accessed July 12, 2008.

[10] YA Maldonado. Current controversies in vaccination: Vaccine safety. *JAMA*, 288(24):3155–3158, 2002.

[11] A Wakefield, S Murch, A Anthony, J Linnell, D Casson, M Malik, M Berelowitz, A Dhillon, M Thomson, and P Harvey. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, 351(9103):637–641, 1998.

[12] US Food and Drug Administration. Thimerosal in vaccines. `http://www.fda.gov/CbER/vaccine/thimerosal.htm`, Accessed July 8, 2008.

[13] E Gangarosa, A Galazka, C Wolfe, L Phillips, R Gangarosa, E Miller, and R Chen. Impact of anti-vaccine movements on pertussis control: the untold story. *The Lancet*, 351(9099):356–361, 1998.

[14] R Jacobson, P Targonski, and G Poland. A taxonomy of reasoning flaws in the anti-vaccine movement. *Vaccine*, 25(16):3146–3152, 2007.

[15] GA Poland and RM Jacobson. Understanding those who do not understand: a brief review of the anti-vaccine movement. *Vaccine*, 19:2440–2445, 2001.

[16] RM Wolfe. Content and design attributes of antivaccination web sites. *JAMA*, 287(24):3245–3248, 2002.

[17] RD Silverman. No more kidding around: Restructuring non-medical childhood immunization exemptions to ensure public health protection. *Annals of Health Law*, 12:277–294, 2003.

[18] J Thompson, S Tyson, P Cardhigginson, R Jacobs, J Wheeler, P Simpson, J Bost, K Ryan, and D Salmon. Impact of addition of philosophical exemptions on childhood immunization rates. *American Journal of Preventive Medicine*, 32(3):194–201, 2007.

[19] DA Salmon, M Haber, EJ Gangarosa, L Phillips, NJ Smith, and RT Chen. Health consequences of religious and philosophical exemptions from immunization laws: Individual and societal risk of measles. *JAMA*, 282(1):47–53, 1999.

[20] DA Salmon and AW Siegel. Religious and philosophical exemptions from vaccination requirements and lessons learned from conscientious objectors from conscription. *Public Health Reports*, 116(4):289–296, 2001.

[21] SB Omer, WKY Pan, NA Halsey, S Stokley, LH Moulton, AM Navar, M Pierce, and DA Salmon. Non-medical exemptions to school immunization requirements: Secular trends and association of state policies with pertussis incidence. *JAMA*, 296(14):1757–1763, 2006.

[22] JG Hodge. School vaccination requirements: Legal and social perspectives. *NCSL State Legislative Rep*, 27:1–14, 2002.

[23] D Khalili and A Caplan. Off the grid: Vaccinations among homeschooled children. *The Journal of Law, Medicine & Ethics*, 35(3):1073–1105, Jul 2007.

[24] T May and RD Silverman. 'Clustering of exemptions' as a collective action threat to herd immunity. *Vaccine*, 21(11-12):1048–1051, 2003.

[25] A d'Onofrio, P Manfredi, and E Salinelli. Vaccinating behaviour, information, and the dynamics of SIR vaccine preventable diseases. *Theoretical Population Biology*, 71(3):301–317, 2007.

[26] A Maayan-Metzger, P Kedemfriedrich, and J Kuint. To vaccinate or not to vaccinate - that is the question: why are some mothers opposed to giving their infants hepatitis B vaccine? *Vaccine*, 23(16):1941–1948, 2005.

[27] A Wroe, A Bhan, P Salkovskis, and H Bedford. Feeling bad about immunising our children. *Vaccine*, 23(12):1428–1433, 2005.

[28] Thomas C. Schelling. *Micromotives and Macrobehavior*. W. W. Norton and Company, 1978.

[29] WO Kermack and AG McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, 115(772):700–721, 1927.

[30] JJ Gart. The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics*, 24(3):557–566, 1968.

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

[32] H Lacayo and NA Langberg. The exact and symptotic formulas for the state probabilities in simple epidemics with m kinds of susceptibles. *Journal of Applied Probability*, 19(1):1–9, 1982.

[33] CM Kribs-Zaleta and JX Valesco-Hernández. A simple vaccination model with multiple endemic states. *Mathematical Biosciences*, 164:183–201, 2000.

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

[35] RK Watson. On an epidemic in a stratified population. *Journal of Applied Probability*, 9(3):659–666, 1972.

[36] NG Becker and K Dietz. The effect of household distribution on transmission and control of highly infectious diseases. *Mathematical Biosciences*, 127(2):207–219, 1995.

[37] F Ball, D Mollison, and G Scalia-Tomba. Epidemics with two levels of mixing. *Annals of Applied Probability*, 7(1):46–89, 1997.

[38] F Ball and P Neal. A general model for stochastic SIR epidemics with two levels of mixing. *Mathematical Biosciences*, 180:73–102, 2002.

[39] FG Ball, T Britton, and OD Lyne. Stochastic multitype epidemics in a community of households: Estimation of threshold parameter r and secure vaccination coverage. *Biometrika*, 91(2):345–362, 2004.

[40] RM Anderson RM May. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences*, 72(1):83–111, 1984.

[41] DJ Murrell, U Dieckmann, and R Law. On moment closures for population dynamics in continuous space. *Journal of Theoretical Biology*, 229:421–432, 2004.

[42] D Hiebeler. Moment equations and dynamics of a household SIS epidemiological model. *Bulliten of Mathematical Biology*, 68(6):1315–1333, 2006.

[43] NG Becker and DN Starczak. Optimal vaccination strategies for a community of households. *Mathematical Biosciences*, 139(2):117–132, 1997.

[44] DH Zanette and M Kuperman. Effects of immunization in small-world epidemics. *Physica A: Statistical Mechanics and its Applications*, 309:445–452, 2002.

[45] Z Lu, X Chi, and L Chen. The effect of constant and pulse vaccination on SIR epidemic model with horizontal and vertical transmission. *Mathematical and Computer Modelling*, 36(9):1039–1057, 2002.

[46] PA Briss, LE Rodewald, AR Hinman, AM Shefer, RA Strikas, RR Bernier, VG Carande-Kulis, HR Yusuf, SM Ndiaye, and SM Williams. Reviews of evidence regarding interventions to improve vaccination coverage in children, adolescents, and adults. *American Journal of Preventive Medicine*, 18(1S):97–140, 2000.

[47] PA Gross, AW Hermogenes, HS Sacks, J Lau, and RA Levandowski. The Efficacy of Influenza Vaccine in Elderly Persons: A Meta-Analysis and Review of the Literature. *Annals of Internal Medicine*, 123(7):518–527, 1995.

[48] MJ Fine, MA Smith, CA Carson, F Meffe, SS Sankey, LA Weissfeld, AS Detsky, and WN Kapoor. Efficacy of pneumococcal vaccination in adults. A meta-analysis of randomized controlled trials. *Archives of Internal Medicine*, 154(23):2666–2677, 1994.

[49] J Ward, J Cherry, S Chang, S Partridge, and H Lee. Efficacy of an acellular pertussis vaccine among adolescents and adults. *New England Journal of Medicine*, 353(15):1555–1563, 2005.

[50] S Black, H Shinefield, B Fireman, E Lewis, P Ray, JR Hansen, L Elvin, KM Ensor, J Hackell, and G Siber. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. *The Pediatric Infectious Disease Journal*, 19(3):187, 2000.

[51] K O'Brien, L Moulton, R Reid, and R Weatherholtz. Efficacy and safety of seven-valent conjugate pneumococcal vaccine in American Indian children: group randomised trial. *The Lancet*, 362:355–361, 2003.

[52] N Becker. Estimation for discrete time branching processes with application to epidemics. *Biometrics*, 33(3):515–522, 1977.

[53] D Hiebeler. Populations on fragmented landscapes with spatially structed heterogeneities: Landscape generation and local dispersal. *Ecology*, 81(6):1629–1641, 2000.

[54] DE Hiebeler and AK Criner. Partially mixed household epidemiological model with clustered resistant individuals. *Physical Review E*, 75(2), 2007.

[55] O Diekmann, J Heesterbeek, and J Metz. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990.

[56] K Dietz. The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research*, 2(1):23–41, 1993.

[57] J Li JM Hyman. An intuitive formulation for the reproductive number for the spread of diseases in heterogeneous populations. *Mathematical Biosciences*, 167:65–86, 2000.

[58] C Castillo-Chavez, JX Velasco-Hernandez, and S Fridman. Modeling contact structures in biology. *Frontiers of Theoretical Biology, Lecture Notes in Biomathematics*, 100:454–491, 1994.

[59] Ying-Hen Hsieh, P Driessche, and Lin Wang. Impact of travel between patches for spatial spread of disease. *Bull. Math. Biol.*, 69(4):1355–1375, May 2007.

[60] R Cohen, S Havlin, and D ben Avraham. Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24):247901:1–4, 2003.

[61] M Murray and Z Rasmussen. Measles Outbreak in a Northern Pakistani Village: Epidemiology and Vaccine Effectiveness. *American Journal of Epidemiology*, 151(8):811–819, 2000.

[62] S van den Hof, CMA Meffre, MAE Conyn van Spaendonck, F Woonink, HE de Melker, and RS van Binnendijk. Measles Outbreak in a Community with Very Low Vaccine Coverage, the Netherlands. *Emerging Infectious Diseases*, 7(3 Supplement), 2000.

[63] J Puvimanasinghe, C Arambepola, and N Abeysinghe. Measles Outbreak in Sri Lanka, 1999-2000. *The Journal of Infectious Diseases*, 2003.

[64] C Stein-Zamir, N Abramson, H Shoob, and G Zentner. An outbreak of measles in an ultra-orthodox Jewish community in Jerusalem, Israel, 2007 - An in-depth report. *Eurosurveillance*, 13:1–4, 2008.

[65] JL Richard, V Masserey-Spicher, S Santibanez, and A Mankertz. Measles outbreak in Switzerland - An update relevant for the European football championship (EURO 2008). *Eurosurveillance*, 13(8), 2008.